



universität
wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Kunst sehen, Schönheit messen: Eine mobile Eye-Tracking-Studie zur Beziehung zwischen Blickmetriken und ästhetischer Bewertung im Museumssetting

verfasst von | submitted by

Jan-Ove Wiesner BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt |
Degree programme code as it appears on the
student record sheet:

UA 066 840

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Psychologie

Betreut von | Supervisor:

Univ.-Prof. Dipl.-Psych. Dr. Helmut Leder

Zusammenfassung

Diese Masterarbeit untersucht die Beziehung zwischen Blickverhalten, Pupillenreaktionen und ästhetischen Bewertungen beim Betrachten realer Kunstwerke im Museum. Sie knüpft an Befunde an, die systematische Zusammenhänge zwischen Blickverhalten und Schönheitsbewertungen zeigen und okulomotorische sowie pupillometrische Parameter als Indikatoren ästhetischer Bewertung diskutieren. Unter Verwendung mobiler Eye-Tracking-Messungen während eines freien Rundgangs im Belvedere Museum Wien wurden zentrale Blickmetriken (Fixationsanzahl, mittlere Fixationsdauer und kumulative Betrachtungszeit) sowie der Pupillendurchmesser von 25 Studierenden erfasst und mit kontinuierlichen Schönheitsratings synchronisiert. Das Stimulusmaterial umfasste 26 Gemälde aus vier Kategorien (Individuen, Menschenmengen, Naturdarstellungen und Stilleben). Die Ergebnisse zeigen, dass Blickmetriken zwischen den Bildkategorien variieren: Soziale Motive werden länger und mit mehr Fixationen betrachtet als Stilleben, Naturdarstellungen liegen dazwischen. Bei Individuen-Darstellungen zeigt sich eine signifikant höhere mittlere Fixationsdauer als bei Natur- und Stilleben-Gemälden. Explizite Schönheitsratings lassen sich nur zu einem kleinen Teil der Varianz ($R^2 \approx .17$) durch die betrachteten Blickmetriken erklären, und der Pupillendurchmesser zeigt zwar positiv gerichtete, insgesamt aber schwache und statistisch nicht signifikante Zusammenhänge mit den Schönheitsratings. Damit deuten die Ergebnisse darauf hin, dass sich das Schönheitserleben im Museum nur eingeschränkt in Blickmetriken und Pupillenreaktionen niederschlägt und Eye-Tracking im Feld eher als ergänzendes Diagnostikum, denn als eigenständiges Maß ästhetischer Präferenz fungiert.

Abstract

This master's thesis investigates the relationship between gaze behavior, pupil responses, and aesthetic evaluations when viewing original artworks in a museum setting. It builds on findings demonstrating systematic associations between gaze behavior and beauty judgments and discusses oculomotor and pupillometric parameters as potential indicators of aesthetic evaluation. Using mobile eye tracking during a free-walking task at the Belvedere Museum Vienna, key gaze metrics (number of fixations, mean fixation duration, and total viewing time) as well as pupil diameter of 25 students were recorded and synchronized with continuous beauty ratings. The stimulus set comprised 26 paintings from four categories (individual persons, groups of people, nature scenes, still lifes). The results show systematic differences in gaze behavior across categories. Social motifs were viewed longer and with more fixations than still lifes, while nature depictions occupied an intermediate position. Individual-person depictions show a significantly longer mean fixation duration than nature scenes and still-life paintings. Explicit beauty ratings can be explained by the considered gaze metrics only to a small proportion of their variance ($R^2 \approx .17$), and pupil diameter shows consistently positive, but overall weak and statistically not significant associations with the beauty ratings. Thus, these results indicate that the experience of beauty in the museum is reflected only to a limited extent in gaze metrics and pupil responses, and that eye tracking in the field functions more as a complementary diagnostic tool than as a standalone measure of aesthetic preference.

Inhalt

1. Einleitung	4
2. Theoretischer Hintergrund	8
2.1 Bildkategorien und biologische Relevanz	8
2.1.1 Individuen	9
2.1.2 Menschenmengen	11
2.1.3 Natur	13
2.1.4 Stilleben	15
2.2 Blickmetriken als Prädiktoren ästhetischen Gefallens	16
2.3 Forschungsfragen der vorliegenden Untersuchung	19
3. Methodik	20
3.1 Studiendesign und Teilnehmer*innen	20
3.2 Stimulusmaterial und Datenaufbereitung	23
3.3 Datenauswertung	24
4. Ergebnisse	26
4.1 Bildkategoriespezifische Blickprofile	26
4.2 Blickmetriken als Prädiktoren der Schönheitsurteile	31
5. Diskussion	37
5.1 Interpretation der Ergebnisse und theoretische Einordnung	37
5.2 Die Rolle der Physiologie	39
5.3 Methodische Reflexionen und Limitationen	41
6. Fazit und Ausblick	44
Literaturverzeichnis	47
Abbildungsverzeichnis	55
Tabellenverzeichnis	56
Anhang: Stimulusmaterial	57

1. Einleitung

Wie sehr liegt Schönheit im Auge der betrachtenden Person, und in welchem Maß spiegelt sie sich im Blickverhalten wider? In psychologischen Modellen visueller ästhetischer Bewertung gilt Schönheit nicht als feste Eigenschaft eines Objekts, sondern entsteht aus den Prozessen, mit denen es wahrgenommen und beurteilt wird. Diese Modelle gehen davon aus, dass ästhetische Erfahrungen das Ergebnis mehrerer aufeinanderfolgender Verarbeitungsstufen im Betrachtenden sind. Zunächst werden grundlegende Merkmale wie Form, Farbe und Struktur analysiert. Die so gewonnenen perzeptuellen Informationen werden anschließend mit vorhandenem Wissen, Erwartungen und Bedeutungszuschreibungen verknüpft, sodass ein kohärenter Eindruck des Objekts entsteht. Auf dieser Grundlage bildet sich eine affektiv gefärbte Bewertung, die sich in Urteilen über Gefallen und Schönheit niederschlägt ([Leder et al., 2004](#); [Leder & Nadal, 2014](#)). Diese Sichtweise ist im Forschungsfeld der empirischen Ästhetik verankert, das seit Fechners programmatischen Arbeiten vor rund 150 Jahren danach fragt, wie sich ästhetische Erfahrungen mit wissenschaftlichen Methoden untersuchen lassen. Entsprechend zielt diese Forschung darauf ab, ästhetische Urteile mit experimentellen Designs und psychophysiologischen Maßen zugänglich zu machen ([Fechner, 1876](#); [Nadal & Ureña, 2021](#)).

Eye-Tracking bietet in diesem Kontext eine zentrale methodische Möglichkeit, ästhetische Erfahrungen über das Blickverhalten zu erschließen. Durch die objektive Erfassung von Blickverhalten macht es sichtbar, wohin sich Personen wenden, wie lange sie bestimmte Bildbereiche betrachten und wie sich ihre visuelle Aufmerksamkeit über ein Kunstwerk verteilt. Auf diese Weise werden jene Aufmerksamkeitsprozesse erfasst, die ästhetische Eindrücke vorbereiten und begleiten. Blickverhalten lässt sich dazu als Abfolge von Fixationen und Sakkaden beschreiben. Fixationen sind kurze Phasen stabilen Hinsehens, Sakkaden schnelle Blicksprünge zwischen zwei Fixationen. Die Analyse dieser Abfolgen erlaubt Rückschlüsse auf

zugrunde liegende Verarbeitungsprozesse, etwa darauf, welche Elemente zunächst ins Auge fallen, wie Informationen integriert werden und wo sich die Betrachtung verdichtet.

Blickverhalten wird damit zu einer indirekten, aber aussagekräftigen Spur dafür, wie sich ästhetische Bewertung im Verlauf der Betrachtung entfaltet. Bereits die klassischen Arbeiten von Buswell (1935, zitiert nach [Rosenberg & Klein, 2015](#)) und Yarbus (1967, zitiert nach [Rosenberg & Klein, 2015](#)) zeigen, dass sowohl Aufgabenstellung als auch Bildbedeutung die Abfolge der Blickbewegungen systematisch beeinflussen und begründen damit die psychologische Forschung zu Blickbewegungen.

Zahlreiche Befunde legen nahe, dass ästhetisch bevorzugte Stimuli mit einer längeren Betrachtungsdauer, mehr Fixationen und einer intensiveren visuellen Exploration einhergehen. So werden attraktiver bewertete oder stärker gemochte Gesichter länger betrachtet und erhalten mehr Fixationen ([Faust et al., 2019](#); [Jankowski et al., 2020](#); [Leder et al., 2010](#); [Mitrovic et al., 2016](#)). Ein vergleichbares Muster erhöhter Betrachtungsdauer und Fixationsanzahl bei stärkerem Gefallen zeigt sich auch für Kunstwerke und andere menschengeschaffene Objekte ([Brieber et al., 2014](#); [Mitschke et al., 2017](#)). Ähnliche Zusammenhänge wurden für natürliche Szenen und Landschaften beschrieben, in denen höhere Präferenzurteile mit längerer Betrachtungsdauer und vermehrter visueller Exploration einhergehen ([Franěk et al., 2019](#); [Jacobs, 2011](#); [Valtchanov & Ellard, 2015](#)). Studien, die Gesichts- und Kunststimuli in einem gemeinsamen Versuchsaufbau vergleichen, finden klare Zusammenhänge zwischen Liking-Ratings und Blickmaßen wie Gesamtfixationsdauer und Fixationsanzahl ([Holmes & Zanker, 2012](#); [Goller et al., 2019](#)). Darüber hinaus zeigen neuere Arbeiten, dass sich qualitative Aspekte des Kunsterlebens, etwa unterschiedliche Grade ästhetischer Distanz bzw. Involviertheit, in charakteristischen Blickmustern niederschlagen ([Marin & Leder, 2022](#)).

Ergänzend zum Blickverhalten werden Pupillenreaktionen als physiologische Marker autonomer Aktivierung betrachtet, die in Laborstudien wiederholt mit der Verarbeitung affektiv bedeutsamer und ästhetisch ansprechender Stimuli in Verbindung gebracht wurden. So zeigen Studien mit emotionalen Bildern, dass stärkere affektive Bedeutsamkeit mit stärkeren Pupillenerweiterungen einhergeht ([Bradley et al., 2008](#); [Henderson et al., 2018](#)). Vor diesem Hintergrund wird der Pupillendurchmesser in der vorliegenden Arbeit als zusätzlicher Indikator des ästhetischen Erlebens herangezogen. Physiologisch spiegelt der Pupillendurchmesser zunächst die Anpassung an die Lichtsituation wider. Im Rahmen der vorliegenden Studie wird er jedoch gezielt als Indikator ästhetischen Erlebens genutzt, indem sein Zusammenhang mit Schönheitsratings und den verschiedenen Bildkategorien im realen Museumsraum untersucht wird.

Auch speziell im Kunstkontext wird beschrieben, dass weniger abstrakte, leichter zu verarbeitende Gemälde tendenziell höhere Gefallensurteile und größere Pupillenerweiterungen hervorrufen, was als Hinweis auf einen Zusammenhang zwischen pupillären Reaktionen und positivem ästhetischem Erleben interpretiert wurde ([Kuchinke et al., 2009](#)).

Offen ist jedoch, inwieweit sich diese Effekte, die überwiegend in Laborstudien mit hochkontrollierten Stimuli gezeigt wurden, in reale Betrachtungssituationen wie Museumsbesuche übertragen lassen. Studien, die dieselben Kunstwerke im Labor und im Museum vergleichen, zeigen, dass Originale im Museum länger betrachtet, emotional intensiver und stärker bedeutungsbezogen erlebt und besser erinnert werden als Reproduktionen im Labor ([Brieber et al., 2014](#); [Specker et al., 2017](#)). Theoretische Arbeiten zum Kunsterleben im Museum heben hervor, dass Laborstudien mit kurz präsentierten Reproduktionen in stark standardisierten Settings arbeiten, während Museumsbesuche selbstgesteuerte, sozial eingebettete Interaktionen

mit originalen, räumlich inszenierten Kunstwerken darstellen, die andere Wahrnehmungs-, Bewertungs- und Emotionsprozesse anstoßen ([Pelowski et al., 2017](#)).

Damit zeigen sich systematische Unterschiede in Betrachtungsdauer, Bewegungsverhalten im Raum und subjektivem Erleben zwischen Labor- und Museumssituationen. Dies wirft die Frage auf, in welchem Ausmaß sich etablierte Zusammenhänge zwischen Schönheitserleben, Blickverhalten und Pupillenreaktionen unter realen Bedingungen wiederfinden. Mobiles Eye-Tracking bietet für diese Frage ein wenig intrusives Verfahren. Es ermöglicht die Erfassung prozessbezogener Parameter im Museumsraum, ohne die Situation wesentlich zu verändern, und verbindet so ökologische Validität mit prozessnaher Messung ([Pelowski et al., 2018](#); [Santini et al., 2018](#)).

Die vorliegende Arbeit nutzt diese Möglichkeit, um das Schönheitserleben im realen Museumsraum zu untersuchen, das über explizite Schönheitsratings der betrachteten Gemälde erfasst wird. Im Zentrum stehen zwei Fragen, die in den folgenden Kapiteln vertieft werden: Erstens, wie sich Blickverhalten und Pupillenreaktionen beim Betrachten realer Gemälde in Abhängigkeit von der Bildkategorie unterscheiden. Zweitens, in welchem Ausmaß diese Maße mit expliziten Schönheitsratings zusammenhängen und zu deren Vorhersage in multivariaten Modellen beitragen. Untersucht werden vier Bildkategorien, die sich hinsichtlich sozialer und biologischer Relevanz sowie struktureller Komplexität unterscheiden: Darstellungen einzelner Personen (Individuen), Szenen mit Personengruppen (Menschenmengen), naturbezogene Darstellungen (Natur) und Stilleben mit menschengemachten Objekten (Stilleben). Das anschließende Kapitel erläutert die theoretischen Grundlagen zu diesen Bildkategorien und den eingesetzten Blick- und Pupillenmaßen und leitet daraus die Forschungsfragen der vorliegenden Untersuchung ab.

2. Theoretischer Hintergrund

2.1 Bildkategorien und biologische Relevanz

In dieser Arbeit bezeichnen Bildkategorien inhaltlich und strukturell abgegrenzte Gruppen von Gemäldemotiven. Sie fassen Bildmotive zusammen, die sich hinsichtlich der dargestellten Objekte und Szenen sowie der damit verbundenen sozialen, biologischen und formalen Merkmale unterscheiden. Für die vorliegende Untersuchung sind dabei insbesondere Unterschiede in sozialer Präsenz, biologischer Relevanz und visueller beziehungsweise struktureller Komplexität bedeutsam, da diese Merkmale Wahrnehmung, affektive Reaktionen und ästhetische Bewertungen systematisch beeinflussen. Im Sinne des Modells von [Leder et al. \(2004\)](#) prägen solche Merkmale bereits frühe Stufen der ästhetischen Verarbeitung, insbesondere die perzeptuelle Analyse und die Gedächtnisintegration, und wirken damit auf nachfolgende Prozesse der Klassifikation, kognitiven Meisterung und Bewertung ein. Auch aus empirisch-ästhetischer Perspektive sind Bildmotive mit struktureller Komplexität, semantischem Gehalt und emotionalem Aktivierungspotenzial verbunden. Diese Dimensionen stehen systematisch mit Arousal sowie Bewertungsdimensionen wie Gefallen und Schönheit in Beziehung ([Marin & Leder, 2013](#)).

Die in dieser Arbeit verwendeten Bildkategorien sind nicht als abschließende Typologie zu verstehen, sondern als theoretisch motivierte Gruppierungen, die bestimmte Merkmalskonstellationen bündeln. Unterschieden werden Darstellungen einzelner Personen (Individuen), Szenen mit Personengruppen (Menschenmengen), naturbezogene Darstellungen (Natur) sowie Stillleben mit menschengemachten Objekten und Arrangements (Stillleben).

Analytisch dienen die Bildkategorien nicht nur der inhaltlichen Strukturierung des Stimulusmaterials, sondern fungieren als Moderatorvariablen. Die zentrale Annahme lautet, dass sich die Beziehung zwischen visueller Exploration, physiologischer Aktivierung und expliziten Schönheitsratings in Abhängigkeit von der Bildkategorie verändert. Diese Annahme stützt sich auf theoretische und empirische Befunde zu sozialer Relevanz, naturbezogener Präferenz und visueller Komplexität, die in den nachfolgenden Unterkapiteln für die einzelnen Bildkategorien genauer erläutert werden. Vor diesem Hintergrund wird geprüft, in welchem Ausmaß die Kennwerte Fixationsanzahl, mittlere Fixationsdauer, Gesamtbetrachtungszeit und mittlerer Pupillendurchmesser je nach Bildkategorie unterschiedlich mit dem Schönheitserleben zusammenhängen. Es wird erwartet, dass sich sowohl die Blickmetriken und Pupillendurchmesser zwischen den vier Bildkategorien unterscheiden als auch die Stärke des Zusammenhangs zwischen Blick- und Pupillenmaßen und expliziten Schönheitsratings von der Bildkategorie abhängt.

2.1.1 Individuen

Unter der Kategorie Individuen werden in dieser Arbeit Darstellungen einzelner oder weniger (maximal drei) menschlicher Figuren verstanden, bei denen Gesichter und emotionale Ausdrucksmerkmale klar erkennbar sind und die prominent im zentralen Bildraum erscheinen, etwa in Form von Porträts oder klar hervorgehobenen Hauptfiguren.



Abbildung 1

Beispielgemälde der Kategorie „Individuen“ (vgl. [Stimulusmaterial im Anhang](#))

Aufgrund ihrer hohen sozialen Relevanz besitzen solche Motive eine besondere Priorität für das visuelle System und binden typischerweise ausgeprägte Aufmerksamkeit. Bereits Neugeborene zeigen eine Präferenz für gesichtsähnliche Stimuli ([Johnson et al., 1991](#); [Goren et al., 1975](#)), und im Erwachsenenalter belegen Befunde zu spezialisierten Arealen wie der fusiform face area die effiziente Verarbeitung von Identität, Emotionen und anderen sozial relevanten Merkmalen ([Kanwisher & Yovel, 2006](#)). Im Rahmen des Modells von [Leder et al. \(2004\)](#) lässt sich dies so fassen, dass Personenbilder bereits in der frühen perzeptuellen Analyse und Gedächtnisintegration eine privilegierte Verarbeitung erfahren, weil sie an stark voraktivierte soziale Schemata und Wissensbestände anschließen. Erste Urteile über Eigenschaften wie Attraktivität oder Vertrauenswürdigkeit können bereits nach sehr kurzer Darbietungszeit gebildet werden ([Willis & Todorov, 2006](#)). Solche raschen, weitgehend automatischen Urteile bilden einen Teil der frühen Bewertungsprozesse, die in späteren Stufen zu expliziten Schönheitsurteilen integriert werden. Aktuelle Arbeiten zeigen zudem, dass auch Pupillenreaktionen sensitiv auf soziale Signale und Gesichtsinformationen reagieren und damit

physiologische Korrelate sozialer Relevanz liefern ([Ricou et al., 2024](#)). Individuen-Motive eignen sich daher, um zu prüfen, wie soziale Relevanz mit physiologischer Aktivierung und ästhetischen Präferenzen zusammenhängt.

Eye-Tracking-Studien zu Kunstwerken mit menschlichen Figuren zeigen, dass Blicke systematisch auf Gesichter und zentrale Körperpartien gelenkt werden, selbst wenn andere Bildbereiche bottom-up salient sind ([Massaro et al., 2012](#)). Es ist daher plausibel, dass Individuen-Motive mit erhöhter Fixationsanzahl, längerer mittlerer Fixationsdauer und höherer Gesamtbetrachtungszeit einhergehen, da Betrachter*innen mehr Zeit in die Auswertung von Gesichtszügen, Ausdruck und individueller Ästhetik investieren ([Valuch et al., 2015](#)). Vor diesem Hintergrund wird erwartet, dass bei Individuen-Darstellungen sowohl Blickmetriken als auch Pupillendurchmesser stark mit dem Schönheitsrating korrelieren, da hier soziale Relevanz, personenbezogene Zuschreibungen (z.B. Attraktivität) und ästhetische Bewertung besonders stark miteinander zusammenhängen ([Cheng et al., 2021](#)).

2.1.2 Menschenmengen

Unter Menschenmengen werden Szenen verstanden, in denen zahlreiche Personen in einem gemeinsamen öffentlichen Raum dargestellt sind, etwa Markt-, Fest- oder Alltagsszenen, ohne dass eine einzelne Figur klar als zentrale Hauptperson hervorgehoben wird. In Menschenmengen-Szenen sind Personen typischerweise über den Bildraum verteilt, in Gruppen gegliedert und in unterschiedliche Aktivitäten und Interaktionen eingebunden, sodass sich die visuelle Aufmerksamkeit eher über viele Figuren verteilt.



Abbildung 2

Beispielgemälde der Kategorie „Menschenmengen“ (vgl. [Stimulusmaterial im Anhang](#))

Studien zur Betrachtung sozialer Szenen zeigen, dass Blicke zwar bevorzugt auf Personen und Gesichter gerichtet sind, Fixationen sich in mehrpersonalen Situationen aber breiter über die Szene verteilen und von Interaktionen und Gruppenkonstellationen gesteuert werden ([Birmingham et al., 2008](#); [Villani et al., 2015](#); [Skripkauskaitė et al., 2023](#)).

Im Vergleich zur Darstellung von Individuen ist daher von einem scannenden Betrachtungsstil auszugehen, der sich in erhöhter Fixationsanzahl und möglicherweise längerer Gesamtbetrachtungszeit niederschlagen kann, weil fortlaufend zwischen mehreren Personen und relevanten Bildelementen gewechselt wird ([Birmingham et al., 2009](#)). Gleichzeitig legen Befunde zur Wahrnehmung von Menschenmengen nahe, dass aus vielen Gesichtern und Körperhaltungen zusammenfassende Informationen über kollektive Zustände wie Stimmung oder Bewegungsrichtung extrahiert werden können, also eine Art verdichtete Repräsentation sozialer Signale entsteht ([Sweeny et al., 2013](#); [Whitney & Yamanashi Leib, 2018](#)). Es ist daher offen, ob ästhetische Bewertungen hier stärker gruppenbezogene Eindrücke reflektieren und in der Folge andere Zusammenhänge zwischen Blickmaßen und Ratings zu beobachten sind als bei

Individuen. Zudem könnte die physiologische Reaktion, etwa in Form von Pupillenweitung, stärker durch wahrgenommene soziale Interaktion und kollektive Dynamik als durch einzelne Gesichter bestimmt sein. Für die vorliegende Arbeit wird daher erwartet, dass Menschenmengen intensive und breit verteilte Exploration hervorrufen. Der Zusammenhang zwischen Blickmetriken und Schönheitsratings dürfte jedoch schwächer und weniger direkt ausfallen als bei Darstellungen von Individuen, weil die ästhetische Bewertung hier stärker auf die Gesamtszene als auf einzelne Figuren bezogen ist.

2.1.3 Natur

Unter der Kategorie Natur werden in dieser Arbeit Darstellungen von Naturlandschaften verstanden, in denen Pflanzen, Früchte, Wasser, Felsen, Himmel oder andere natürliche Elemente den Bildraum dominieren und menschliche Figuren, falls vorhanden, höchstens eine untergeordnete Rolle spielen.



Abbildung 3

Beispielgemälde der Kategorie „Natur“ (vgl. [Stimulusmaterial im Anhang](#))

Naturmotive sind häufig reich an Details und hierarchisch organisiert, sodass sowohl lokale Elemente als auch der Aufbau der Szene integriert werden müssen. [Jacobs \(2011\)](#) beschreibt, dass Konfigurationen von Landschaftsmerkmalen, etwa die Kombination von Übersicht und Schutz oder die Präsenz von Wasser und Vegetation, systematisch bevorzugt werden, weil sie neben hoher Aufenthaltsqualität auch funktionale Bedeutungen wie Sicherheit, Orientierung und Ressourcenverfügbarkeit signalisieren. Solche Bedeutungszuschreibungen greifen damit direkt in die Stufe der expliziten Klassifikation ein und verbinden biologische Relevanz mit ästhetischer Bewertung.

Darstellungen von Natur werden im Vergleich zu urbanen oder technisch geprägten Szenen häufig als ästhetischer, angenehmer und erholsamer bewertet. Studien berichten, dass beim Betrachten von Natur weniger, dafür längere Fixationen auftreten und die Scanpfade weniger fragmentiert sind, was auf eine vergleichsweise mühelose, aber involvierte Verarbeitung hindeutet ([Valtchanov & Ellard, 2015](#); [Franěk et al., 2018](#)). Daraus ergibt sich die Erwartung, dass Natur-Darstellungen mit hoher Gesamtbetrachtungszeit, moderater Fixationsanzahl und längerer mittlerer Fixationsdauer einhergehen, was als Hinweis auf eine intensivere Auseinandersetzung mit Struktur und Ästhetik der Landschaft interpretiert werden kann. Unklar ist jedoch, ob diese Muster primär Erholung und Faszination widerspiegeln und in welchem Ausmaß sie tatsächlich mit expliziten Schönheitsratings korrelieren. Die vorliegende Arbeit nutzt daher Natur-Darstellungen als eigene Kategorie, um zu prüfen, ob Blick- und Pupillenmaße, die häufig mit Erholung und Faszination in Verbindung gebracht werden, systematisch mit dem expliziten Schönheitserleben im Museumssetting kovariieren. Zugleich wird untersucht, wie sich diese Zusammenhänge von denen sozial geprägter Bildkategorien unterscheiden.

2.1.4 Stilleben

Unter der Kategorie Stilleben werden in dieser Arbeit Darstellungen menschengemachter Objekte verstanden, etwa Alltagsgegenstände, Gefäße, Geschirr oder arrangierte Nahrungsmittel, bei denen gestaltete Formen und Materialien den Bildraum dominieren. Die ausgewählten Gemälde zeigen somit komponierte Arrangements von Dingen, die in ihrem Aufbau und ihrer Oberfläche kulturell und funktional geprägt sind.



Abbildung 4

Beispielgemälde der Kategorie „Stilleben“ (vgl. [Stimulusmaterial im Anhang](#))

Solche Motive weisen oft klare geometrische Strukturen, Symmetrien und wiederkehrende Muster auf und aktivieren kulturell und individuell geprägte Präferenzen für Ordnung, Proportion und Funktionalität ([Bertamini et al., 2019](#)). Eye-Tracking-Arbeiten zeigen, dass Blickmuster systematisch zwischen natürlichen und menschengemachten bzw. urbanen Szenen variieren und dass Parameter wie Fixationsanzahl und Fixationsdauer zuverlässig Unterschiede im Verarbeitungsmodus abbilden, wobei urbane Szenen typischerweise mehr, kürzere Fixationen und stärker explorative Blickmuster hervorrufen, während natürliche Umgebungen mit weniger Fixationen und längeren Fixationsdauern einhergehen ([Franěk et al.,](#)

2019; Batool et al., 2022). Für gestaltete Produkte im Industriedesign konnte zudem gezeigt werden, dass Anzahl, Dauer und räumliche Verteilung von Fixationen systematisch mit ästhetischen Präferenzen korrelieren und zur Vorhersage von Produktattraktivität genutzt werden können (Khalighy et al., 2015). Studien zu architektonischen und anderen menschengemachten Stimuli legen darüber hinaus nahe, dass steigende visuelle Komplexität mit intensiverer visueller Exploration einhergeht, etwa mehr Fixationen und stärker verteilten Blickpunkten, während einfachere, klar strukturierte Gestaltungen eher fokussierte, kürzere Blicksequenzen begünstigen (Beder et al., 2024).

Vor diesem Hintergrund ist zu erwarten, dass Stilleben in der vorliegenden Studie eine mittlere bis hohe Gesamtbetrachtungszeit und eine moderate bis erhöhte Fixationsanzahl aufweisen, wenn Formen, Oberflächen und räumliche Beziehungen systematisch erkundet werden. Da hier unmittelbare biologische Relevanz geringer und gelernte Gestaltungsnormen wichtiger sind, dürfte die Übersetzung von visueller Exploration in ästhetische Urteile stärker von individuellen Erfahrungen, Expertise und Präferenzen abhängen als bei sozialen oder Naturmotiven. Bei Stilleben dürften Blickmetriken stärker kognitive Strukturierungsprozesse als unmittelbare affektive Reaktionen widerspiegeln. Auch Pupillenreaktionen sind hier eher als unspezifische Indikatoren von Anstrengung und Involvierung zu verstehen. Die Kategorie Stilleben dient in dieser Studie somit als Kontrastbedingung zu biologisch und sozial hoch relevanten Inhalten.

2.2 Blickmetriken als Prädiktoren ästhetischen Gefallens

In dieser Arbeit werden Fixationsanzahl (FC), Gesamtbetrachtungszeit (TFD) und mittlere Fixationsdauer (MFD) als zentrale Kennwerte der visuellen Exploration herangezogen. FC gibt an, wie häufig Betrachter*innen den Blick auf ein Gemälde richten, und erlaubt damit

Rückschlüsse darauf, in welchem Ausmaß ein Motiv wiederholt Aufmerksamkeit anzieht. TFD erfasst die kumulierte Dauer aller Fixationen auf einem Gemälde und bildet so die Betrachtungsintensität ab, also den Umfang visueller Ressourcen, die ein Bild bindet. MFD beschreibt die mittlere Dauer einzelner Fixationen und wird häufig als Hinweis auf die Tiefe lokaler Verarbeitung interpretiert, etwa im Sinne ausführlicherer Analyse oder semantischer Integration bei längeren Fixationen. Gerade weil MFD im Unterschied zu FC und TFD weniger direkt von der Gesamtdauer der Betrachtung beeinflusst wird, eignet sich dieses Maß besonders zur Erfassung kategorialer Unterschiede in der Tiefe visueller Verarbeitung.

Wie bereits in der Einleitung dargestellt, zeigt die Forschungslage über verschiedene Motivklassen hinweg, dass ästhetisch bevorzugte Stimuli typischerweise mit längerer Betrachtungsdauer, höherer Fixationsanzahl und insgesamt intensiverer visueller Exploration einhergehen ([Goller et al., 2019](#); [Holmes & Zanker, 2012](#); [Jankowski et al., 2020](#); [Leder et al., 2016](#); [Faust et al., 2019](#); [Brieber et al., 2014](#); [Marin & Leder, 2022](#)).

Attraktivere oder stärker gemochte Gesichter werden länger betrachtet und erhalten mehr Fixationen als weniger attraktive Vergleichsstimuli, was als Hinweis interpretiert wird, dass positiv bewertete Reize Aufmerksamkeit stärker binden und wiederholt aufgesucht werden ([Goller et al., 2019](#); [Holmes & Zanker, 2012](#); [Jankowski et al., 2020](#); [Leder et al., 2016](#)).

Vergleichbare Muster wurden auch für Kunstwerke und andere menschengemachte Objekte beschrieben ([Brieber et al., 2014](#); [Mitschke et al., 2017](#)). Für MFD fallen die Effekte insgesamt kleiner aus, sie weisen jedoch in dieselbe Richtung: Unter Bedingungen klarer Präferenzunterschiede sind mittlere Fixationsdauern auf bevorzugten Bildern länger, was auf eine intensivere inhaltliche Verarbeitung hindeutet ([Goller et al., 2019](#); [Jankowski et al., 2020](#)). Zusammengefasst sprechen diese Befunde dafür, FC und TFD vor allem als Maße der

Zuwendung und Betrachtungsintensität, MFD hingegen stärker als Maß der lokalen Verarbeitungstiefe zu verstehen.

Ergänzend zu diesen Blickmaßen wird in der vorliegenden Arbeit der Pupillendurchmesser als physiologischer Indikator prozessualer Aspekte ästhetischen Erlebens berücksichtigt. Physiologisch dient die Pupille zunächst der Regulierung des Lichteinfalls, sodass Veränderungen der physischen und wahrgenommenen Luminanz einen erheblichen Teil der Varianz in der Pupillengröße erklären. Insbesondere Helligkeitsunterschiede zwischen Reizen und Veränderungen der Beleuchtung führen zu raschen, reflexartigen Verengungen oder Erweiterungen, die nicht unmittelbar mit subjektiven Bewertungen verknüpft sind ([Bradley et al., 2008](#); [Henderson et al., 2018](#)). Über diese basale Lichtanpassung hinaus gilt die Pupille in der psychophysiologischen Forschung jedoch als sensibles Fenster auf emotionales Arousal und autonome Aktivierung. Studien mit affektiven Bildern zeigen konsistent, dass emotional ansprechendere Stimuli bei kontrollierter Luminanz stärkere Pupillenweitungen hervorrufen als neutrale Bilder, unabhängig davon, ob sie positiv oder negativ bewertet werden ([Bradley et al., 2008](#); [Henderson et al., 2018](#); [Laeng et al., 2012](#)).

Auch im Kunstkontext finden sich Hinweise auf einen Zusammenhang zwischen Pupillendilatation und ästhetischem Erleben, wenngleich die Befundlage weniger eindeutig ist. [Kuchinke et al. \(2009\)](#) berichteten, dass weniger abstrakte und leichter zu verarbeitende Gemälde tendenziell höhere Gefallensurteile und stärkere Pupillenerweiterungen hervorrufen als stärker fragmentierte Varianten kubistischer Gemälde. Neuere Arbeiten replizieren jedoch keine robuste, eindeutig positive Beziehung zwischen Präferenzurteilen und Pupillendilatation ([Elschner et al., 2018](#)). Stattdessen spricht ein Teil der neueren Literatur dafür, Pupillenreaktionen vorsichtig als unspezifische Marker kognitiver Anstrengung, Verarbeitungsflüssigkeit und affektiver Involvierung zu interpretieren, nicht aber als direkten Index des Erlebens von Schönheit. Hinzu

kommt, dass aktuelle dynamische Paradigmen die zeitliche und kontextuelle Komplexität pupillärer Reaktionen betonen. [Spee et al. \(2024\)](#) zeigen, dass Pupillenverläufe in einem engen, aber nicht linearen Zusammenspiel mit Aufgabenanforderungen, Bildstruktur sowie Gestalt- und Aha-Erlebnissen stehen. Pupillenreaktionen erscheinen damit als dynamische Marker von Bedeutungsfindung und Involvierung, deren Beziehung zu expliziten ästhetischen Urteilen von Kontext, Aufgabe und Stimulusklasse moderiert wird.

Für die vorliegende Untersuchung ergibt sich daraus eine gemeinsame theoretische Ableitung: Blickmetriken und Pupillendurchmesser erfassen unterschiedliche, sich jedoch ergänzende Facetten ästhetischen Erlebens. Während FC, TFD und MFD vor allem beschreiben, wie intensiv und in welcher Weise ein Gemälde visuell exploriert wird, bildet der Pupillendurchmesser stärker automatisch ablaufende affektive und aktivierungsbezogene Prozesse ab. Daraus folgt für die vorliegende Arbeit die Frage, ob und in welchem Ausmaß explizite Schönheitsurteile mit diesen Blickmustern und pupillären Reaktionen einhergehen und ob sich diese Zusammenhänge zwischen unterschiedlichen Bildkategorien systematisch unterscheiden. Im Museumssetting dienen FC, TFD und MFD daher einerseits der Beschreibung bildkategoriespezifischer Betrachtungsprofile und andererseits der Prüfung ihres Beitrags zur Vorhersage expliziter Schönheitsratings; der Pupillendurchmesser ergänzt diese Analyse um eine physiologische Perspektive auf affektive und kognitive Involvierung.

2.3 Forschungsfragen der vorliegenden Untersuchung

Ausgehend von den theoretischen Überlegungen verfolgt die vorliegende Studie zwei zentrale Forschungsfragen. Erstens wird untersucht, in welchem Ausmaß sich die mittlere Fixationsdauer (MFD) sowie der Pupillendurchmesser zwischen den vier Bildkategorien Individuen, Menschenmengen, Stillleben und Natur unterscheiden (F1). Konkret wird erwartet,

dass soziale Motive (Individuen, Menschenmengen) längere mittlere Fixationsdauern aufweisen als Stilleben (H1a) und dass der Pupillendurchmesser bei sozialen Motiven erhöht ist (H1b). Weitere Blickmetriken wie TFD und FC werden ergänzend explorativ ausgewertet, ohne dass hierfür spezifische gerichtete Hypothesen formuliert werden.

Zweitens wird geprüft, in welchem Ausmaß die mittlere Fixationsdauer und der Pupillendurchmesser das explizite Schönheitsrating von Kunstwerken vorhersagen (F2). Es wird angenommen, dass die mittlere Fixationsdauer in einem linearen gemischten Modell mit zufälligen Intercepts für Personen die Schönheitsratings positiv vorhersagt (H2a) und dass der Pupillendurchmesser in einem solchen Modell ebenfalls ein positiver Prädiktor der Schönheitsratings ist (H2b). Zusätzlich wird explorativ untersucht, ob sich die Stärke dieser Zusammenhänge zwischen den Bildkategorien unterscheidet, ohne dass hierfür eine spezifische Richtungshypothese formuliert wird.

Insgesamt wird also geprüft, ob die Bildkategorie die Zusammenhänge zwischen mittlerer Fixationsdauer, Pupillenreaktion und Schönheitsurteil verändert.

3. Methodik

3.1 Studiendesign und Teilnehmer*innen

Die Datenerhebung erfolgte im Rahmen einer Feldstudie im Museum Belvedere in Wien, mit dem Ziel, die Kunstrezeption unter alltagsnahen Museumsbedingungen zu untersuchen. Zur Erfassung des Blickverhaltens während des Museumsbesuchs wurde ein mobiles Eye-Tracking-System eingesetzt. Die Untersuchung fand im Januar 2024 statt und dauerte etwa 40 Minuten pro Person; die für die Analyse dieser Arbeit relevante freie Betrachtungsphase im Versuchsraum war auf maximal 15 Minuten begrenzt. Die Stichprobe bestand aus $N = 114$ Studierenden der

Universität Wien, die für ihre Teilnahme eine Vergütung in Form von Versuchspersonenstunden (Labor-Credits) erhielten. Für die vorliegende Auswertung wurden nur Datensätze berücksichtigt, die die definierten Qualitätskriterien erfüllten und für die geplanten Analysen geeignet waren, sodass die endgültige Eye-Tracking-Stichprobe $n = 25$ umfasste (siehe Abschnitt 3.2). Das Durchschnittsalter der in dieser Analyse inkludierten Teilnehmenden lag bei $M = 22.17$ Jahren ($SD = 3.46$; Spannweite: 19–34 Jahre). Alle Teilnehmenden berichteten eine normale oder korrigierte Sehfähigkeit.

Das Untersuchungsdesign wurde als Free-Walking-Task mit paralleler kontinuierlicher Erfassung von Schönheitsratings umgesetzt. Die Teilnehmenden erhielten die Instruktionen, sich frei im ausgewählten Versuchsraum zu bewegen und die ausgestellten Gemälde so zu betrachten, wie sie es bei einem normalen Museumsbesuch tun würden. Sie sollten alle Gemälde ansehen und anschließend zu dem Punkt am Eingang des Raumes zurückkehren, an dem ihr Rundgang begonnen hatte. Der Rundgang war in seiner Dauer grundsätzlich selbstgesteuert, wurde jedoch nach spätestens 15 Minuten von der Versuchsleitung beendet, falls die Teilnehmenden den Aufenthalt in dem Raum bis dahin noch nicht eigenständig beendet hatten. Auf diese Weise sollte eine natürliche, individuell variierende Exploration bei gleichzeitig vergleichbarer maximaler Betrachtungsdauer gewährleistet werden. Die Reihenfolge der betrachteten Gemälde, die Verweildauer vor einzelnen Werken und die gewählten Laufwege wurden nicht experimentell kontrolliert, sondern als charakteristische Merkmale der natürlichen Rezeptionssituation zugelassen.



Abbildung 5

Der Versuchsraum „Spielezimmer“ im Belvedere Museum

Die Aufzeichnung der Blickbewegungen erfolgte mit dem mobilen Eye-Tracking-System Tobii Pro Glasses 3 ([Tobii Pro, 2024](#)), das Fixationen, Sakkaden und Pupillenweite mit hoher zeitlicher Auflösung in der natürlichen Betrachtungssituation erfasste. Vor Beginn des Rundgangs wurde für jede Person eine Kalibrierung mit der standardisierten Tobii-Pro-Routine durchgeführt; bei unzureichender Kalibrierungsqualität wurde der Vorgang wiederholt, bis eine stabile Zuordnung der Blickpunkte erreicht war.

Parallel zur visuellen Exploration bewerteten die Teilnehmenden kontinuierlich die wahrgenommene Schönheit des jeweils betrachteten Kunstwerks. Die Erfassung der subjektiven Ästhetikurteile erfolgte über die webbasierte Forschungsplattform [emoTouch, 2025](#)), wobei die Teilnehmenden hierfür ihre eigenen Smartphones verwendeten. Auf einer kontinuierlichen Analogskala von 0 bis 100, deren Endpunkte mit „gar nicht schön“ und „sehr schön“ verankert waren, gaben sie während des gesamten Rezeptionsvorgangs ihr Urteil ab. Durch die kontinuierliche Analogskala konnten auch kurzzeitige Schwankungen im ästhetischen

Empfinden zeitlich hochaufgelöst abgebildet werden. Die Synchronisation der Blickdaten mit den Schönheitsratings erfolgte über identische Zeitstempel, sodass jedem Blickereignis der zeitlich nächstliegende Wert zugeordnet werden konnte. Dieser methodische Ansatz ermöglichte einen kombinierten Datensatz, der explizite subjektive Urteile mit impliziten, verhaltens- und physiologiebasierten Maßen des Rezeptionsprozesses verknüpft.

Das Untersuchungsdesign wurde von der Ethikkommission der Fakultät für Psychologie der Universität Wien positiv beurteilt. Alle Teilnehmenden wurden schriftlich und mündlich über den Ablauf der Studie, den Umgang mit ihren personenbezogenen und aufgezeichneten Daten sowie ihr Recht auf jederzeitigen Abbruch informiert. Vor Beginn der Erhebung gaben sie ihre schriftliche Einwilligung.

3.2 Stimulusmaterial und Datenaufbereitung

Zur präzisen Zuordnung der Blickdaten zu den Kunstwerken wurden die Rohdaten in mehreren Schritten gefiltert und validiert. In einem ersten Schritt wurden nur jene Datensätze für die Analyse herangezogen, bei denen mindestens 95% der registrierten Fixationen durch die Software Tobii Pro Lab ([Tobii Pro, 2022](#)) automatisch einem Referenzbild zugeordnet werden konnten. Der Schwellenwert von 95% wurde als konservative methodische Setzung gewählt, um einen Kompromiss zwischen Datenquantität und Mapping-Genauigkeit zu erreichen und nur Datensätze mit hoher Zuverlässigkeit der Bildzuordnung in die weiteren Analysen einzubeziehen. Anschließend erfolgte eine manuelle Validierung aller gemappten Fixationen, um systematische Fehler des automatischen Mappings zu korrigieren. Artefakte, etwa durch Blinzeln, Signalverluste oder Reflexionen, wurden identifiziert und ausgeschlossen, sodass die weitere Analyse ausschließlich auf validen Fixationen basierte. Zusätzlich wurden nur jene Fälle in die Auswertung einbezogen, in denen ein durchgängiger, technisch einwandfreier Verlauf der

kontinuierlichen Schönheitsratings vorlag (z. B. kein vorzeitiger Abbruch oder Ausfall der emoTouch-Messung). Nach Anwendung dieser Kriterien reduzierte sich die für die Eye-Tracking-Analysen verfügbare Stichprobe auf $n = 25$ Teilnehmende.

Das Stimulusmaterial umfasste 26 Gemälde, die nach ihrer Motivik vier inhaltlich klar abgrenzbaren Kategorien zugeordnet wurden: Natur (Naturlandschaften, $n = 10$), Individuen (Darstellungen einzelner oder weniger Personen mit klar erkennbaren Gesichtern und emotionalen Ausdrücken, $n = 6$), Menschenmengen (Darstellungen von Personengruppen ohne eindeutig hervorgehobene Hauptfigur, $n = 5$) und Stillleben (von Menschen geschaffene Objekte und Stillleben, $n = 5$). Für die Analyse der Blickbewegungen wurde jedes Gemälde als eigenständige Area of Interest (AOI) definiert, die den jeweils sichtbaren Gemäldeausschnitt umfasste. Eine detaillierte Übersicht über Titel, Künstler:innen und Entstehungsjahre der Werke findet sich im Anhang ([Stimulusmaterial](#)). Die Reihenfolge der Betrachtung ergab sich aus dem individuellen Rundgang der Teilnehmenden im Rahmen des Free-Walking-Tasks und wurde nicht experimentell kontrolliert, um die natürliche Rezeptionssituation möglichst unverändert abzubilden.

3.3 Datenauswertung

Die Eye-Tracking-Rohdaten wurden auf Fixationsebene bereinigt, indem ausschließlich Fixationen mit gültigen Pupillenmessungen beider Augen und einer Dauer zwischen 60 und 1500 ms berücksichtigt wurden. Kürzere, vermutlich sakkadische Ereignisse sowie längere, potenziell Artefakt behaftete Fixationen wurden damit ausgeschlossen. Pupillenzeitreihen wurden zusätzlich von Blinzern und Signalabbrüchen bereinigt, bevor Mittelwerte pro Fixation berechnet wurden.

Die so bereinigten Fixationen wurden mit den kontinuierlichen Schönheitsratings aus emoTouch zusammengeführt. Die Zuordnung erfolgte über Versuchspersonen-IDs und Zeitinformation, wobei jeder Fixation der zeitlich nächstgelegene Rating-Wert zugewiesen wurde. Anschließend wurden alle Fixationen, die außerhalb der definierten AOIs lagen, verworfen, sodass nur Blickereignisse, die innerhalb der einzelnen Gemälde lagen, in die weitere Analyse eingingen.

Im nächsten Schritt wurden die Daten aggregiert. Zunächst wurden für jede Kombination aus Person und Bild die Gesamtbetrachtungszeit, die Fixationsanzahl, die mittlere Fixationsdauer und der mittlere Pupillendurchmesser sowie das mittlere Schönheitsrating berechnet. Darauf aufbauend wurden diese Kennwerte pro Person und Bildkategorie gemittelt, sodass für jede teilnehmende Person und jede der vier Kategorien (Individuen, Menschenmengen, Natur, Stillleben) durchschnittliche Blickmetriken und ein entsprechendes mittleres Schönheitsrating vorlagen. Diese pro Person nach Bildkategorie aggregierten Daten bildeten die Grundlage für alle deskriptiven und inferenzstatistischen Auswertungen.

Zur Beantwortung der ersten Forschungsfrage (F1) wurden auf Basis der personenweise nach Bildkategorien aggregierten Daten deskriptive Statistiken für FC, TFD, MFD und PD pro Kategorie berechnet. Ergänzend wurden für alle vier Maße einfaktorielle Varianzanalysen mit der Bildkategorie als Faktor durchgeführt, um kategoriale Unterschiede inferenzstatistisch zu prüfen. Dabei erfolgten die Analysen für TFD und FC explorativ, während MFD und PD im Zentrum der hypothesentestenden Auswertung standen.

Zur Beantwortung der zweiten Forschungsfrage (F2), in welchem Ausmaß Blickmetriken und Pupillendurchmesser das explizite Schönheitsrating vorhersagen und ob sich diese Zusammenhänge zwischen den Bildkategorien unterscheiden, wurden mehrere korrelative und modellbasierte Analysen durchgeführt. Zunächst wurden Pearson-Korrelationen zwischen Rating

und den Blickmetriken (TFD, FC, MFD, PD) getrennt nach Kategorie berechnet und tabellarisch sowie grafisch dargestellt, um kategoriespezifische lineare Zusammenhänge zu beschreiben. Im Anschluss wurden alle kontinuierlichen Variablen (Schönheitsrating, TFD, FC, MFD, PD) z-standardisiert, um Skalierungsunterschiede zu reduzieren und Regressionskoeffizienten vergleichbar zu machen. Auf dieser Basis wurde zunächst ein multiples lineares Regressionsmodell (MLR) geschätzt, in dem das z-standardisierte Schönheitsrating durch z-standardisierte TFD, FC, MFD und PD sowie die Bildkategorie als festen Effekt vorhergesagt wurde.

Da mehrere Beobachtungen pro Person vorlagen, wurden zusätzlich lineare gemischte Modelle (LMM) mit zufälligen Intercepts für Teilnehmende spezifiziert, um interindividuelle Unterschiede im allgemeinen Bewertungsniveau zu berücksichtigen. Als Prädiktoren gingen die Bildkategorie sowie die z-standardisierten Blickmetriken in das Modell ein. Die Ergebnisse der deskriptiven Analysen, Korrelationen und Modelle werden im folgenden Kapitel im Hinblick auf F1 und F2 berichtet.

4. Ergebnisse

4.1 Bildkategoriespezifische Blickprofile

In die finale Analyse flossen insgesamt 138 866 gültige Fixationen ein. Die MFD lag im Mittel bei 561 ms ($SD = 328$ ms, Range: 60–1483 ms), der mittlere PD bei 5.30 mm ($SD = 0.75$, Range: 2.93–7.60), und die kontinuierlichen Schönheitsratings verteilten sich mit einem Mittelwert von 50.62 ($SD = 27.69$, Range: 1–100) über den gesamten Skalenbereich. Für die folgenden Analysen standen nach Aggregation der Fixationsdaten 615 Beobachtungen auf der Ebene „Person innerhalb einer Bildkategorie“ zur Verfügung. Die zugrunde liegenden 26

Gemälde verteilten sich auf vier Bildkategorien (10 Natur-Gemälde, 6 Individuen-Gemälde, 5 Menschenmengen-Gemälde und 5 Stillleben-Gemälde). Tabelle 1 fasst Mittelwerte, Standardabweichungen und Wertebereiche von TFD, FC, MFD und PD nach Bildkategorie zusammen. Die deskriptiven Kennwerte zeigen unterschiedliche Rezeptionsmuster zwischen den Kategorien. Menschenmengen-Gemälde wiesen die höchste mittlere TFD ($M = 2:47$ min) und FC ($M = 294$) auf, gefolgt von Individuen (TFD $M = 2:20$ min, FC $M = 226$) und Natur (TFD $M = 2:13$ min, FC $M = 242$), während Stillleben die niedrigste TFD ($M = 1:25$ min) und FC ($M = 159$) zeigten. Die MFD war bei Individuen am höchsten ($M = 621$ ms), lag bei Menschenmengen ($M = 538$ ms) und Natur ($M = 514$ ms) im mittleren Bereich und war bei Stillleben ($M = 478$ ms) am geringsten, was bereits deskriptiv die in H1a erwarteten längeren MFD-Werte für soziale Motive gegenüber Stillleben widerspiegelt. Der PD variierte zwischen den Kategorien nur moderat; Individuen-, Menschenmengen- und Stillleben-Gemälde wiesen ähnliche mittlere PD im Bereich von etwa 5.4–5.5 mm auf, während Natur-Gemälde geringfügig niedrigere Werte zeigten ($M = 5.31$ mm).

Tabelle 1*Deskriptive Statistik der Blickmetriken (TFD, FC, MFD, PD) nach Bildkategorie*

	Individuen			Menschenmengen			Natur			Stilleben		
	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range
TFD (min)	2:20	1.17	0:45–4:44	2:47	1.67	0:40–6:06	2:13	1.18	0:34–5:38	1:25	1.06	0:15–4:13
FC	226	117	84–496	294	155	95–614	242	96	69–445	159	93	36–396
MFD (ms)	621	100	489–962	538	122	292–805	514	101	354–741	478	106	319–654
PD (mm)	5.47	0.73	4.07–6.96	5.44	0.78	3.92–6.94	5.31	0.75	3.94–6.89	5.46	0.66	4.21–6.76

Anmerkung. $N = 25$. *TFD* = Gesamtbetrachtungszeit (Total Fixation Duration, in Minuten); *FC* = Fixationsanzahl (Fixation Count); *MFD* = mittlere Fixationsdauer (Mean Fixation Duration, in Millisekunden); *PD* = mittlerer Pupillendurchmesser (Pupil Diameter, in Millimetern). *M* und *SD* stehen für Mittelwert bzw. Standardabweichung. Die Werte im Bereich „Range“ geben Minimum und Maximum an.

Abbildung 6 ergänzt diese tabellarische Übersicht um Boxplots der vier Blickmetriken und des Pupillendurchmessers nach Bildkategorie und illustriert damit die Verteilung und Streuung der Werte innerhalb der Kategorien.

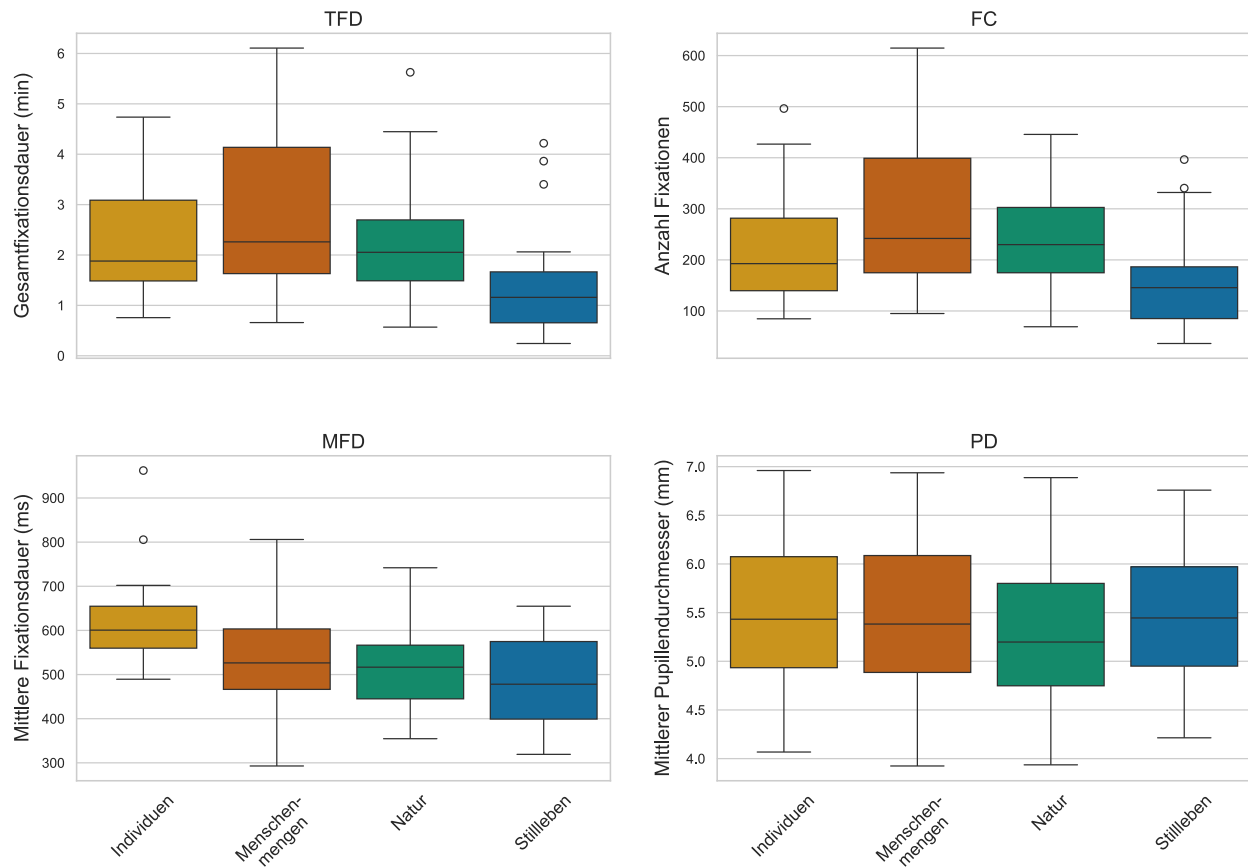


Abbildung 6

Boxplots der Blickmetriken nach Bildkategorie, basierend auf den pro Person und Bildkategorie aggregierten Daten: Gesamtfixationsdauer (TFD), Fixationsanzahl (FC), mittlere Fixationsdauer (MFD) und mittlerer Pupillendurchmesser (PD)

Auf z-standardisierter Ebene (siehe Abbildung 7) werden kategoriespezifische Blickprofile sichtbar, die Unterschiede im Gesamtmuster von TFD, FC, MFD und PD verdeutlichen. Soziale Motive (Individuen- und Menschenmengen-Gemälde) zeigen überdurchschnittliche Werte in mehreren Blickmetriken, während Stilleben durchgängig unter

dem Stichprobenmittel liegen und Natur-Gemälde meist im mittleren Bereich. Die Blickprofile legen nahe, dass sich die vier Bildkategorien nicht nur in einzelnen Kennwerten, sondern in ihrem gesamten Muster über TFD, FC, MFD und PD unterscheiden.

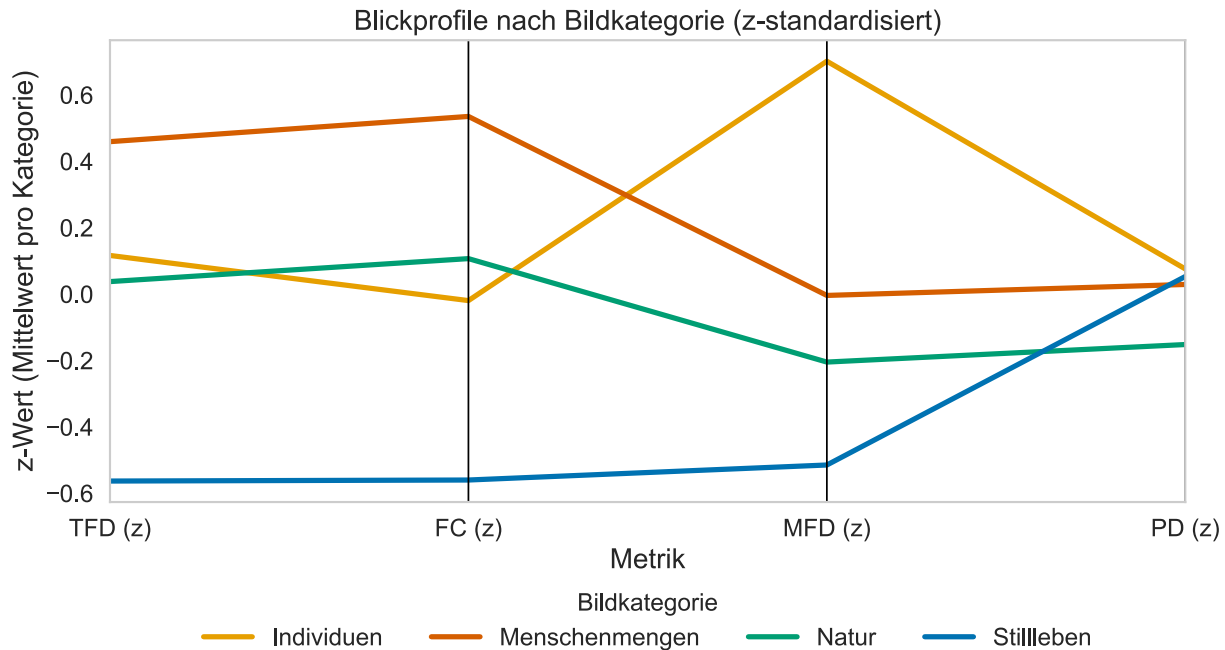


Abbildung 7

Z-standardisierte Blickprofile der vier Bildkategorien über Gesamtbetrachtungsdauer (TFD), Fixationsanzahl (FC), mittlere Fixationsdauer (MFD) und mittleren Pupillendurchmesser (PD), basierend auf den pro Person und Bildkategorie aggregierten Daten

Zur Beantwortung der ersten Forschungsfrage (F1), ob sich die Blickmetriken zwischen den Bildkategorien unterscheiden, wurden einfaktorielle Varianzanalysen mit der Bildkategorie als Faktor für TFD, FC, MFD und PD durchgeführt. Die Analysen ergaben signifikante Haupteffekte der Bildkategorie auf die Gesamtfixationsdauer, $F(3, 91) = 4.56, p = .005, \eta^2p = .13$, die Fixationsanzahl, $F(3, 91) = 5.17, p = .002, \eta^2p = .15$, und die mittlere Fixationsdauer, $F(3, 91) = 7.95, p < .001, \eta^2p = .21$, nicht jedoch auf den Pupillendurchmesser, $F(3, 91) = 0.26, p = .85, \eta^2p = .01$.

Für die mittlere Fixationsdauer wurden Post-hoc-Paarvergleiche berechnet. Für den Pupillendurchmesser entfielen solche Vergleiche, da kein signifikanter Haupteffekt der Bildkategorie vorlag. Bonferroni-korrigierte Post-hoc-Tests zeigten, dass Individuen-Gemälde signifikant längere mittlere Fixationsdauern aufwiesen als Natur- ($p = .003$) und Stillleben-Darstellungen ($p < .001$), während die übrigen Paarvergleiche keine signifikanzkorrigierten Unterschiede in der MFD ergaben (alle $p \geq .09$). Insgesamt sprechen diese Befunde dafür, dass sich mindestens einige Bildkategorien hinsichtlich der mittleren Fixationsdauer unterscheiden und dass insbesondere die längeren Fixationen auf Individuen-Darstellungen die Annahme aus H1a unterstützen, während sich der mittlere Pupillendurchmesser über Kategorien hinweg nicht systematisch unterscheidet und H1b damit nicht gestützt wird. Die mittlere Fixationsdauer und der Pupillendurchmesser werden im Folgenden als zentrale Indikatoren für die hypothesentestenden Analysen herangezogen, während TFD und FC ergänzend zur Charakterisierung der Betrachtungsmodi berichtet werden.

4.2 Blickmetriken als Prädiktoren der Schönheitsurteile

Die Zusammenhänge zwischen Blickmetriken und Schönheitsrating wurden zunächst explorativ anhand von Pearson-Korrelationen auf Basis der pro Person nach Bildkategorien aggregierten Daten untersucht. Die Korrelationskoeffizienten fielen insgesamt überwiegend schwach aus; nur wenige Zusammenhänge erreichten einen Betrag von etwa $r \approx 0,30$ oder mehr. Am deutlichsten zeigten sich Effekte bei Individuen-Darstellungen: Hier korrelierte die mittlere Fixationsdauer moderat positiv mit dem Schönheitsrating ($r \approx 0.40$), und auch der Pupillendurchmesser war positiv mit höheren Bewertungen verbunden ($r \approx 0.33$). Bei Stillleben ergab sich ebenfalls ein moderater positiver Zusammenhang zwischen Pupillendurchmesser und Schönheitsrating ($r \approx 0.35$), während die Korrelation mit der Fixationsdauer nahezu null war ($r \approx$

–0.01). Im Kontrast dazu blieben die Zusammenhänge bei Menschenmengen- und Natur-Darstellungen sehr gering. Für Menschenmengen-Gemälde zeigten sich nur sehr schwache positive Korrelationen zwischen Schönheitsrating und mittlerer Fixationsdauer bzw. Pupillendurchmesser ($r \approx 0.07$ bzw. $r \approx 0.09$), und bei Natur-Gemälden lagen die entsprechenden Koeffizienten praktisch bei null (MFD $r \approx 0.02$; PD $r \approx -0.04$). Insgesamt deutet dieses Muster darauf hin, dass MFD und PD vor allem bei Individuen- und teilweise bei Stillleben mit höheren Schönheitsratings einhergehen, während sie bei Natur- und Menschenmengen-Darstellungen kaum systematisch mit den Bewertungen verknüpft sind. Die vollständigen Korrelationskoeffizienten sind in **Tabelle 2** zusammengefasst.

Tabelle 2

Pearson-Korrelationen zwischen Schönheitsrating und Blickmetriken nach Bildkategorie

Bildkategorie	TFD	FC	MFD	PD
Individuen	-0.06	-0.08	0.40	0.33
Menschenmengen	-0.08	-0.06	0.07	0.09
Natur	0.24	0.30	0.02	-0.04
Stillleben	-0.09	0.00	-0.01	0.35

Anmerkung. N = 25. TFD = Gesamtbetrachtungszeit; FC = Fixationsanzahl; MFD = mittlere Fixationsdauer; PD = mittlerer Pupillendurchmesser.

Zur Illustration der in Tabelle 2 berichteten Zusammenhänge zeigt Abbildung 8 die Streudiagramme zwischen Schönheitsrating und den zentralen Blickmetriken getrennt nach Bildkategorie. Die Panels verdeutlichen, dass sich bei Darstellungen von Individuen eine sichtbare positive Steigung der Regressionslinien für MFD und Pupillendurchmesser ergibt,

während die Punktwolken für Menschenmengen- und Naturdarstellungen keine klaren Trends erkennen lassen.

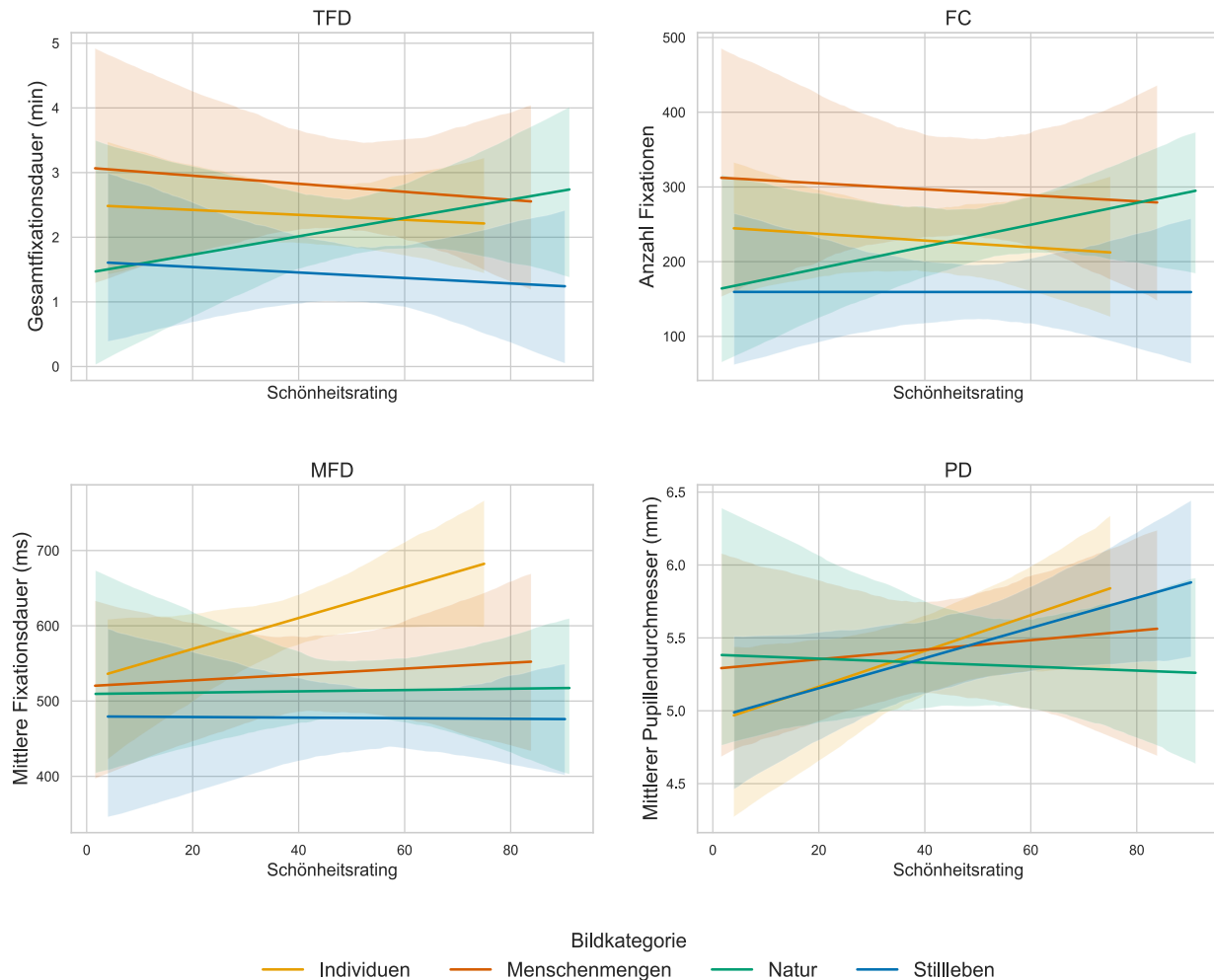


Abbildung 8

Streudiagramme der Zusammenhänge zwischen Schönheitsrating und Blickmetriken (TFD, FC, MFD, PD) nach Bildkategorie

Für Stillleben zeigt sich ein ähnliches Muster wie bei Individuen, allerdings vor allem für den Pupillendurchmesser.

Zur Beantwortung der zweiten Forschungsfrage (F2) wurden lineare Modelle mit z-standardisiertem Schönheitsrating als Kriterium und Blickmetriken sowie Bildkategorie als

Prädiktoren berechnet. Alle kontinuierlichen Prädiktoren (Schönheitsrating und Blickmetriken) wurden vor der Modellierung z-standardisiert, sodass die geschätzten Regressionskoeffizienten als Effekt einer Erhöhung um eine Standardabweichung interpretierbar sind. Im ersten Schritt wurde eine multiple lineare Regression (MLR) mit dem Schönheitsrating als abhängiger Variable und TFD, FC, MFD und PD sowie Bildkategorie als Prädiktoren gerechnet (Schätzung mit der Methode der kleinsten Quadrate). Im MLR-Modell erklärte die Kombination der Prädiktoren 17 % der Varianz des Schönheitsratings ($R^2 = .17$, $adj. R^2 = .10$, $F(7, 87) = 2.49$, $p = .022$). Dies weist auf einen kleinen, aber statistisch signifikanten Beitrag der Blickmetriken und der Bildkategorie zur Vorhersage des Ratings hin. Höhere TFD-Werte sagten ein geringeres Schönheitsrating vorher ($\beta = -1.77$, $p = .002$), während eine höhere Fixationsanzahl ($\beta = 1.64$, $p = .002$) sowie längere mittlere Fixationsdauern ($\beta = 0.55$, $p = .004$) mit höheren Schönheitsratings verbunden waren. Der Pupillendurchmesser zeigte einen positiven Zusammenhang mit dem Schönheitsrating ($\beta = 0.20$, $p = .051$), der das Signifikanzniveau von $\alpha = .05$ knapp verfehlte und daher als statistisch nicht signifikant zu interpretieren ist. Für die Bildkategorien ergab sich, dass Natur im Vergleich zur Referenzkategorie Individuen signifikant höhere Schönheitsratings vorhersagte ($\beta = 0.65$, $p = .028$), während Stillleben ($\beta = 0.55$, $p = .080$) und Menschenmengen ($\beta = 0.16$, $p = .61$) nicht signifikant von der Referenzkategorie Individuen abwichen. Im MLR-Modell ergibt sich damit ein Muster, in dem höhere Schönheitsratings mit kürzerer Gesamtfixationsdauer, zugleich aber mit mehr Fixationen, längeren mittleren Fixationsdauern und einem positiv, aber nicht signifikant mit Schönheitsratings verknüpften Pupillendurchmesser einhergehen. Angesichts der hohen inhaltlichen und statistischen Nähe von TFD und FC ist die Interpretation der gegenläufigen Regressionsgewichte jedoch vorsichtig vorzunehmen, während MFD und Pupillendurchmesser als inhaltlich zentralere Prädiktoren für H2a und H2b im Fokus stehen.

Um der hierarchischen Datenstruktur mit wiederholten Messungen pro Person Rechnung zu tragen, wurde ergänzend ein lineares gemischtes Modell (LMM) mit denselben festen Effekten und zufälligen Intercepts für Teilnehmende geschätzt. Dieses Modell basierte auf 95 Beobachtungen, die sich aus den pro Person und Bildkategorie aggregierten Daten ergaben (25 Personen mit jeweils bis zu 4 Bildkategorien). Das LMM wies eine merkliche Varianzkomponente für die zufälligen Intercepts der Teilnehmenden auf (Varianz der Intercepts ≈ 0.55 ; Residualvarianz ≈ 0.42), was darauf hinweist, dass Unterschiede zwischen den Personen zur Gesamtvarianz der Schönheitsratings beitragen und im Modell entsprechend berücksichtigt werden. In diesem LMM waren die Effekte der Blickmetriken TFD und FC nicht signifikant (β zwischen -0.59 und 0.57 , alle $p \geq .23$). Die MFD zeigte einen positiven, aber ebenfalls nicht signifikanten Effekt ($\beta = 0.33$, $p = .057$), und der Pupillendurchmesser war positiv, jedoch nicht signifikant mit dem Schönheitsrating verknüpft ($\beta = 0.21$, $p = .181$). Demgegenüber sagten die Bildkategorien Natur ($\beta = 0.68$, $p = .002$) und Stilleben ($\beta = 0.51$, $p = .030$) im LMM signifikant höhere Schönheitsratings im Vergleich zur Referenzkategorie Individuen voraus, während Menschenmengen keinen Unterschied zur Referenzkategorie zeigte ($\beta = 0.13$, $p = .57$). Ein Likelihood-Ratio-Test zwischen einem Modell mit und ohne Bildkategorie zeigte, dass die Bildkategorie als Faktor signifikant zur Modellgüte beitrug ($\chi^2(3) = 7.94$, $p = .047$).

Insgesamt sprechen die linearen (Misch-) Modelle dafür, dass die klassischen Fixationsmaße (TFD, FC, MFD) nur einen kleinen Anteil der Varianz im Schönheitsrating erklären und ihre Effekte im LMM, nach expliziter Modellierung interindividueller Unterschiede, weitgehend nicht signifikant sind. TFD und FC zeigen im LMM keine signifikanten Zusammenhänge mit dem Schönheitsrating (alle $p \geq .23$), die MFD weist einen positiven, aber nicht signifikanten Effekt auf ($\beta = 0.33$, $p = .057$), und auch der Pupillendurchmesser ist positiv,

aber nicht signifikant ($\beta = 0.21, p = .181$). Die entsprechenden Schätzwerte der MLR und des LMM sind in Tabelle 3 zusammengefasst.

Tabelle 3

Lineare Modelle zum Zusammenhang von Schönheitsrating, Bildkategorie und Blickmetriken (z-standardisiert)

Prädiktor	MLR				LMM			
	<i>b</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>b</i>	<i>SE</i>	<i>z</i>	<i>p</i>
Intercept	-0.35	0.21	-1.66	.102	-0.35	0.21	-1.64	.102
Bildkategorie								
Stilleben	0.55	0.31	1.77	.080	0.51	0.51	2.17	.030
Menschenmengen	0.16	0.31	0.52	.607	0.13	0.23	0.57	.571
Natur	0.65	0.29	2.23	.028	0.68	0.22	3.16	.002
Blickmetriken								
TFD (min, z)	-1.77	0.57	-3.13	.002	-0.59	0.52	-1.14	.255
FC (z)	1.64	0.52	3.14	.002	0.57	0.48	1.21	.228
MFD (ms, z)	0.55	0.19	2.94	.004	0.33	0.17	1.91	.057
PD (mm, z)	0.20	0.10	1.98	.051	0.21	0.16	1.34	.181

Anmerkungen. MLR = multiple lineare Regression (Methode der kleinsten Quadrate); LMM = lineares gemischtes Modell mit zufälligen Intercepts für Teilnehmende. Alle kontinuierlichen Prädiktoren wurden z-standardisiert. TFD = Total Fixation Duration; FC = Fixationsanzahl; MFD = mittlere Fixationsdauer; PD = Pupillendurchmesser.

Referenzkategorie = Individuen. Für das MLR-Modell ergab sich $R^2 = .17$, $adj. R^2 = .10$, $F(7, 87) = 2.49$, $p = .022$.

Für das LMM ergab der Likelihood-Ratio-Test für den Effekt der Bildkategorie $\chi^2(3) = 7.94$, $p = .047$. $N = 95$

Beobachtungen, basierend auf den pro Person und Bildkategorie aggregierten Daten.

Im Zusammenhang mit F2 deutet dieses Muster darauf hin, dass die mittlere Fixationsdauer zwar deutlich zwischen den Bildkategorien unterscheidet (vgl. Abschnitt 4.1) und

in beiden Modellen positiv mit den Schönheitsratings zusammenhängt, im LMM jedoch keine Signifikanz erreicht (MLR: $\beta = 0.55$, $p = .004$; LMM: $\beta = 0.33$, $p = .057$). Auch der Zusammenhang von Pupillendurchmesser und Schönheitsrating ist in beiden Modellen positiv, jedoch statistisch nicht signifikant (MLR: $\beta = 0.20$, $p = .051$; LMM: $\beta = 0.21$, $p = .181$).

Da H2a und H2b ausdrücklich für das lineare gemischte Modell formuliert wurden, ist dieses Modell für ihre Bewertung maßgeblich. Beide Hypothesen gelten daher formal als nicht bestätigt.

5. Diskussion

5.1 Interpretation der Ergebnisse und theoretische Einordnung

Die Ergebnisse zeigen kategoriespezifische Unterschiede im Blickverhalten, insbesondere hinsichtlich Gesamtbetrachtungszeit, Fixationsanzahl und mittlerer Fixationsdauer. Diese Muster lassen sich als Unterschiede in sozialer Relevanz und explorativer Verarbeitung interpretieren: So scheinen soziale Motive und Naturdarstellungen im Museum stärkere visuelle Zuwendung hervorzurufen als Stillleben, während Individuen-Darstellungen auf eine vertiefte lokale Verarbeitung hindeuten. Damit stimmen die Befunde insgesamt mit theoretischen Annahmen überein, wonach biologisch und sozial relevante Reize Aufmerksamkeit bevorzugt binden und die Art der visuellen Exploration mitprägen. Diese kategorialen Unterschiede im Blickverhalten fügen sich in Befunde der Aufmerksamkeits- und Sozialkognitionsforschung ein, die zeigen, dass Gesichter und Körper bereits in frühen Verarbeitungsphasen priorisiert werden und automatische Orientierungsreaktionen auslösen. Gleichzeitig legen die längeren Betrachtungszeiten für Natur-Darstellungen nahe, dass auch naturhafte Motive eine erhöhte explorative Verarbeitung anstoßen, möglicherweise weil sie sowohl visuelle Komplexität als auch positive affektive

Konnotationen vereinen. Stillleben scheinen demgegenüber eine geringere intrinsische Anziehungskraft zu besitzen: Sie werden eher kurz, selektiv und mit geringerer Fixationsdichte betrachtet, was auf eine funktionale, eher informationsorientierte als ästhetisch-genussorientierte Rezeption hindeuten könnte.

Bezüglich des Schönheitsratings zeigten sich kategorial differenzierte, insgesamt jedoch schwache Zusammenhänge zwischen Blickmetriken und Bewertung. Fixationsanzahl und Gesamtbetrachtungszeit korrelierten bei Natur-Gemälden positiv mit dem Schönheitsrating, bei Stillleben und Menschenmengen dagegen kaum oder negativ. Die mittlere Fixationsdauer wies insgesamt nur sehr geringe Zusammenhänge mit dem Rating auf, lediglich bei Individuen-Darstellungen fanden sich moderate positive Tendenzen. Diese Effekte blieben jedoch in gemischten Modellen nach Kontrolle individueller Unterschiede nicht bestehen und bewegen sich insgesamt im Bereich kleiner Effekte. In Übereinstimmung damit beschreibt die Forschung zu feldnahen Studien urbaner Ästhetik, dass zwischen der Betrachtungsdauer und der ästhetischen Bewertung keine konsistenten Zusammenhänge nachweisbar sind ([Chana et al., 2023](#)). Betrachtungszeit ist damit kein direkter Indikator ästhetischer Präferenz, sondern reflektiert eher das Ausmaß explorativer Verarbeitung.

Auch die Regressionsanalysen bestätigen, dass einfache Dauer- oder Häufigkeitsmaße des Blickverhaltens nur in begrenztem Ausmaß mit den Schönheitsratings zusammenhängen und deren Varianz entsprechend schwach erklären. Nach Einbezug interindividueller Unterschiede verloren TFD, FC und MFD ihre Signifikanz, während vor allem die Bildkategorie und tendenziell der Pupillendurchmesser relevant blieben. Die Modelle erklären nur einen kleinen Anteil der Varianz, was nahelegt, dass Blickverhalten allein keine hinreichende Grundlage für ästhetische Bewertungen bildet. Entscheidend scheint die Kombination aus Bildinhalt, individuellen Präferenzen und physiologischer Involvierung zu sein ([Leder et al., 2004](#)).

Vor diesem Hintergrund sind die Gaze-Profile eher Signaturen unterschiedlicher Betrachtungsmodi als direkte Marker von Schönheit. Dass die Zusammenhänge zwischen Blickmetriken und Schönheitsratings im LMM nicht mehr signifikant sind, unterstreicht, dass stabile, personenspezifische Unterschiede das Blick- und Bewertungsverhalten prägen. Bei Natur-Werken zeigten sich positive Zusammenhänge zwischen explorativen Blickmaßen und Schönheitsrating. Bei Stilleben fielen die Zusammenhänge schwächer und uneinheitlicher aus. Damit zeigt sich, dass der Zusammenhang zwischen Betrachtungszeit und Gefallen kontextabhängig ist und nicht verallgemeinert werden kann. Der Kunstrahmen des Museums dürfte diesen Effekt zusätzlich moduliert haben, da Kontexteinflüsse auf ästhetische Bewertung belegt sind und negative Inhalte oft positiver bewertet werden, wenn sie als Kunst deklariert sind (Gerger et al., 2014). Insgesamt stützen die Befunde die Annahme, dass Blickmetriken nur in Relation zu spezifischen Inhalten und Kontextbedingungen informativ sind, nicht jedoch als universelle Marker ästhetischer Präferenz.

5.2 Die Rolle der Physiologie

Die Pupillometrie erwies sich in dieser Studie nicht als verlässlicher Indikator ästhetischer Erfahrung. Der Pupillendurchmesser war in den Korrelationsanalysen und in beiden Regressionsmodellen zwar durchgängig positiv gerichtet, die Effekte blieben jedoch klein und statistisch nicht signifikant (MLR: $\beta \approx 0.20$, $p \approx .051$; LMM: $\beta \approx 0.21$, $p \approx .181$). Dies relativiert neuroästhetische Perspektiven, die ästhetische Erfahrung als verkörperten Prozess mit physiologischer Beteiligung verstehen (Faust et al., 2019).

Befunde zu emotionalen Bildern und Stimuli zeigen, dass stärkere affektive Bedeutsamkeit mit ausgeprägterer Pupillenreaktion einhergeht, Pupillenreaktionen jedoch multimodal beeinflusst werden und nicht exklusiv ästhetisches Gefallen abbilden. Helligkeit,

kognitive Anstrengung, Überraschung oder Unsicherheit können ebenfalls Pupillenweitung verursachen.

Vor diesem Hintergrund wiegen insbesondere die Einschränkungen der Messsituation im Feld schwer. Die Erhebung erfolgte unter realen Lichtbedingungen des Museums. Helligkeit, Reflexionen und Kontraste variierten zwischen den Gemälden, sodass der Pupillendurchmesser teilweise durch Leuchtdichte und Lichtquellen mitbestimmt worden sein könnte ([Castellotti et al., 2020](#)). Zugleich zeigen experimentelle Arbeiten, dass Pupillenreaktionen selbst bei angeglicherer Luminanz systematisch durch Aufmerksamkeitszustand und die Verarbeitung bedeutsamer Informationen sowie andere Reizmerkmale wie Orientierung beeinflusst werden ([Kang et al., 2014](#); [Parrella et al., 2025](#)). Somit bleibt die Pupillenreaktion im Feld ein komplexes Signal, das neben ästhetischem Gefallen auch kognitive Anstrengung, Überraschung oder Aufmerksamkeitswechsel widerspiegeln kann ([Laeng et al., 2012](#)). Erschwerend kommt hinzu, dass keine standardisierten Baseline-Messungen durchgeführt wurden, wodurch Unterschiede im Ausgangsniveau nicht getrennt von phasischen Reaktionen auf die Kunstwerke geschätzt werden konnten.

Für die empirische Ästhetik im Feld eröffnet dies mehrere Perspektiven. Pupillenmessungen können auch unter realen Museumsbedingungen als ergänzendes Instrument eingesetzt werden, sollten aber in multimodale Designs mit weiteren physiologischen und subjektiven Indikatoren integriert werden. Durch die Kombination von Pupillometrie mit anderen physiologischen Maßen und expliziten Berichten lassen sich gemeinsame physiologische Reaktionsmuster identifizieren und licht- oder aufgabenbedingte Effekte besser abgrenzen. Der Vergleich von MLR- und LMM-Ergebnissen verdeutlicht zudem, dass stabile personenspezifische Unterschiede im individuellen Reaktionsniveau einen wesentlichen Teil der Varianz ausmachen und ohne explizite Modellierung zu Überschätzungen der Zusammenhänge

mit Schönheitsratings führen können. Insgesamt deuten die Ergebnisse nicht darauf hin, dass der Pupillendurchmesser als Maßwert für Schönheitserleben fungiert. Stattdessen lässt er sich als empfindlicher, jedoch unspezifischer physiologischer Indikator affektiv-bewertender Involvierung verstehen, der einen begrenzten Beitrag zu einem multimodalen Verständnis ästhetischer Prozesse leistet.

5.3 Methodische Reflexionen und Limitationen

Die vorliegenden Befunde sind vor dem Hintergrund mehrerer methodischer Einschränkungen zu interpretieren. Eine zentrale Limitation betrifft das Stimulusset und die Präsentationsbedingungen. Die verwendeten Gemälde unterschieden sich in ihrer physischen Größe, Detaildichte und Anzahl pro Kategorie; zudem wurden sie im bestehenden Hängungskontext des Museums präsentiert. Diese Unterschiede führen dazu, dass Intensitätsmaße wie TFD und FC nicht nur das Interesse an einem Werk, sondern auch die verfügbare Bildfläche und die Expositionshäufigkeit widerspiegeln und daher nicht ausschließlich als reiner Einfluss der Bildinhalte zu interpretieren sind. Insbesondere die Beobachtung, dass Menschenmengen-Darstellungen die höchste TFD und FC aufweisen, kann daher nicht ausschließlich auf ihre soziale Relevanz zurückgeführt werden, sondern ist zumindest teilweise durch Stimulus- und Kontextfaktoren mitbestimmt. Diese Maße sollten als Beschreibung unterschiedlicher Betrachtungsintensität und nicht als direkte Indikatoren ästhetischer Relevanz interpretiert werden. Zudem unterscheidet sich die Wahrnehmung im Museumsraum grundsätzlich von der im Labor. Sie begünstigt längere, intensivere Betrachtung und ist Teil einer verkörperten Interaktion mit Raum und Werk ([Kühnapfel et al., 2024](#)).

Zudem ist zu berücksichtigen, dass die Betrachtung von Kunstwerken im Museum nicht auf das Werk selbst beschränkt bleibt, sondern regelmäßig durch kontextuelle Informationen

ergänzt wird. Dies ist auch vor dem Hintergrund des Prozessmodells ästhetischer Erfahrung von [Leder et al. \(2004\)](#) bedeutsam, wonach ästhetische Urteile nicht allein aus der unmittelbaren visuellen Verarbeitung eines Objekts hervorgehen, sondern durch Vorwissen, Kontext, Gedächtnisintegration und explizite Klassifikation mitgeprägt werden. [Reitstätter et al. \(2020\)](#) zeigen, dass Besucher*innen einen erheblichen Teil ihrer Betrachtungszeit auf Labels und andere Begleitinformationen verwenden und dass diese Informationen die ästhetische Wahrnehmung sowie die Dauer und Struktur der Werkbetrachtung mitprägen. Für die vorliegende Untersuchung bedeutet dies, dass ein Teil der ästhetischen Verarbeitung somit außerhalb der erfassten Blickbewegungen stattgefunden haben könnte. Dies ist insbesondere für die Interpretation der Zusammenhänge zwischen Blickmetriken und Schönheitsratings relevant. Wenn ästhetische Urteile nicht ausschließlich während der direkten Betrachtung des Gemäldes entstehen, sondern auch durch begleitende Informationen, Vorwissen und situative Einbettung mitgeformt werden, kann die Vorhersagekraft werkbezogener Blickmaße systematisch unterschätzt werden. Die in der vorliegenden Studie beobachteten schwachen Zusammenhänge zwischen Blickverhalten und Schönheitsurteil sind daher nicht notwendigerweise allein als Ausdruck geringer Kopplung zwischen visueller Exploration und ästhetischer Bewertung zu verstehen, sondern möglicherweise auch als Folge einer nur teilweisen Erfassung des tatsächlichen Rezeptionsprozesses.

Eine weitere Einschränkung betrifft die Stichprobe und die statistische Power. Die Eye-Tracking-Analysen basieren auf 25 Personen, die aus einem größeren Studienkontext selektiert wurden und überwiegend Studierende repräsentieren. Auch wenn diese Stichprobengröße für explorative Feldstudien im Museumssetting üblich ist, begrenzt sie die Sensitivität für kleinere Effekte, insbesondere in Modellen mit mehreren Prädiktoren und kategorialen Faktoren. Nicht signifikante Ergebnisse, etwa der fehlende Haupteffekt der

Bildkategorie auf den Pupillendurchmesser oder die ausbleibende Signifikanz der Blickmetriken im LMM, sind daher nicht als Hinweis auf das Fehlen relevanter Effekte zu verstehen, sondern könnten teilweise auch auf eine begrenzte Teststärke zurückzuführen sein. Eine Replikation mit größerer und heterogenerer Stichprobe wäre notwendig, um die beobachteten Muster robuster zu prüfen und mögliche Interaktionseffekte zwischen Bildkategorie, Blickverhalten und Bewertung verlässlicher zu erfassen.

Schließlich ist zu berücksichtigen, dass die Studie in einem expliziten Kunstrahmen mit einer klaren Bewertungsaufgabe durchgeführt wurde. Die Teilnehmenden wurden gebeten, während der Betrachtung kontinuierlich die Schönheit der Werke zu bewerten, was die Aufmerksamkeit auf evaluative Aspekte lenken und spontane Rezeptionsprozesse modifizieren kann. Der kontinuierliche Bewertungsmodus könnte dabei sowohl das Blickverhalten als auch die Pupillenreaktion beeinflussen, etwa durch zusätzliche kognitive Belastung oder durch stärkere Fokussierung auf bestimmte Bildmerkmale, die für das Urteil relevant erscheinen. Zukünftige Arbeiten könnten hier mit variierenden Aufgaben (z. B. freie Betrachtung ohne Rating vs. explizite Bewertung) oder zeitlich getrennten Messphasen arbeiten, um die Effekte der Bewertungsanforderung auf Blick- und Pupillenmaße systematischer zu untersuchen.

Insgesamt stellt die Studie einen ersten feldnahen Schritt zur Untersuchung der Kopplung von Blickverhalten, Pupillenreaktion und ästhetischer Bewertung im Museum dar. Sie zeigt, dass kategoriespezifische Betrachtungsmodi und entsprechende physiologische Korrelate trotz realweltlicher Variabilität erfassbar sind. Zugleich wird deutlich, dass größere, stärker kontrollierte und methodisch erweiterte Untersuchungen notwendig sind, um die identifizierten Muster zu bestätigen und weiter auszudifferenzieren.

6. Fazit und Ausblick

Die vorliegende Arbeit untersuchte, in welchem Ausmaß Blickverhalten und Pupillenreaktionen im Museumsraum Aufschluss über das Schönheitserleben beim Betrachten realer Gemälde geben. Im Zentrum standen vier Bildkategorien (Individuen, Menschenmengen, Natur und Stilleben) sowie die Frage, ob sich aus Blickmetriken und Pupillendurchmesser Rückschlüsse auf explizite Schönheitsurteile ziehen lassen. Die Ergebnisse zeigen zunächst, dass sich die untersuchten Bildkategorien in ihrem Blickprofil systematisch unterscheiden. Damit erweisen sich die eingesetzten Eye-Tracking-Maße als geeignet, unterschiedliche Modi visueller Exploration im Museumssetting sichtbar zu machen. Für die Vorhersage expliziter Schönheitsurteile war ihre Aussagekraft jedoch insgesamt begrenzt. Auch der Pupillendurchmesser zeigte zwar tendenziell positive, aber keine signifikanten Zusammenhänge mit den Schönheitsratings.

Ein zentraler Befund der Arbeit liegt somit in der Differenz zwischen der deskriptiven Sensitivität der Blickmetriken für bildkategoriespezifische Betrachtungsmuster und ihrer vergleichsweise geringen Eignung als Prädiktoren individueller Schönheitsurteile. Blickverhalten scheint im Feld vor allem Informationen darüber zu liefern, wie unterschiedliche Bildtypen betrachtet werden, nicht jedoch in hinreichendem Maß darüber, wie schön ein Werk erlebt wird. Diese Differenz verweist darauf, dass ästhetische Bewertung im Museumsraum nicht auf die unmittelbare visuelle Exploration reduziert werden kann. Vielmehr sprechen die Befunde dafür, dass explizite Schönheitsurteile durch ein Zusammenspiel aus Bildinhalt, situativem Kontext, individuellen Präferenzen und personenspezifischen Verarbeitungsweisen entstehen. Damit stehen die Ergebnisse im Einklang mit prozessorientierten Modellen ästhetischer Erfahrung, nach

denen Wahrnehmung, Gedächtnisintegration, Kontextverarbeitung und bewertende Urteilsbildung eng miteinander verflochten sind.

Zugleich verdeutlicht die Arbeit die Grenzen einer direkten Übertragung von Laborbefunden auf feldnahe Museumssituationen. Während kontrollierte Laborstudien häufig engere Kopplungen zwischen Blickverhalten und Präferenz berichten, fielen die Zusammenhänge unter natürlichen Rezeptionsbedingungen in der vorliegenden Untersuchung schwächer und instabiler aus. Dies lässt sich plausibel damit erklären, dass das Museumssetting eine komplexere Form ästhetischer Erfahrung hervorbringt: Die Betrachtung erfolgt selbstgesteuert, eingebettet in räumliche Orientierung, Kontextinformationen, individuelle Bewegungsmuster und wechselnde Aufmerksamkeitsfoki. Eye-Tracking erfasst in diesem Rahmen zwar zentrale Aspekte visueller Zuwendung, bildet das Schönheitserleben jedoch nur ausschnittshaft ab.

Für die Interpretation der Pupillometrie ergibt sich ein ähnliches Bild. Der Pupillendurchmesser erwies sich in der vorliegenden Studie nicht als verlässlicher Marker expliziter ästhetischer Bewertung. Unter Feldbedingungen ist seine Aussagekraft zusätzlich dadurch eingeschränkt, dass pupilläre Reaktionen nicht nur mit affektiver Bedeutsamkeit, sondern auch mit Luminanz, Aufmerksamkeitswechseln und kognitiver Beanspruchung zusammenhängen. Für die empirische Ästhetik bedeutet dies, dass Pupillenmaße im Museumsraum sinnvoll nutzbar sein können, jedoch nicht isoliert interpretiert werden sollten.

Aus diesen Befunden ergeben sich mehrere Perspektiven für die weitere Forschung. Erstens erscheinen größere und heterogenere Stichproben notwendig, um kleinere Effekte zuverlässiger zu erfassen und interindividuelle Unterschiede systematischer zu modellieren. Zweitens sollten feldnahe Designs verstärkt kontextbezogene Variablen einbeziehen, etwa die Nutzung von Labels, räumliche Bewegungsmuster oder Unterschiede in der Präsentationssituation, um den tatsächlichen Rezeptionsprozess vollständiger abzubilden.

Drittens wäre es sinnvoll, Blick- und Pupillendaten stärker mit weiteren physiologischen, verhaltensbezogenen und subjektiven Maßen zu kombinieren, um ästhetische Erfahrung als mehrdimensionalen Prozess zu erfassen. Besonders aufschlussreich dürften darüber hinaus direkte Vergleiche zwischen Labor- und Museumssettings sein, da sie präziser bestimmen könnten, welche Zusammenhänge zwischen Blickverhalten, Physiologie und Schönheitserleben kontextstabil sind und welche wesentlich vom Rezeptionsrahmen abhängen.

Insgesamt leistet die Arbeit einen feldnahen Beitrag zur empirischen Ästhetik, indem sie zeigt, dass mobile Eye-Tracking-Verfahren auch im Museum differenzierte Muster visueller Exploration erfassen können. Zugleich macht sie deutlich, dass solche Maße explizite ästhetische Urteile nicht ersetzen, sondern nur als ergänzende Indikatoren innerhalb eines umfassenderen Verständnisses ästhetischer Erfahrung interpretiert werden sollten. Der Gewinn künftiger Forschung dürfte daher weniger in der Suche nach einem einzelnen impliziten Marker des Schönheitserlebens liegen als in der Entwicklung integrativer Modelle, die Blickverhalten, Physiologie, Kontext und subjektive Bewertung gemeinsam berücksichtigen.

Literaturverzeichnis

- Batool, A., Rutherford, P., McGraw, P., Ledgeway, T., & Altomonte, S. (2022). Gaze correlates of view preference: Comparing natural and urban scenes. *Lighting Research & Technology*, 54(6), 576–594. <https://doi.org/10.1177/14771535211055703>
- Beder, D., Pelowski, M., & Imamoğlu, Ç. (2024). Influence of complexity and gestalt principles on aesthetic preferences for building façades: An eye tracking study. *Journal of Eye Movement Research*, 17(2). <https://doi.org/10.16910/jemr.17.2.4>
- Belvedere. (o. J.). *Sammlung Online*. Belvedere Museum Wien. Abgerufen 14. April 2026, von <https://sammlung.belvedere.at/start>
- Bertamini, M., Rampone, G., Makin, A. D. J., & Jessop, A. (2019). Symmetry preference in shapes, faces, flowers and landscapes. *PeerJ*, 7, e7078. <https://doi.org/10.7717/peerj.7078>
- Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Social attention and real-world scenes: The roles of action, competition and social content. *Quarterly Journal of Experimental Psychology*, 61(7), 986–998. <https://doi.org/10.1080/17470210701410375>
- Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research*, 49(24), 2992–3000. <https://doi.org/10.1016/j.visres.2009.09.014>
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45(4), 602–607. <https://doi.org/10.1111/j.1469-8986.2008.00654.x>
- Brieber, D., Nadal, M., Leder, H., & Rosenberg, R. (2014). Art in time and space: Context modulates the relation between art experience and viewing time. *PLoS ONE*, 9(6), e99019. <https://doi.org/10.1371/journal.pone.0099019>

- Castellotti, S., Conti, M., Feitosa-Santana, C., & Del Viva, M. M. (2020). Pupillary response to representations of light in paintings. *Journal of Vision*, 20(10), 14.
<https://doi.org/10.1167/jov.20.10.14>
- Chana, K., Mikuni, J., Schnebel, A., & Leder, H. (2023). Reading in the city: Mobile eye-tracking and evaluation of text in an everyday setting. *Frontiers in Psychology*, 14, 1205913. <https://doi.org/10.3389/fpsyg.2023.1205913>
- Cheng, Y., Liu, W., Yuan, X., & Jiang, Y. (2021). The eyes have it: Perception of social interaction unfolds through pupil dilation. *Neuroscience Bulletin*, 37(11), 1595–1598.
<https://doi.org/10.1007/s12264-021-00739-z>
- Elschner, S. G., Hübner, R., & Dambacher, M. (2018). Do fluency-induced pupillary responses reflect aesthetic affect? *Psychology of Aesthetics, Creativity, and the Arts*, 12(3), 294–303. <https://doi.org/10.1037/aca0000139>
- emoTouch. (2025). *emoTouch Web [Webanwendung]*. <https://www.emotouch.de/de/home>
- Faust, N. T., Chatterjee, A., & Christopoulos, G. I. (2019). Beauty in the eyes and the hand of the beholder: Eye and hand movements' differential responses to facial attractiveness. *Journal of Experimental Social Psychology*, 85, 103884.
<https://doi.org/10.1016/j.jesp.2019.103884>
- Fechner, G. T. (1876). *Vorschule der Aesthetik*. Breitkopf & Härtel.
- Franěk, M., Petružálek, J., & Šefara, D. (2019). Eye movements in viewing urban images and natural images in diverse vegetation periods. *Urban Forestry & Urban Greening*, 46, 126477. <https://doi.org/10.1016/j.ufug.2019.126477>
- Franěk, M., Šefara, D., Petružálek, J., Cabal, J., & Myška, K. (2018). Differences in eye movements while viewing images with various levels of restorativeness. *Journal of Environmental Psychology*, 57, 10–16. <https://doi.org/10.1016/j.jenvp.2018.05.001>

- Gerger, G., Leder, H., & Kremer, A. (2014). Context effects on emotional and aesthetic evaluations of artworks and IAPS pictures. *Acta Psychologica*, 151, 174–183.
<https://doi.org/10.1016/j.actpsy.2014.06.008>
- Goller, J., Mitrovic, A., & Leder, H. (2019). Effects of liking on visual attention in faces and paintings. *Acta Psychologica*, 197, 115–123. <https://doi.org/10.1016/j.actpsy.2019.05.008>
- Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, 56(4), 544–549.
- Henderson, R. R., Bradley, M. M., & Lang, P. J. (2018). Emotional imagery and pupil diameter. *Psychophysiology*, 55(6), e13050. <https://doi.org/10.1111/psyp.13050>
- Holmes, T., & Zanker, J. M. (2012). Using an oculomotor signature as an indicator of aesthetic preference. *I-Perception*, 3(7), 426–439. <https://doi.org/10.1068/i0448aap>
- Jacobs, M. (2011). Psychology of the visual landscape. *Research in Urbanism Series*, 41–54.
<https://doi.org/10.7480/RIUS.2.206>
- Jankowski, T., Francuz, P., Oleś, P., Chmielnicka-Kuter, E., & Augustynowicz, P. (2020). The effect of painting beauty on eye movements. *Advances in Cognitive Psychology*, 16(3), 213–227. <https://doi.org/10.5709/acp-0298-4>
- Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40(1–2), 1–19.
[https://doi.org/10.1016/0010-0277\(91\)90045-6](https://doi.org/10.1016/0010-0277(91)90045-6)
- Kang, O. E., Huffer, K. E., & Wheatley, T. P. (2014). Pupil dilation dynamics track attention to high-level information. *PLoS ONE*, 9(8), e102463.
<https://doi.org/10.1371/journal.pone.0102463>

- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1476), 2109–2128. <https://doi.org/10.1098/rstb.2006.1934>
- Khalighy, S., Green, G., Scheepers, C., & Whittet, C. (2015). Quantifying the qualities of aesthetics in product design using eye-tracking technology. *International Journal of Industrial Ergonomics*, 49, 31–43. <https://doi.org/10.1016/j.ergon.2015.05.011>
- Kuchinke, L., Trapp, S., Jacobs, A. M., & Leder, H. (2009). Pupillary responses in art appreciation: Effects of aesthetic emotions. *Psychology of Aesthetics, Creativity, and the Arts*, 3(3), 156–163. <https://doi.org/10.1037/a0014464>
- Kühnapfel, C., Fingerhut, J., Brinkmann, H., Ganster, V., Tanaka, T., Specker, E., Mikuni, J., Güldenpfennig, F., Gartus, A., Rosenberg, R., & Pelowski, M. (2024). How do we move in front of art? How does this relate to art experience? Linking movement, eye tracking, emotion, and evaluations in a gallery-like setting. *Empirical Studies of the Arts*, 42(1), 86–146. <https://doi.org/10.1177/02762374231160000>
- Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, 7(1), 18–27. <https://doi.org/10.1177/1745691611427305>
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95(4), 489–508. <https://doi.org/10.1348/0007126042369811>
- Leder, H., Mitrovic, A., & Goller, J. (2016). How beauty determines gaze! Facial attractiveness and gaze duration in images of real world scenes. *I-Perception*, 7(4), 2041669516664355. <https://doi.org/10.1177/2041669516664355>

- Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode – Developments and challenges in empirical aesthetics. *British Journal of Psychology*, 105(4), 443–464. <https://doi.org/10.1111/bjop.12084>
- Leder, H., Tinio, P. P., Fuchs, I. M., & Bohrn, I. (2010). When attractiveness demands longer looks: The effects of situation and gender. *Quarterly Journal of Experimental Psychology*, 63(9), 1858–1871. <https://doi.org/10.1080/17470211003605142>
- Marin, M. M., & Leder, H. (2013). Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music. *PLoS ONE*, 8(8), e72412. <https://doi.org/10.1371/journal.pone.0072412>
- Marin, M. M., & Leder, H. (2022). Gaze patterns reveal aesthetic distance while viewing art. *Annals of the New York Academy of Sciences*, 1514(1), 155–165. <https://doi.org/10.1111/nyas.14792>
- Massaro, D., Savazzi, F., Di Dio, C., Freedberg, D., Gallese, V., Gilli, G., & Marchetti, A. (2012). When art moves the eyes: A behavioral and eye-tracking study. *PLoS ONE*, 7(5), e37285. <https://doi.org/10.1371/journal.pone.0037285>
- Mitrovic, A., Tinio, P. P. L., & Leder, H. (2016). Consequences of beauty: Effects of rater sex and sexual orientation on the visual exploration and evaluation of attractiveness in real-world scenes. *Frontiers in Human Neuroscience*, 10, Article 122. <https://doi.org/10.3389/fnhum.2016.00122>
- Mitschke, V., Goller, J., & Leder, H. (2017). Exploring everyday encounters with street art using a multimethod design. *Psychology of Aesthetics, Creativity, and the Arts*, 11(3), 276–283. <https://doi.org/10.1037/aca0000131>
- Nadal, M., & Ureña, E. (2021). One hundred years of empirical aesthetics: Fechner to Berlyne (1876–1976). *PsyArXiv*. <https://doi.org/10.31234/osf.io/c92y7>

- Parrella, M. A., Gurung, I., & Grubb, M. A. (2025). Luminance matching in cognitive pupillometry is not enough: The curious case of orientation. *eNeuro*, 12(11), ENEURO.0238-25.2025. <https://doi.org/10.1523/ENEURO.0238-25.2025>
- Pelowski, M., Forster, M., Tinio, P. P. L., Scholl, M., & Leder, H. (2017). Beyond the lab: An examination of key factors influencing interaction with ‘real’ and museum-based art. *Psychology of Aesthetics, Creativity, and the Arts*, 11(3), 245–264. <https://doi.org/10.1037/aca0000141>
- Pelowski, M., Leder, H., Mitschke, V., Specker, E., Gerger, G., Tinio, P. P. L., Vaporova, E., Bieg, T., & Husslein-Arco, A. (2018). Capturing aesthetic experiences with installation art: An empirical assessment of emotion, evaluations, and mobile eye tracking in Olafur Eliasson’s “Baroque, Baroque!”. *Frontiers in Psychology*, 9, 1255. <https://doi.org/10.3389/fpsyg.2018.01255>
- Reitstätter, L., Brinkmann, H., Santini, T., Specker, E., Dare, Z., Bakondi, F., Miscená, A., Kasneci, E., Leder, H., & Rosenberg, R. (2020). The display makes a difference: A mobile eye tracking study on the perception of art before and after a museum’s rearrangement. *Journal of Eye Movement Research*, 13(2). <https://doi.org/10.16910/jemr.13.2.6>
- Ricou, C., Rabadan, V., Mofid, Y., Aguilon-Hernandez, N., & Wardak, C. (2024). Pupil dilation reflects the social and motion content of faces. *Social Cognitive and Affective Neuroscience*, 19(1), nsae055. <https://doi.org/10.1093/scan/nsae055>
- Rosenberg, R., & Klein, C. (2015). The moving eye of the beholder: Eye tracking and the perception of paintings. In J. P. Huston, M. Nadal, F. Mora, L. F. Agnati, & C. J. C. Conde (Hrsg.), *Art, Aesthetics, and the Brain* (S. 79–108). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199670000.003.0005>

- Santini, T., Brinkmann, H., Reitstätter, L., Leder, H., Rosenberg, R., Rosenstiel, W., & Kasneci, E. (2018). The art of pervasive eye tracking: Unconstrained eye tracking in the Austrian Gallery Belvedere. *Proceedings of the 7th Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction*, 1–8. <https://doi.org/10.1145/3208031.3208032>
- Skripkauskaite, S., Mihai, I., & Koldewyn, K. (2023). Attentional bias towards social interactions during viewing of naturalistic scenes. *Quarterly Journal of Experimental Psychology*, 76(10), 2303–2311. <https://doi.org/10.1177/17470218221140879>
- Specker, E., Tinio, P. P. L., & Van Elk, M. (2017). Do you see what I see? An investigation of the aesthetic experience in the laboratory and museum. *Psychology of Aesthetics, Creativity, and the Arts*, 11(3), 265–275. <https://doi.org/10.1037/aca0000107>
- Spee, B. T. M., Arato, J., Mikuni, J., Tran, U. S., Pelowski, M., & Leder, H. (2024). The dynamics of experiencing Gestalt and Aha in cubist art: Pupil responses and art evaluations show a complex interplay of task, stimuli content, and time course. *Frontiers in Psychology*, 15, 1192565. <https://doi.org/10.3389/fpsyg.2024.1192565>
- Sweeny, T. D., Haroz, S., & Whitney, D. (2013). Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *Journal of Experimental Psychology: Human Perception and Performance*, 39(2), 329–337. <https://doi.org/10.1037/a0028712>
- Tobii Pro. (2024). *Tobii Pro Glasses 3 User Manual* (Version 1.13). [User Manual]. Tobii AB. <https://go.tobii.com/tobii-pro-glasses-3-user-manual>
- Tobii Pro. (2022). *Tobii Pro Lab* (Version 1.207) [Computer software]. Tobii AB. https://go.tobii.com/tobii_pro_lab_user_manual

- Valtchanov, D., & Ellard, C. G. (2015). Cognitive and affective responses to natural scenes: Effects of low level visual properties on preference, cognitive load and eye-movements. *Journal of Environmental Psychology*, 43, 184–195.
<https://doi.org/10.1016/j.jenvp.2015.07.001>
- Valuch, C., Pflüger, L. S., Wallner, B., Laeng, B., & Ansorge, U. (2015). Using eye tracking to test for individual differences in attention to attractive faces. *Frontiers in Psychology*, 6, Article 42. <https://doi.org/10.3389/fpsyg.2015.00042>
- Villani, D., Morganti, F., Cipresso, P., Ruggi, S., Riva, G., & Gilli, G. (2015). Visual exploration patterns of human figures in action: An eye tracker study with art paintings. *Frontiers in Psychology*, 6, Article 1636. <https://doi.org/10.3389/fpsyg.2015.01636>
- Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, 69(1), 105–129. <https://doi.org/10.1146/annurev-psych-010416-044232>
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17(7), 592–598. <https://doi.org/10.1111/j.1467-9280.2006.01750.x>

Abbildungsverzeichnis

Abbildung 1 <i>Beispielgemälde der Kategorie „Individuen“</i>	10
Abbildung 2 <i>Beispielgemälde der Kategorie „Menschenmengen“</i>	12
Abbildung 3 <i>Beispielgemälde der Kategorie „Natur“</i>	13
Abbildung 4 <i>Beispielgemälde der Kategorie „Stilleben“</i>	15
Abbildung 5 <i>Der Versuchsraum „Spielezimmer“</i>	22
Abbildung 6 <i>Boxplots der Blickmetriken nach Bildkategorie</i>	29
Abbildung 7 <i>Z-standardisierte Blickprofile der vier Bildkategorien</i>	30
Abbildung 8 <i>Streudiagramme zwischen Schönheitsrating und Blickmetriken</i>	33

Tabellenverzeichnis

Tabelle 1 <i>Deskriptive Statistik der Blickmetriken nach Bildkategorie</i>	28
Tabelle 2 <i>Korrelationen zwischen Schönheitsrating und Blickmetriken</i>	32
Tabelle 3 <i>Lineare Modelle zur Vorhersage des Schönheitsratings</i>	36

Anhang: Stimulusmaterial

Tabelle A1

Deskriptive Statistiken der Schönheitsratings für die Bildkategorie „Individuen“

Index	Künstler:in	Titel	Kategorie	Schönheits- bewertung		Gemälde
				M	SD	
B3	Franz	Der	Individuen	46.58	26.16	
	Sebald	Schlachtenmaler				
	Unterberger	August Querfurt				
	(1752)					
B7	Franz	Unterhaltung im	Individuen	45.80	28.44	
	Christoph	Freien mit Tanz				
	Janneck					
	(um 1740)					

B8 Franz [Unterhaltung im](#) Individuen 45.84 26.78
 Christoph [Freien mit Wein](#)
 Janneck (um [und Gesang](#)
 1740)



C1 Adam [Der galante Helfer](#) Individuen 48.34 26.47
 Johann
 Braun
 (1785)



C3 Adam [Der Künstler in](#) Individuen 40.49 25.41
 Johann [seinem Atelier](#)
 Braun
 (undatiert)



C5	E. K.	Die	Individuen	42.99	23.96
	Lautter	Spitzenklöpplerin			
	(undatiert)				



Anmerkung. $N = 25$. Das Schönheitsrating wurde auf einer Skala von 0 (gar nicht schön) bis 100 (sehr schön) erhoben. M = Mittelwert, SD = Standardabweichung. Die Abbildungen dienen der Identifikation der verwendeten Gemälde aus der Online-Sammlung des Belvedere, Wien (Belvedere, o. J.).

Tabelle A2

Deskriptive Statistiken der Schönheitsratings für die Bildkategorie „Menschenmengen“

Index	Künstler:in	Titel	Kategorie	Schönheits- bewertung		Gemälde
				M	SD	
B1	August Querfurt (undatiert)	Beizerfolg auf Reiher	Menschen- mengen	51.65	25.28	
B2	August Querfurt (undatiert)	Aufbruch zur Beizjagd	Menschen- mengen	47.75	25.61	
B4	Karl Josef Aigen (undatiert)	Jahrmarkt vor einem Stadttor	Menschen- mengen	48.10	25.13	

B5 Karl Josef [Kirchweihfest](#) Menschen- 47.81 27.27
Aigen mengen
(undatiert)



B6 Franz de [Jahrmakrt mit](#) Menschen- 42.16 25.24
Paula Ferg [zinnenbekröntem](#) mengen
(1724) [Bergfried](#)



Anmerkung. $N = 25$. Das Schönheitsrating wurde auf einer Skala von 0 (gar nicht schön) bis 100 (sehr schön) erhoben. M = Mittelwert, SD = Standardabweichung. Die Abbildungen dienen der Identifikation der verwendeten Gemälde aus der Online-Sammlung des Belvedere, Wien (Belvedere, o. J.).

Tabelle A3

Deskriptive Statistiken der Schönheitsratings für die Bildkategorie „Natur“

Index	Künstler:in	Titel	Kategorie	Schönheits- bewertung		Gemälde
				M	SD	
A1	Johann Christian Brand (1758)	Laxenburg vom Münkendorfer Lusthaus gegen Südwest	Natur	53.80	21.85	
A2	Johann Christian Brand (1758)	Laxenburg vom Lusthaus auf der Hanawiese gegen Mödling	Natur	55.45	23.67	
A3	Johann Christian Brand (1787)	Aussicht von der Albrechtsburg in Klosterneuburg	Natur	52.83	22.03	

A4	Johann Christian Brand (1790)	<u>Donaulandschaft, vom Bisamberg aus gesehen</u>	Natur	50.30	24.21	
A5	Josef Heideloff (1779)	<u>Weinernte vor Wien mit Blick nach Klosterneuburg, den Leopoldsberg, Bisamberg und Korneuburg</u>	Natur	47.63	25.89	
A6	Johann Christian Brand (1758)	<u>Laxenburg vom Hayd-Lusthaus gegen Maria Lanzendorf</u>	Natur	48.99	27.20	
A7	Johann Christian Brand (1759)	<u>Laxenburg von der Schneiderau gegen Guntramsdorf und Mödling</u>	Natur	53.04	26.71	

D1 David [Ideale Landschaft](#) Natur 58.19 24.93
 Richter
 d.Ä.
 (1700/
 1735)



D2 David [Ideale Landschaft](#) Natur 57.63 26.80
 Richter [\(Abendstimmung\)](#)
 d.Ä.
 (1700/
 1735)



D3 Franz [Blumen und](#) Natur 59.09 27.67
 Werner [Früchte](#)
 Tamm
 (1715)



Anmerkung. $N = 25$. Das Schönheitsrating wurde auf einer Skala von 0 (gar nicht schön) bis 100 (sehr schön) erhoben. M = Mittelwert, SD = Standardabweichung. Die Abbildungen dienen der Identifikation der verwendeten Gemälde aus der Online-Sammlung des Belvedere, Wien (Belvedere, o. J.).

Tabelle A4

Deskriptive Statistiken der Schönheitsratings für die Bildkategorie „Stilleben“

Index	Künstler:in	Titel	Kategorie	Schönheits- bewertung		Gemälde
				M	SD	
C2	Unbekannter Künstler (1640/1700)	Stilleben mit Käse	Stilleben	45.64	25.44	
C6	Anna Maria Punz (1754)	Stilleben mit Krug und Tulpe	Stilleben	50.66	24.54	

C8	Anna Maria Punz (1754)	Stillleben mit Küchengeschirr, Zwiebel und Kohlrabi	Stillleben	48.94	26.46
----	---------------------------	--	------------	-------	-------



C9	Joannes de Cordua (undatiert)	Stillleben mit Kupfergeschirr	Stillleben	47.22	26.96
----	-------------------------------------	----------------------------------	------------	-------	-------



C10	Anna Maria Punz (1754)	Stillleben mit Apfelkorb	Stillleben	49.52	28.35
-----	---------------------------	-----------------------------	------------	-------	-------



Anmerkung. $N = 25$. Das Schönheitsrating wurde auf einer Skala von 0 (gar nicht schön) bis 100 (sehr schön) erhoben. M = Mittelwert, SD = Standardabweichung. Die Abbildungen dienen der Identifikation der verwendeten Gemälde aus der Online-Sammlung des Belvedere, Wien (Belvedere, o. J.).

Tabelle A5

Deskriptive Statistiken der Schönheitsratings für die aus der weiteren Analyse exkludierten Gemälde

Index	Künstler:in	Titel	Kategorie	Schönheits- bewertung		Gemälde
				M	SD	
C 4	Felice Boselli (undatiert)	Jagdstillleben mit Katze	Aus Analyse exkludiert	42.98	25.66	
C 7	Martin Ferdinand Quadal (1781)	Katzen und Fische	Aus Analyse exkludiert	51.87	24.93	

Anmerkung. $N = 25$. Das Schönheitsrating wurde auf einer Skala von 0 (gar nicht schön) bis 100 (sehr schön) erhoben. M = Mittelwert, SD = Standardabweichung. Die Abbildungen dienen der Identifikation der verwendeten Gemälde aus der Online-Sammlung des Belvedere, Wien (Belvedere, o. J.).