



universität  
wien

# DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Das Ganze ist mehr als die Summe seiner Teile“

Chancen und Herausforderungen empirischer Zugänge zu den  
Kunstwissenschaften

verfasst von / submitted by  
Laura Commare

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of  
Doktorin der Philosophie (Dr. phil.)

Wien/ Vienna, 2018

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on the  
student record sheet:

A 792 315

Dissertationsgebiet lt. Studienblatt /  
field of study as it appears on the student record  
sheet:

Kunstgeschichte

Betreut von / Supervisor:

Univ.-Prof. Dr. Raphael Rosenberg

Betreut von / Supervisor:

Univ.-Prof. Dr. Helmut Leder

---

---

# DANKSAGUNG

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die zur Entstehung dieser Dissertation wesentlich beigetragen haben.

Für die Betreuung bedanke ich mich an erster Stelle bei Herrn Professor Raphael Rosenberg, der mit viel Rat, Zeit und Geduld die Entstehung dieser Arbeit begleitet hat. Auch bei meinem Co-Betreuer Herrn Professor Helmut Leder möchte ich mich bedanken. Seine psychologische Expertise war bei der Erarbeitung und Durchführung der empirischen Studien, die das Kernstück und die Analyseeinheiten dieser Arbeit stellen, unerlässlich.

Dem Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF), der dieses Projekt größtenteils finanziert hat, bin ich ebenfalls zu großem Dank verpflichtet. Darüber hinaus möchte ich mich ganz herzlich bei meiner Kollegin Hanna Brinkmann bedanken – für die langjährige gute und intensive Zusammenarbeit, für die gemeinsam ausgestandenen Höhen und Tiefen und für das allzeit hilfreiche Feedback. Vielen Dank auch an David Brieber, Johanna Aufreiter, Jane Boddy und Mario Thalwitzer für die vielen Gespräche und Anregungen.

Großer Dank gebührt meinen Eltern, Roman Hoffmann und Cornell Schreiber. Ihnen möchte ich diese Arbeit widmen, denn sie haben mein Vorhaben stets befürwortet, niemals an mir gezweifelt und mir immer mit einem offenen Ohr, Empathie und Humor beigestanden.

---



---

# INHALTSVERZEICHNIS

<b>Einleitung</b>	<b>6</b>
<b>I Empirische Ansätze in der Kunstwissenschaft</b>	<b>13</b>
1.1 Spielarten empirischer Ansätze in der Kunstwissenschaft	19
1.2 Allgemeine Klassifikationsmerkmale empirischer Ansätze	27
1.3 Chancen und Möglichkeiten empirischer Ansätze in der Kunstwissenschaft	32
1.4 Kritik und Probleme empirischer Ansätze in der Kunstwissenschaft	41
1.5 Zusammenschau der positiven und negativen Kritik	48
<b>II Studien</b>	<b>55</b>
2.1 More Than the Sum of Its Parts. Perceiving Complexity in Painting	59
<i>Einleitung</i>	60
<i>Zusammenfassung und Bemerkungen zur Studie</i>	68
<i>Studententext</i>	71
<i>Diskussion</i>	83
<i>Fazit und Beitrag der empirischen Methoden</i>	115
2.2 Abstract Art as a Universal Language?	120
<i>Einleitung</i>	121
<i>Bemerkungen zur Studie</i>	124
<i>Studententext</i>	125
<i>Diskussion</i>	127
<i>Fazit und Beitrag der empirischen Methoden</i>	133
2.3 Aesthetic Echoes in the Beholder's Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art	137 137
<i>Einleitung</i>	138
<i>Bemerkungen zur Studie</i>	140
<i>Studententext</i>	141
<i>Diskussion</i>	155
<i>Fazit und Beitrag der empirischen Methoden</i>	159
2.4 Zwischenfazit	163

---

<b>III Methodenreflexion</b>	<b>165</b>
3.1 Methodenübersicht	167
3.2 Datenerhebung	174
<i>Merkmale der Befragten</i>	175
<i>Merkmale der Fragen/Messung</i>	180
<i>Merkmale der Interview- bzw. Testsituation</i>	185
3.3 Auswertung und Interpretation	190
<i>Offene Fragen</i>	191
<i>Standardisierte Fragen</i>	192
<i>Eye-Tracking</i>	198
<i>Statistik</i>	204
<i>Publikationsform</i>	210
3.4 Zwischenfazit	212
<b>Diskussion</b>	<b>213</b>
REFERENZEN	234
ABBILDUNGSVERZEICHNIS	265
TABELLENVERZEICHNIS	269
ANHANG	271
<i>Stellungnahme zur Arbeitssteilung I</i>	272
<i>Stellungnahme zur Arbeitssteilung II</i>	273
<i>Stellungnahme zur Arbeitssteilung III</i>	274
<i>Stellungnahme zur Arbeitssteilung IV</i>	275
<i>Expertisefragebogen</i>	276
<i>Abstract (D)</i>	281
<i>Abstract (E)</i>	282

---

# Einleitung

---

„Braucht es derzeit tatsächlich einen kritischen Kommentar zur Lage?“

*Ulrich Pfisterer, Hans im Glück: Kunstgeschichte heute<sup>1</sup>*

Den Geisteswissenschaften wurden im Laufe der Zeit eine verblüffend lange Reihe an *turns* attestiert. So wurde zum Beispiel ein *linguistic*, ein *pictorial* oder *iconic*, ein *feminist* und auch ein *digital turn* im Verlauf der letzten Jahrzehnte ausgerufen.<sup>2</sup> Die Frage nach der Angemessenheit des Begriffs *turn* für die hier jeweils bezeichneten Beobachtungen außen vor gelassen, ist doch die Tatsache interessant, dass der Begriff des *turns* überhaupt zum Tragen kommt; impliziert er doch die Annahme, es handle sich bei den beobachteten Entwicklungen um bereits im Institutionalisierungsprozess inbegriffene Elemente der jeweiligen Fächer. Der Begriff des *turns* steht also, so könnte man salopp formulieren, für die Beobachtung: es tut sich etwas im Fach! Je nach *turn* werden Themenfelder neu erschlossen, das methodische Vorgehen ausgeweitet oder auch Fragen unter neuen Gesichtspunkten betrachtet. Die fachliche *comfort zone*, die vertraute, immer gleiche und verlässliche Umgebung des Denkens, die Basis aller fachlicher Sicherheit beginnt sich aufzulösen. Zumindest könnte man so meinen – insbesondere angesichts der erhitzten Gemüter, welche die Diskurse um die diversen *turns* notwendig zu begleiten scheinen. Hier werden nicht selten, wie Ulrich Pfisterer in Bezug auf die Kunstwissenschaften so treffend formulierte, „grenzpolizeiliche Befangenheiten“ in der Manier der „letzte[n] Gralshüter der Meisterwerke europäisch-westlicher Hochkunst und `solider` Methoden“ postuliert.<sup>3</sup> Etwas weniger despektierlich könnte man auch festhalten: wann immer in den letzten Jahren ein

---

<sup>1</sup> Pfisterer 2016: 5.

<sup>2</sup> Vgl. z.B. Bachmann-Medick 2006, Boehm 1994, Mitchell 1994.

<sup>3</sup> Pfisterer 2016: 5.

---

*turn* ausgerufen wurde, schieden sich die Geister. Ein besonders erbitterter Positionsstreit wird momentan im Rahmen um den so bezeichneten *empirical turn* geführt.<sup>4</sup> Wenn auch fraglich bleibt, ob es sich bei diesem tatsächlich um eine paradigmatische Wende in den Geisteswissenschaften handelt, so ist doch unbestreitbar, dass insbesondere im Verlauf des letzten Jahrzehnts eine verstärkte Zuwendung zu als *empirisch* beschreibbaren, methodischen Ansätzen zu verzeichnen ist. Dabei ist freilich das Ausmaß, in welchem eine solche Ausweitung des klassischen Methodenrepertoires betrieben wird, in den jeweiligen Disziplinen der Geisteswissenschaften unterschiedlich stark ausgeprägt. Während etwa die Philosophie und die Linguistik bereits vor Jahrzehnten den empirischen Zugang zum Fach durch die Etablierung von Subdisziplinen, wie etwa der Psycholinguistik oder auch der experimentellen Philosophie, institutionalisiert haben, scheint sich die Empirie in den Kunstwissenschaften nur mühsam Raum zu erkämpfen. Dennoch ist auch für die Kunstwissenschaften unbestreitbar, dass empirische Zugänge zum Fach längst keine unbedeutenden Randerscheinungen mehr darstellen. Spätestens seit der Etablierung von Forschungsschwerpunkten, wie etwa der *digitalen Kunstgeschichte*, der *Neuroarthistory* oder auch der *empirischen Bildwissenschaft*, ist die Empirie in der Mitte des Faches angekommen.

Die Bedeutung, welche dem Erstarken des empirisch-ausgeprägten Teilbereichs des Faches zukommt, zeigt sich nicht zuletzt darin, dass eine Institution wie die *Zeitschrift für Kunstgeschichte* seit 2015 Positionspapiere zur aktuellen Lage des Faches veröffentlicht, die sich hauptsächlich der empirischen Ausweitung der Disziplin widmen.<sup>5</sup> Hierbei sind freilich

---

<sup>4</sup> Thiel 2012.

<sup>5</sup> Dazu ist dem Editorial des 78. Bandes der *Zeitschrift für Kunstgeschichte* (2015) zu entnehmen: „Wenn die *Zeitschrift für Kunstgeschichte* weiterhin als ein Forum für das gesamte Fach dienen soll, wird es von entscheidender Bedeutung sein, in verstärktem Maße auch die jüngste Kunst sowie aktuelle Fragestellungen und neue theoretische Herausforderungen zur Geltung zu bringen.“ (Frohne et al. 2015: 2). Bisher (Stand Oktober 2017) sind zu empirischen Themen (im weiteren Sinne) bereits sieben Beiträge erschienen – von Whitney Davis (Davis 2015), Peter Geimer (Geimer 2015), Ulrich Pfisterer (Pfisterer 2016), Hubertus Kohle (Kohle 2016), Max Marmor (Marmor 2016), Raphael Rosenberg (Rosenberg: 2016) und Béatrice Joyeux-Prunel (Joyeux-Prunel 2016).

---

längst nicht alle Stimmen, die zum Empirischen in der Kunstgeschichte Stellung beziehen, positiv. Die eingangs erwähnten erhitzten Gemüter finden hier ebenso ihren Raum, wie hoffnungsfrohe Stimmen. Vielleicht am treffendsten lässt sich die Landschaft des Diskurses mit den von Ulrich Pfisterer gewählten Worten beschreiben: „Jedenfalls gibt es eine grobe Dreier-Typologie der Herangehensweisen an die Aufgabe einer >Positionsbestimmung<: adhortative Texte, die mahnend und prophezeiend zukünftige Aufgaben und Perspektiven einfordern; lamentierende, die berechtigt oder habituell unzufrieden mit dem Status quo sind, häufig ohne sich die Blöße einer eigenen Zukunftsidee zu geben; schließlich distanziert konstatierenden Positionen, die von einer vermeintlichen (teils durch Ironie zusätzlich abgesicherten) Außen- und Übersicht die vielfältigen Tendenzen und Positionen ihrer jeweiligen Gegenwart sezieren wollen.“<sup>6</sup> Pfisterers Position in seine eigene Typologie einzuordnen dürfte in Anbetracht der ironischen Distanz, die seinen Text durchzieht, nicht weiter schwerfallen. Tatsächlich ist die hier greifbare, humoristisch-neutrale Perspektive auf die Lage jedoch eher ein Einzelfall als die Regel. Weit häufiger inkorporieren die postulierten Positionen den Typus des „mahnenden Propheten“ oder des „lamentierenden Unzufriedenen“, sodass der Eindruck einer Spaltung entsteht; einer Spaltung in zwei sich in Opposition befindende Lager. Gerade quantitativ-empirische Ansätze werden häufig skeptisch beäugt und als fachfremde Über- und Eingriffe seitens der Natur- und Sozialwissenschaften in den geisteswissenschaftlichen Hegemonialbereich verstanden. Dabei scheint der Grad der Ablehnung linear mit dem Standardisierungsgrad der zur Anwendung kommenden Methode zu wachsen. Damit ist gemeint, dass insbesondere klar kontrollierte Datenerhebungssettings, die meist von statistischen Auswertungen begleitet werden, dazu in der Lage sind, die Geister zu scheiden. Von den einen wird eine solche Erweiterung der Kunstwissenschaften begeistert begrüßt, von den anderen im besten Fall kritisch beäugt, meist aber vehement verdammt.

---

<sup>6</sup> Pfisterer 2016: 4.

---

Angesichts dieser im öffentlichen Fachdiskurs allzu präsenten Spaltung muss Pfisterers rhetorische Frage nach der Notwendigkeit eines kritischen Kommentars zur Lage des Faches Kunstgeschichte wider ihre Suggestion beantwortet werden.<sup>7</sup> Ja, es braucht derzeit tatsächlich einen kritischen Kommentar zur Lage. Es braucht einen kritischen Kommentar, oder vielmehr eine kritische Analyse und Aufarbeitung der opponierenden Positionen, weil beiderseits Missverständnisse bestehen und weil diejenigen Fragen, bei welchen die größte Uneinigkeit besteht, Fragen nach den *Zielen*, den *Möglichkeiten* und der *Sinnhaftigkeit* einer empirisch-arbeitenden Kunstwissenschaft in den Subdiskursen der jeweiligen Positionen isoliert bleiben. Eine kritische Reflexion, die sowohl die Bedenken der „Unzufriedenen“ als auch die Hoffnungen der „Propheten“ inkludiert, findet nicht statt. Ein Zustand, der nur zu sich weiter verhärtenden Fronten führen wird, ohne letztlich Antwort auf die Frage geben zu können, worauf die zu konstatierende Spaltung im Wesentlichen zurückzuführen ist. Resultiert die Spaltung aus den, von Pfisterer so despektierlich als „grenzpolizeiliche Befindlichkeiten“ beschriebenen, fachpolitischen Ansichten oder Haltungen, oder gibt es tatsächlich gute Gründe empirische Zugänge für die Kunstwissenschaften als unangemessen zu empfinden?

Mit dieser Arbeit möchte ich einen Beitrag zu einem Brückenschlag leisten. Dazu werde ich im ersten Teil zunächst einen groben Überblick über die Spielarten empirischer Zugänge in den Kunstwissenschaften geben. Im Konkreten soll damit für den weiteren Verlauf der Diskussion eine Referenzgrundlage geschaffen werden. Hierzu gehört neben der Darstellung und Differenzierung der Tätigkeitsfelder der empirisch-motivierten Kunstwissenschaft auch die grundlegende Definition empirischer Ansätze im Allgemeinen, sowie die Abgrenzung empirischer Methoden von klassischen Methoden oder Techniken. Die Schaffung einer solchen definatorischen Grundlage erlaubt nicht nur die methodologische Klassifikation der jeweiligen Ansätze, sondern legt gleichzeitig auch die Eckpfeiler des Möglichkeitsraumes der unterschiedlichen Methoden offen. Somit werden sowohl die Chancen als auch die

---

<sup>7</sup> „Braucht es derzeit tatsächlich einen kritischen Kommentar zur Lage?“, ebd.: 5.

---

Grenzen empirischer Methoden transparent. Daran anschließend werde ich sowohl die Argumentation der positiven Kritik als auch die Argumentation der negativen Kritik an ausgewählten Beispielen darstellen. Hierzu werde ich einen kurzen, als paradigmatisch zu begreifenden Überblick zu den oppositionellen Positionen geben, um in der Folge die wesentlichen, in diesem Wechselspiel der Positionen zu Tage tretenden, Uneinigkeiten konkret zu adressieren. Inhaltlich beziehend auf die zuvor geschaffene zweiteilige Referenzgrundlage, bestehend aus den konkreten Tätigkeitsfeldern der empirisch-arbeitenden Kunstwissenschaft und der grundlegenden Definition empirischer Ansätze, sollen hier zum einen Missverständnisse identifiziert werden; zum anderen sollen aber auch Fragen und Probleme benannt werden, die nicht auf ein unzureichendes Verständnis von den Möglichkeiten und Zielen empirischer Methoden reduzierbar sind.

Der zweite Teil dieser Untersuchung greift die zuvor identifizierten Fragen und Probleme, die empirischen Ansätzen gegenüber bestehen, auf und bereitet eine konkrete Untersuchungsgrundlage für ihre Analyse und Diskussion vor. Dazu werden drei empirische Studien vorgestellt, an deren Beispiel die zuvor abstrakt-theoretisch diskutierten Möglichkeiten bzw. Grenzen des empirischen Arbeitens greifbar werden. Die hierfür ausgewählten Studien, die aufgrund des kumulativen Charakters dieser Arbeit dem Fließtext in ihrer ursprünglichen Publikationsform beigelegt sind, sind dem Kontext rezeptionsästhetischer Fragestellungen zuzuordnen. Das erste Beispiel ist eine umfangreiche Studie zur Erfassung und Definition des Konzeptes der Komplexität in der Malerei; die beiden weiteren Beispiele sind kleinere empirische Studien zu unterschiedlichen Fragen, die von der bisherigen Forschung an die Betrachtung von und die Auseinandersetzung mit abstrakter Kunst gestellt wurden. Die Studien greifen auf unterschiedliche Messmethoden zurück, wodurch an ihnen ein breites Methodenspektrum anschaulich wird. Diese Methodenvielfalt hat für die Diskussion der Chancen, Grenzen und Probleme empirischer Ansätze den Vorteil, dass auch Unterschiede zwischen einzelnen Vorgehensweisen und Messmethoden identifiziert und zur Diskussion gestellt werden können. Gerahmt werden die einzelnen Studien jeweils durch eine knappe theoretische Einbettung der untersuchten Ideen



---

sowie eine Diskussion der empirischen Ergebnisse und des spezifischen Nutzens der empirischen Herangehensweise für die Beantwortung der Fragestellung.

Der dritte Teil der Arbeit ist einer ausführlichen Methodenreflexion gewidmet. Diese wird explizit auf die Probleme und Einschränkungen der vorgestellten Studien und ihrer Ergebnisse eingehen, um auch die Grenzen der hier gewählten empirischen Zugänge deutlich offenzulegen und so einen Anknüpfungspunkt für die Diskussion der negativen Kritik an einer empirisch-arbeitenden Kunstwissenschaft zu schaffen.

In einem abschließenden Fazit werde ich schließlich einen begründeten Antwortversuch auf die Frage nach dem Wert und den Voraussetzungen empirischer Ansätze für die Kunstwissenschaften entwerfen. Im Konkreten werde ich dazu – auf der Basis der im zweiten und dritten Teil erarbeiteten Detailbetrachtungen – die gegenüber empirischen Ansätzen geäußerten Kritikpunkte im Einzelnen prüfen. Es werden also sowohl die am konkreten Beispiel nachvollzogenen Aspekte positiver als auch negativer Kritik gleichermaßen berücksichtigt. Es ist damit gesichert, dass die abschließende Diskussion das Spannen eines vermittelnden Brückenschlags zwischen den von Pfisterer so bezeichneten „mahrenden Propheten“ und „lamentierenden Unzufriedenen“ erlaubt, da eine integrierte Perspektive auf den bestehenden Diskurs realisiert wird.

---

*Kapitel* **1**

Empirische Ansätze in der  
Kunstwissenschaft

---

Der Forschungsstand zur Frage nach empirischen Methoden in der Kunstwissenschaft im Allgemeinen und deren Wert oder Nutzen im Speziellen ist wenig umfangreich.<sup>8</sup> Hauptsächlich finden sich hier Versuche unterschiedliche Methoden systematisch zu gruppieren. Eine tiefergehende methodologische Analyse ist hingegen noch immer ausstehend. Die meisten der bisherigen Arbeiten sind zudem der *digitalen Kunstgeschichte* zuzuordnen; sie beschäftigen sich deswegen hauptsächlich mit *digitalen* Methoden. *Digitale* Methoden sind zwar, wie im weiteren Verlauf noch deutlicher werden wird, häufig auch als *empirische* Methoden beschreibbar, doch handelt es ganz ausdrücklich *nicht* bei jedem Zugang, der auf eine digitale Methode bzw. ein digitales Arbeitswerkzeug zugrückt, auch notwendig um einen empirischen Forschungsansatz.

Systematische Methodenkategorisierungen im Bereich der digitalen Kunstgeschichte wurden zuletzt beispielsweise von Angela Dressen (*Grenzen und Möglichkeiten der digitalen Kunstgeschichte und der Digital Humanities – eine kritische Betrachtung der Methoden*, 2017)<sup>9</sup> und auch Hubertus Kohle (*Digitale Kunstgeschichte*, 2013, siehe auch: Jannidis, Kohle, & Rehbein (Hg.), *Digital Humanities. Eine Einführung*, 2017)<sup>10</sup> unternommen.<sup>11</sup> Hier finden sich in beiden Fällen Überblicke zu den verschiedenen Feldern der digitalen Kunstgeschichte und den dort zur Anwendung kommenden Methoden. Dressen und Kohle legen dabei für ihre Überblicke unterschiedliche Categoriesysteme zugrunde; während Kohle seine Kategorisierung weitestgehend entlang der Anwendungsbereiche der unterschiedlichen

---

<sup>8</sup> Die hier besprochenen Ansätze und Arbeiten entspringen ausschließlich der *zeitgenössischen* empirischen Forschung in den Kunstwissenschaften. Um 1900 fand bereits ein erster „empirical turn“ statt (siehe auch Kapitel 2.4 *Kritik und Probleme empirischer Ansätze in der Kunstwissenschaft*). Dieser verläuft nach dem ersten Weltkrieg jedoch im Sand und steht in keiner direkten Relation zur den neueren empirischen Tendenzen in den Kunstwissenschaften (vgl. z.B. Lanzoni/Brain/Young 2012). Dieser erste „empirical turn“ ist daher nicht Gegenstand der folgenden Betrachtungen.

<sup>9</sup> Dressen 2017.

<sup>10</sup> Jannidis/Kohle/Rehbein 2017, Kohle 2013, außerdem: *Kunstgeschichte digital* (Kohle 1997).

<sup>11</sup> Für eine Darstellung des Forschungsstandes in den USA vgl. auch Drucker/Helmreich/Lincoln/Rose 2015, Drucker 2013, Zorich 2012, Drucker 2009.

---

Methoden führt (Suchen, Analysieren, Publizieren, Präsentieren/Rekonstruieren)<sup>12</sup>, orientieren sich die von Dressen gewählten Kategorien hauptsächlich an den methodischen Ansätzen (Mapping und Networking, Text-Encoding, Data Mining und Big Data, Kognitive Methoden/Empirische Bildwissenschaft, Computer Vision).<sup>13</sup> Der wesentliche Unterschied dieser beiden Zugänge ist im Zweck der jeweiligen Schriften auszumachen: Kohles systematische Zusammenstellung dient tatsächlich einer Übersicht und dem Festschreiben eines momentanen *status quo*; Dressens Aufsatz ist hingegen eher als kritische Analyse zu verstehen, deren Ziel es ist, die digitale Kunstgeschichte im Bezug zum Feld der *Digital Humanities* zu diskutieren, zu positionieren und abzugrenzen. Ihre Besprechung kommt dabei zu einer wesentlichen Unterscheidung: Sie hält fest, dass der Gebrauch von digitaler Technologie nicht notwendig mit einem *Digital Humanities*-Ansatz gleichzusetzen sei – ganz im Gegenteil, es sei dringend erforderlich zu unterscheiden zwischen dem Einsatz von digitalen Arbeitswerkzeugen einerseits und „computergestützter Forschung“ andererseits.<sup>14</sup> Mit ersterem ist eine im weiteren Sinne klassische Herangehensweise bezeichnet, die durch den Einsatz digitaler Arbeitswerkzeuge lediglich erleichtert wird. Unter zweiterem lässt sich hingegen ein Vorgehen fassen, das über die schlichte Anwendung von digitalen Arbeitswerkzeugen hinausgeht: Hier ist das digitale Arbeitswerkzeug tatsächlich Methode. Ansätze, die Dressen – etwas missverständlich formuliert – unter computergestützter Forschung fasst, *benötigen* den digitalen Zugang; er ist wesentlicher Bestandteil des Erkenntnisprozesses. Ohne ein digitales Methodenrepertoire wäre in diesen Fällen die angestrebte Erkenntnis nicht zu leisten.<sup>15</sup>

Die von Dressen geforderte Unterscheidung, die leider ohne weitere Diskussion verbleibt, ist eine Wesentliche. Hier deutet sich bereits an, dass die dem weiten Feld der digitalen Kunstgeschichte zugeschriebenen Methoden aus einer methodologischen und epistemologischen Perspektive unterschiedlichen Parametern unterworfen sind. Salopp

---

<sup>12</sup> Vgl. auch Kohle 2018: 10, und Kohle 2008.

<sup>13</sup> Dressen 2017.

<sup>14</sup> Dressen 2017:14; vgl. auch Drucker 2013, Pias 2003.

<sup>15</sup> Siehe auch Jannidis/Kohle/Rehbeim 2018: 13-15.

---

gesprochen: Digitale Kunstgeschichte ist nicht gleich digitale Kunstgeschichte. Die hier subsummierten Zugänge zum Fach unterscheiden sich nicht nur hinsichtlich dem, was sie jeweils inhaltlich bewirken oder erzielen können, sondern auch hinsichtlich der Geisteshaltungen, welche den methodischen Zugängen zugrunde liegen: während die von Dressen so benannten digitalen Arbeitswerkzeuge eine Stütze des altbekannten hermeneutischen Forschungsprozesses darstellen, basiert die „computergestützte Forschung“ auf der Annahme, dass Erkenntnis bestmöglich durch eine (objektive) empirisch-analytische Perspektive auf Bild- und Textmaterial gewonnen wird. Im Grunde genommen lässt sich die Unterscheidung also auf die erkenntnistheoretischen Prämissen der jeweiligen Methoden herunterbrechen.

Das eine diesbezügliche Unterscheidung bisher nicht, oder zumindest nicht in ausreichendem Maße, geleistet wurde, mag eine der Quellen der angesprochenen Missverständnisse und folglich auch wesentlichen Kritikpunkte sein. Das zeigt sich anhand der sorgfältigen Lektüre der zweiten größeren Schriftengruppe zu digitalen/empirischen Methoden: der Diskussion von spezifischen Vor- und Nachteilen des digitalen Herangehens an Bild- und Kunstwerke. Die hierzu geleisteten Diskussionsbeiträge lassen sich in zwei gegensätzliche Gruppen einordnen – I) digitale-Methoden-bejahende Beiträge, die spezifische Methoden und ihre Möglichkeiten vorstellen und II) Schriften, die eine Skepsis gegenüber der Anwendung eines digitalen Methodenrepertoires ausdrücken, dabei aber eine allgemeinere Perspektive wählen.

## I.

Beispiele für bejahende Beiträge wären etwa der Aufsatz *Ein Bild sagt mehr als tausend Pixel? Über den Einsatz digitaler Methoden in den Bild- und Objektwissenschaften* von Celia Krause und Ruth Reiche (2014)<sup>16</sup> oder auch der, nur als Abstract publizierte Vortrag, den Peter Bell und Lisa Diekmann, 2016 auf 3. *Tagung des Verbands Digital Humanities im*

---

<sup>16</sup> Krause/Reichle 2014.

---

*deutschsprachigen Raum e.V.* in Leipzig hielten (*Die Kunst als Ganzes. Heterogene Bilddatensätze als Herausforderung für die Kunstgeschichte und die Computer Vision*).<sup>17</sup> Diese Form von Beiträgen, ob nun gesprochen oder schriftlich, stellt bestimmte Methoden oder methodische Aspekte heraus und setzt im diskursiven Stil zu ihrer (voraussetzenden) Verteidigung an. Krause und Reiche diskutieren, wenn auch nicht ausschließlich, so doch schwerpunktmäßig den Nutzen von unterschiedlichen *digitalen Bildinformationen* sowie dem digitalen Bild zugeordneten *Metainformationen*. Bells diverse Vorträge und verwandte Aufsätze zum Thema digitale Kunstgeschichte beleuchten hingegen meist Aspekte der *computergestützten semantischen Erschließung* von Bildinhalten (Computer Vision). Die beiden angeführten Beispiele beziehen sich also einerseits eher auf formale Bildaspekte im digitalen Raum und andererseits eher auf die digital-gestützte semantische Erschließung von dargestellten Inhalten. In beiden Fällen ist der Ton der Argumentation positiv, das Fazit wohlwollend.

## II.

Diesen positiven Stimmen gegenüber stehen unter anderem die argumentative Auseinandersetzung Jürgen Zimmers *Kunstforschung – Kunstwissenschaft – Kunstgeschichte → CyberArtHistory?* (1997)<sup>18</sup> oder auch Claire Bishops scharf-betitelter Aufsatz *Against Digital Art History* (2018).<sup>19</sup> In beiden Fällen wird auf deutlich allgemeinerer Ebene das Methodenrepertoire der digitalen Kunstgeschichte kritisiert. Interessant ist dabei, dass die Kritik in Zimmers und Bishops Ausführungen in ihren zentralen Punkten die Gleiche ist – trotz der 20 Jahre, welche die beiden Schriften trennen. In beiden Fällen wird die Abkehr vom Original und die Zuwendung zur digitalen Reproduktion als

---

<sup>17</sup> Diekmann/Bell 2016. Bells Position wurde darüber hinaus in diversen geisteswissenschaftlichen und auch nicht-geisteswissenschaftlichen Artikeln veröffentlicht: z. B. Bell/Ommer 2017, Bell/Ommer 2016a, Bell/Ommer 2016b, Bell/Ommer 2015, Monroe/ Bell/ Ommer 2014.

<sup>18</sup> Zimmer 1997.

<sup>19</sup> Bishop 2018.

---

missliche Entwicklung gedeutet.<sup>20</sup> Zimmer schreibt dazu beispielsweise die folgenden sehr sprechenden Worte: „Der Fortschritt wird fortschreiten, und es wird genügend Kunsthistoriker geben, die [...] meinen, daß schon genug geforscht sei und sie sich nur mit Texten und Ideen, allenfalls noch mit Abbildungen, aber nicht mehr mit den Werken selbst beschäftigen und an der Grundlagenforschung keinen Anteil haben. Sie können politisch korrekt reden und schreiben.“<sup>21</sup> Bishop, die Zimmers grundsätzliche Haltung teilt, sich aufgrund des 20-jährigen zeitlichen Abstands aber bereits mit der von Zimmer befürchteten Fortentwicklung der digitalen Kunstgeschichte konfrontiert sieht, hebt noch einen weiteren Aspekt digitaler Zugänge besonders hervor: Sie kritisiert die häufig durch eine *systematische* und *logische* Perspektive bestimmte Forschung der digitalen Kunstgeschichte scharf, indem sie betont, dass damit interessante Ambiguitäten und Unvorhersehbarkeiten nicht erfasst werden können. Darüber hinaus beschreibt sie den Rückgriff auf *empirische* und *statistische* Methoden als unbefriedigend, da die hier erzielbaren Ergebnisse ihrer Ansicht nach im Vergleich zum erforderlichen Aufwand banal wären.<sup>22</sup>

Interessant an diesem Diskurs zur *digitalen* Kunstgeschichte ist, dass es insbesondere die *empirischen* Aspekte der digitalen Forschung sind, welche vordergründig Diskussionsbedarf aufzeigen. Dressens Methodenkategorisierung kommt zu dem Schluss, dass diesbezüglich eine *Unterscheidung notwendig* ist. Methoden(gruppen)-spezifische Beiträge, wie etwa die genannten Beispiele, unterliegen immer dann einem *besonderen Rechtfertigungsdruck*, wenn ihre methodischen Zugänge einem analytisch-empirischem Paradigma entsprechen; und auch die den digitalen Methoden gegenüber geäußerte *Kritik* richtet sich größtenteils gegen die hier aufscheinende Empirie. Eine Auseinandersetzung mit der erkenntnistheoretischen Basis digitaler Methoden ist folglich nicht nur ausstehend, sondern unumgänglich. Ich werde daher im Folgenden einen kurzen Überblick zu den diversen Spielarten *empirischer* Methoden geben, der sich insofern von den bisher geleisteten Überblicksarbeiten unterscheidet als er

---

<sup>20</sup> Vgl. auch Parsons 2005, Elkins 1997.

<sup>21</sup> Zimmer 1997: 10.

<sup>22</sup> Vgl. Bishop 2018.

---

die erkenntnistheoretischen Prämissen der methodischen Ansätze offenlegt und damit deren direkte Adressierbarkeit ermöglicht. Da empirische Ansätze nicht ausschließlich im Kontext der digitalen Kunstgeschichte zu finden sind, wird der folgende Überblick zwar schwerpunktmäßig empirische Methoden behandeln, welche diesem Spektrum zuzuordnen sind, auf vereinzelte Querverweise zu anderen Subfeldern der Kunstgeschichte aber dennoch nicht verzichten. Wie bereits angekündigt, schließe ich an diesen Überblick im weiteren Verlauf dieses Kapitels eine allgemeine Definition empirischer Methoden an und komme schließlich auf die positive und negative Kritik zurück, welche die Diskurse um digitale/empirische Methoden begleitet.

### 1.1 *Spielarten empirischer Ansätze in der Kunstwissenschaft*

Was im Kontext der Kunstwissenschaft als empirischer Ansatz verstanden werden kann, ist sehr divers. Empirische Zugänge zur Kunstwissenschaft können den klassischen Natur- und Sozialwissenschaften methodisch sehr nahe sein, zum Beispiel, wenn sie die Analyse von menschlichem Verhalten oder chemischen Substanzen zur Grundlage haben. Im Kontrast dazu verstehen wir aber auch die computergestützte Analyse eines einzelnen Bildbeispiels als empirischen Zugang zum Werk. Eine computergestützte Werkanalyse lässt sich als empirisch begreifen, weil sie im Vergleich zu einem klassisch-hermeneutischen Forschungsansatz einem anderen erkenntnistheoretischen Paradigma unterworfen ist: während hermeneutisch-orientierte Forschungsansätze die subjektive Fähigkeit zur Erkenntnis betonen, indem sie Erkenntnis als Auslegung oder Interpretation verstehen, also einem erkenntnistheoretischen Explikationismus folgen, basieren die erkenntnistheoretischen Annahmen, welche empirischen Ansätzen zugrunde liegen, auf der Ansicht, Erkenntnis könne nur über Beobachtung und logisch-einwandfreie Ableitung gewonnen werden. Die erkenntnistheoretische Grundlage empirischer Ansätze setzt dabei lediglich die Beobachtung als zentrales Merkmal voraus, *nicht* aber ein auf Quantifizierung basierendes Vorgehen. Empirische Ansätze können demnach durchaus auch bei



---

Einzelfallstudien zum Tragen kommen und setzen keineswegs die Befragung von Versuchspersonen oder die chemische Analyse diverser Malschichten im Werk voraus.

Legt man eine erkenntnistheoretische Perspektive zugrunde, sind auch viele klassische Methodenzugänge der Kunstwissenschaft als empirisch zu verstehen. Beispielsweise könnte man ganz allgemein all jene Zugänge anführen, die auf vergleichende Beobachtungen in der Analyse von Bild- und Textmaterialien zurückgreifen. Hier werden, zumindest partiell, während einer inhaltlichen Auseinandersetzung empirische Anteile realisiert. Trotz der Tatsache, dass vielen klassischen Methodenzugängen empirische Anteile zugeschrieben werden können, werden heutzutage hauptsächlich solche Ansätze als empirisch bezeichnet, die ursprünglich in den Natur- und Sozialwissenschaften sowie der Informatik geprägt wurden – Ansätze also, die in den Geisteswissenschaften keine Tradition aufweisen. Im aktuellen Diskurs um (digitale) empirische Zugänge zum Feld der Kunstwissenschaft stehen ebenfalls letztere Zugänge im Fokus, also empirische Zugänge, die den natur- und sozialwissenschaftlichen Fächern sowie der Informatik verwandt sind. Ich werde daher im Weiteren *keine* Beispiele diskutieren, die als klassische Zugänge mit empirischem Anteil verstanden werden können, sondern stattdessen diejenigen empirischen Zugänge schwerpunktmäßig in den Blick nehmen, welche den Anstoß der bestehenden Kontroverse gebildet haben und noch immer bilden.

Forschungsansätze, die im zeitgenössischen Diskurs als empirische Ansätze begriffen werden, lassen sich in der Regel, den Schlagworten *digitale Kunstgeschichte*, *empirische Bildwissenschaft* und *Neuroarthistory* (oder Neuronale Kunstgeschichte) zuordnen. Dabei sind die Begriffe *digitale Kunstgeschichte*, *empirische Bildwissenschaft* und *Neuroarthistory*, sowie die von ihnen bezeichneten inhaltlichen bzw. methodischen Bereiche, nicht eindeutig als disjunkte Konzepte oder Forschungsbereiche zu verstehen. Es handelt sich hierbei vielmehr um stark-überlappende Themen- und Methodengebiete, die sich je nach Definition sogar in eine hierarchische Struktur fassen lassen.

---

Die *digitale Kunstgeschichte*, die man im breiteren Sinne als groben Überbegriff für empirische Ansätze in der Kunstwissenschaft führen kann, ist ein Forschungs- und Methodengebiet, welches sich kunstwissenschaftlichen Fragen- und Problemstellungen mittels *digitaler Methoden* nähert.<sup>23</sup> Mit *digitalen Methoden* ist eine ganze Bandbreite an methodischen Werkzeugen bezeichnet, deren verbindendes Merkmal ein *computer- und datengestützter Zugang zur bestehenden Aufgabenstellung* ist.<sup>24</sup> Es sind jedoch nicht alle digitalen Methoden notwendigerweise auch *empirische Methoden*: So ist beispielsweise ein nicht unwesentlicher Teilbereich der digitalen Kunstgeschichte der wissenschaftlichen Dokumentation von Objekten und Sachverhalten gewidmet.<sup>25</sup> Hierzu werden unter Rückgriff auf verschiedene Methoden und Techniken der Digitalisierung sowohl Bestands- als auch Forschungsdatenbanken von Kunstwerken, Quellen und Literaturkorpora geschaffen. Es wäre allerdings nicht ganz korrekt, Dokumentationsarbeit *grundsätzlich* als nicht empirisch oder als ohne-empirischen-Anteil zu beurteilen.<sup>26</sup> Einige Dokumentationsformen inkludieren qua Format eine empirische Analyse. Raumbezogene Dokumentationen, wie zum Beispiel Geoinformationssysteme (GIS), wären hierfür ein Beispiel.<sup>27</sup> GIS-Systeme dokumentieren bestimmte Sachverhalte ortsbezogen. Sie lassen sich somit an die visuelle Darstellung einer

---

<sup>23</sup> Vgl. z.B. Klinke 2018: 19, Klinke 2016.

<sup>24</sup> Im Folgenden soll lediglich ein Überblick zu den wichtigsten *empirischen* Methoden und Anwendungsfeldern gegeben werden. Ich verfolge nicht den Anspruch, eine vollständige Liste aller methodischen Zugänge zu geben, welche der digitalen Kunstgeschichte zugeschrieben werden können. Diejenigen Teilaspekte der digitalen Kunstgeschichte, welche sich dem Suchen, Publizieren und Vermitteln widmen, sind erklärtermaßen nicht Teil dieses Überblicks. Eine zusammenfassende Darstellung dieser Aspekte findet sich in der 2013 von Hubertus Kohle veröffentlichten Schrift zur digitalen Bildwissenschaft (Kohle 2013).

<sup>25</sup> Ein anderes Beispiel für digitale Methoden, welche nur bedingt als empirisch verstanden werden können, wäre die virtuelle Modellierung und Rekonstruktion von Bauwerken oder Innenausstattungen (z. B. Lengyel 2011, Hoppe 2001). Hier werden zumeist hermeneutische Erkenntnisse mittels digitaler Methoden in ein neues Format umgesetzt. Das Ziel dieser Arbeiten ist daher meist eher Anschaulichkeit als das Erschließen neuer Erkenntnisse durch Beobachtung.

<sup>26</sup> Vgl. hierzu auch Lev Manovich, der die Strukturierungsarbeit, welche einer Datenbank notwendigerweise zugrunde liegt, als „projection of the ontology of a computer onto culture itself“ beschreibt (Manovich 2002, S.223).

<sup>27</sup> Es gibt selbstverständlich auch andere Datenbankformate, welche zum Teil empirisch erarbeitet werden. Für einen detaillierteren Überblick zu diesem Thema verweise ich auf die bereits erwähnte Schrift *Digitale Bildwissenschaft* von Hubertus Kohle (Kohle 2013).

---

Landkarte angliedern, wodurch räumliche Verteilungsstrukturen herausgearbeitet werden.<sup>28</sup> GIS-Systeme dieser Art sind im Grunde genommen nichts anderes als eine spezifische Form der Netzwerkanalyse. Die Netzwerkanalyse im Allgemeinen stellt einen relationalen Forschungsansatz dar. Im Fokus stehen hier nicht die spezifischen Charakteristika bestimmter Objekte oder Personen, sondern deren Verbindungen und Verknüpfungen zu anderen Objekten oder Personen. Damit wird die beobachtbare Struktur, in welche die Untersuchungsobjekte eingegliedert sind, zur Einheit der Analyse. Die Natur dieser Struktur kann je nach Frage variieren. So kann etwa, wie im Fall von GIS-Systemen, eine räumliche Struktur abgebildet sein; alternativ kann auch eine soziale oder zeitliche Struktur im Zentrum der Analyse stehen. Netzwerkanalysen unterschiedlichster Art gehören zu den derzeit am häufigsten zur Anwendung kommenden empirisch-digitalen Methoden.<sup>29</sup>

Neben relationalen Analysen kommen empirische Ansätze auch in unterschiedlichster Form bei der Bildanalyse zum Einsatz. Im Falle der technischen Kunstgeschichte, deren Fokus auf der chemischen und bzw. oder technischen Analyse des Objektes liegt, ist der empirische Zugang nur bedingt als digital zu verstehen. Darüber hinaus werden die empirischen Zugänge der technischen Kunstgeschichte als bereits etablierte Methoden begriffen und bergen daher wenig Konfliktpotenzial. Sie dienen häufig der Sichtbarmachung bestimmter Strukturen, die mit dem bloßen Auge aufgrund der Natur des betrachteten Objektes nicht erkennbar sind. Dazu zählen etwa die Herausarbeitung von unter der Malschicht verborgenen Unterzeichnungen mittels Infrarot- oder Röntgenscans oder auch der Abgleich der chemischen Zusammensetzung zweier Farbproben.<sup>30</sup> Häufig werden die mittels dieser Ansätze gewonnenen Daten oder Bilder wieder in den durch klassische Methoden und Zugänge charakterisierten Forschungsprozess eingegliedert.<sup>31</sup> Die (weitere) Analyse dieses Materials basiert demnach in vielen Fällen nicht auf einem empirisch-analytischen Paradigma, sondern folgt einem klassischen hermeneutischen Ansatz.

---

<sup>28</sup> Z. B. Gagliardi/ Petridis 2015.

<sup>29</sup> Z. B. Schich et al. 2017, Kaiser 2014; vgl. auch Dressen 2017.

<sup>30</sup> Z. B. Bin et al. 2017, Beck 1995.

<sup>31</sup> Z. B. IIsink 2016.

---

Die Ansätze der technischen Kunstgeschichte sind jedoch keineswegs die einzigen bildanalytischen Methoden der Kunstwissenschaft, welche auf einen empirischen Grundgedanken zurückgreifen. Insbesondere zur semantischen Erschließung von Kunstwerken wurden über die letzten Jahrzehnte einige digitale Methoden entwickelt, deren Zugang zum Bild dezidiert als empirisch zu verstehen ist. Allgemein lassen sich diese Ansätze als Versuch begreifen die Analyse von kulturhistorischen Bildaspekten zu automatisieren: das Erkennen und Zuordnen von Inhalten und Bedeutungen wird hier nicht mehr – wie in klassisch-hermeneutisch Ansätzen – durch einzelne Expert\_innen geleistet, sondern auf computer-basierte Verfahren ausgelagert. Diese basieren je nach Ansatz auf algorithmischen Lösungen, die im Sinne einer künstlichen Intelligenz eine automatische (semantische) Bilderkennung und Klassifikation ermöglichen,<sup>32</sup> oder aber sie basieren auf der dezentralen Akkumulation von Wissen. Damit ist gemeint, dass Erkenntnis nicht durch eine spezifische Perspektive gewonnen wird, sondern durch die Gleichzeitigkeit vieler unterschiedlicher Perspektiven. Die inhaltliche Erschließung von Kunstwerken sowie die Auseinandersetzung mit Bedeutung wird hier nicht über ein subjektives, sondern über ein „kollektives Bildbewusstsein“ geleistet.<sup>33</sup> In beiden Fällen, sowohl in der algorithmisch-basierten Bildanalyse als auch im Falle einer dezentralen Erkenntnisleistung, kommt der subjektive Fokus einer hermetischen Perspektive nicht zum Tragen. Ganz im Gegenteil, die Subjektivität der Forscher\_innen-Perspektive wird hier weitestgehend umgangen. Aufgrund des aktuellen Entwicklungsstandes von (für solche Ansätze) zwingend erforderlichen Tools und Technologien, sind diese Zugänge bisher jedoch nur ansatzweise in der Lage, eine echte automatisierte, semantische Analyse zu leisten.

Im Gegensatz dazu sind die technischen Möglichkeiten einer formalen Bildanalyse, und hier insbesondere einer formalen Gruppierung, bereits hinreichend gut entwickelt, um breitflächig einsatzfähig zu sein. Ansätze, die den *Visual Analytics* zuordenbar sind, nutzen

---

<sup>32</sup> Z. B. Bell/Ommer 2015, Monroy/Bell/Ommer 2014.

<sup>33</sup> Ein Beispiel hierfür wären Bilddatenbankformate, deren Verschlagwortung auf der Anwendung von Social Tagging Ansätzen beruht, wie etwa das an der Ludwig-Maximilians-Universität in München entwickelte ARTigo Projekt (vgl. Scherz 2017, Commare 2011; Kohle 2011).

---

digitale Methoden für die empirische Analyse großer (Bild-)Datenmengen. Beispiele hierfür wären etwa quantitative Farbverteilungsanalysen<sup>34</sup> oder auch eine algorithmische Stilklassifikation, welche auf objektiver Basis visuelle Gemeinsamkeiten zwischen Stilen identifiziert.<sup>35</sup>

Alle bisher genannten Einsatzbereiche empirischer Ansätze haben im Kern das Objekt zum Gegenstand. Es sind Werkzeuge, die der formalen sowie inhaltlichen Analyse von einzelnen Kunstwerken oder Werkgruppen dienen. Gemein ist ihnen der Versuch über empirisch-analytische Zugänge die Subjektivität klassischer Erkenntnisansätze zu überkommen bzw. diese um eine objektive Perspektive zu erweitern und zu bereichern. Aufgrund des objektbasierten Fokus ist jedoch nur das Kunstwerk selbst und nicht die Subjektivität der jeweiligen Einzelperspektiven Gegenstand dieser Ansätze. Die Frage nach der Relevanz von subjektiven Perspektiven in der Wahrnehmung und Rezeption von Kunst ist in der kunstwissenschaftlichen Forschung dennoch bereits lange präsent. Sowohl die klassische Rezeptionsforschung als auch die Rezeptionsästhetik sind dafür eindrückliche Belege. Es handelt sich in beiden Fällen um Ansätze, die – aufgrund der Natur ihrer Sache – Anleihen bei soziologischen und wahrnehmungspsychologischen Theorien machen.<sup>36</sup> Dieser Fokus auf die Betrachtenden sowie die Anleihen aus den traditionellerweise empirisch-arbeitenden Sozialwissenschaften prädestinieren diese Forschungsgebiete für den Einsatz empirischer Methoden. Es ist daher auch kaum verwunderlich, dass wir gerade in diesem Kontext eine ganze Reihe unterschiedlicher empirischer Zugänge zum Fach finden.

---

<sup>34</sup> Z. B. Manovich 2015.

<sup>35</sup> Z. B. Saleh et al. 2016, Arora/Elgammal 2012.

<sup>36</sup> In der klassischen Rezeptionsforschung werden, wie etwa von Michael Baxandall, kulturelle Determinanten des subjektiven Zugangs zu Kunst als Katalysatoren für Stilentwicklungstendenzen verstanden (Baxandall 1972; vgl. auch Caviness 2001). In der Rezeptionsästhetik, deren zeitgenössische Spielart im Wesentlichen auf Wolfgang Kemp zurückzuführen ist, steht die Frage nach dem Ort der Bedeutungskonstruktion im Mittelpunkt der Auseinandersetzung. Damit ist die Frage nach der Relevanz des betrachtenden Subjekts im Verhältnis zum betrachteten Bild gemeint (Kemp 1992, Kemp 1986).

---

Insbesondere die Rezeptionsästhetik, die eine enge Verwandtschaft zur Wahrnehmungspsychologie aufweist, ist ein Spielfeld für eine Vielzahl von empirischen Ansätzen. So werden Wahrnehmung und Rezeptionshandlungen unter anderem mittels physiologischen und behavioralen Markern gemessen. Zu jenen Methoden, die herangezogen werden um kognitive und emotionale Reaktionen auf Kunst zu messen, zählen physiologische Messungen, wie zum Beispiel Eye-Tracking zur Vermessung des Blickverhaltens und der Verteilung von Aufmerksamkeit, Gesichtselektromyographie zur Messung der Valenz emotionaler Reaktionen, oder auch Hautleitwertmessungen zur Erfassung des Erregungsgrades über den Zeitraum der Kunstbetrachtung hinweg.<sup>37</sup> Neben physiologischen Messungen werden insbesondere für die Erfassung der kognitiven Reaktion auf Kunstwerke auch behaviorale Maße herangezogen. Dazu zählen alle Arten von Fragebögen, die eine Bewertungs- oder Beschreibungshandlung erfordern, aber auch Reaktionszeiten sowie Verweildauern bei der Auseinandersetzung mit einem Kunstwerk und die Analyse der spezifischen Aufgaben, welche die Betrachtung eines Werkes begleiten.<sup>38</sup> Alle diese Messmethoden und -instrumente können als digitale Methoden im weiteren Sinne verstanden werden, denn in der Regel werden die erhobenen Daten in eine digitale Form übertragen und die weitere Analyse dieser Daten greift in jedem Fall auf digitale Werkzeuge zurück. Der spezielle Bereich empirischer Zugänge zur Rezeptionsforschung und zur Rezeptionsästhetik wird dennoch häufig eigens übertitelt. Raphael Rosenberg, der hierzu den Begriff der *empirischen Bildwissenschaft* (engl. *cognitive research in art history*) prägte, begründet die Einführung eines neuen Terminus mit der zusätzlichen Spezifität empirisch-experimenteller Forschung am Menschen, welche bei vielen digital *am Objekt* operierenden Methoden nicht gegeben ist: “Using those devices produces large amounts of data that have to be analysed with specific software. The collection of psychophysiological measurements (by art historians) might therefore be considered as part of digital art history. I prefer to call

---

<sup>37</sup> Z. B. Rosenberg/Klein 2015, Brinkmann et al. 2014, Leder et al. 2014, Massaro et al. 2012, Rosenberg 2011.

<sup>38</sup> Z. B. Brinkmann 2017, Brieber et al. 2014, Commare/Brinkmann 2016, Yarbus 1967.

---

it ‘cognitive research in art history’ (or in German: empirische Bildwissenschaft), to underline the psychophysiological, experimental element that precedes digital processing.”<sup>39</sup>

Der *empirischen Bildwissenschaft* eng verwandt ist die *Neuroarthistory*. Sie ist der empirischen Bildwissenschaft ähnlich, oder vielleicht sogar als ein Teilbereich dieser zu verstehen, da auch hier die Betrachtenden und die Messbarkeit ihrer Reaktionen auf Kunst im Fokus der Betrachtung stehen. Allerdings geht der Ansatz der *Neuroarthistory*, sofern man sie denn überhaupt als kohärenten eigenständigen Bereich begreifen kann,<sup>40</sup> über Fragen nach physiologischen und behavioralen Korrelaten hinaus. Sie setzt einen dezidierten Schwerpunkt auf biologische Determinanten der Rezeption und deren neuronale Verankerung. Während also die Maße der empirischen Bildwissenschaft in der Regel noch im Zusammenhang mit eher phänomenologischen Fragen gelesen und verstanden werden können, legt die *Neuroarthistory* ihren Schwerpunkt auf die Frage nach der *Verankerung* der ästhetischen Erfahrung im biologischen System ‚Mensch‘. Sie greift dazu auf Maße wie die funktionale Magnetresonanztomographie und Elektromyographie zurück, welche beide auf unterschiedliche Weise Aktivitätsmuster diverser Gehirnareale bei der Betrachtung und Auseinandersetzung mit Kunst aufzeichnen.<sup>41</sup> Für die *Neuroarthistory* wie auch für die empirische Bildwissenschaft gilt, dass sie zwar der digitalen Kunstgeschichte zugeordnet werden könnte, durch die Spezifität der hier so zentralen empirisch-experimentellen Forschung am Menschen jedoch ein eigenständiger Begriff gerechtfertigt ist. Eine Abgrenzung von der empirischen Bildwissenschaft ist hingegen nicht zwingend nötig und strenggenommen hauptsächlich aufgrund inhaltlicher, nicht aber aufgrund methodischer Überlegungen sinnvoll.

---

<sup>39</sup> Rosenberg 2016: 308.

<sup>40</sup> Rosenberg merkt hierzu an, dass es zur Bezeichnung der ‚*Neuroarthistory*‘ keinen verbindlichen Terminus gibt; so wird dasselbe Feld unter anderem auch als ‚neural art history‘ bezeichnet (Rosenberg 2016: 314, Fußnote 6).

<sup>41</sup> Z. B. Onians 2007, Clausberg 1999, Zeki 1999, Zeki 1998.

---

Diesen kurzen Überblick zum Anwendungsfeld empirischer Methoden in den Kunstwissenschaften abschließend, können wir festhalten, dass empirische Ansätze nicht bestimmten disjunkten Überbegriffen oder Themenfeldern zuzuordnen sind. Sie kommen in einem breiten Spektrum zur Anwendung, der sowohl einen Objekt-Fokus als auch einen Betrachtenden-Fokus erlaubt. Die Grenzen der verschiedenen Themenfelder sind dabei in methodologischer Hinsicht fließend. Die empirischen Zugänge der *digitalen Kunstgeschichte* sowie die empirischen Zugänge der *empirischen Bildwissenschaft* und der *Neuroarthistory* sind zwar im Detail unterschiedlich, basieren aber auf demselben erkenntnistheoretischen Paradigma. Die Unterscheidung in unterschiedliche ‚empirische Felder‘ ist letztlich inhaltlicher und nicht methodologischer Natur. Empirische Methoden sind damit eindeutig nicht eine Erscheinung eines bestimmten Feldes oder Themas, sondern in allen Bereichen der Kunstwissenschaft präsent.<sup>42</sup>

### 1.2 Allgemeine Klassifikationsmerkmale empirischer Ansätze<sup>43</sup>

Bisher wurde der Begriff *empirisch* umgangssprachlich belassen und nicht ausführlich beleuchtet. Es ist allerdings weder eindeutig noch trivial, was dieser sehr weite und dehnbare Begriff im Zuge einer Methodendiskussion bedeutet (oder besser was er bedeuten sollte). So sind, wie bereits angesprochen, je nach Definition auch viele klassische Methoden der Kunstgeschichte als *empirisch* zu verstehen – etwa wenn auf vergleichende Beobachtungen, die Analyse von Texten oder die Auswertung diversen Bild- und Textmaterials zurückgegriffen wird. Freilich unterscheidet sich solch ein empirischer Zugang deutlich von den sehr technischen Versuchsaufbauten experimenteller Forschung. Heute werden, wie das vorangegangene Kapitel aufgezeigt hat, vor allem letztere, also traditionell

---

<sup>42</sup> Die hier dargelegten Anwendungsbereiche sind dabei nur eine holzschnittartige Auswahl von Beispielen, welche vielfach ergänzt und erweitert werden könnte.

<sup>43</sup> Dieses Kapitel ist in leicht abgewandelter Form als Teil eines Positionspapieres bereits publiziert. Siehe Commare/Brinkmann 2017: 168-171. Die Mehrfachautor\_innenschaft des Textes ist durch das arbeitsteilige Vorgehen bedingt. Eine ausführliche Aufstellung der jeweiligen Einzelleistungen findet sich im Anhang (Anhang *Stellungnahme zur Arbeitssteilung IV*).



---

naturwissenschaftlich-geprägte Zugänge, als *empirisch* bezeichnet. Diese sind jedoch auch nur eine Spielart empirischer Methoden und nicht notwendigerweise als paradigmatisch zu verstehen.

Alle empirischen Methoden können zwei grundlegenden Klassifikationsmerkmalen folgend unterteilt und eingegrenzt werden: dem Datenformat und dem Erhebungsformat. Entlang dieser beiden Merkmale können empirische Methoden in ihren Grundzügen definiert werden, da diese Merkmale hinreichend sind, um die grundlegenden Ziele, Möglichkeiten und Grenzen der unterschiedlichen Ansätze aufzuzeigen. Je nachdem wo auf dieser „Skala“ ein bestimmter empirischer Ansatz angesiedelt ist, desto deutlicher wird er sich von einem anderen, ebenfalls empirischen Ansatz unterscheiden. Empirische Methoden umspannen somit ein breites Feld.

Empirische Methoden reichen von *qualitativen Zugängen*, die den klassischen Ansätzen der Kunstgeschichte durchaus nahe sind, bis hin zu stark formalisierten, häufig technikbasierten *quantitativen Zugängen*, wie sie etwa die Neuroarthistory oder auch die empirische Bildwissenschaft sowie Teile der digitalen Kunstgeschichte verwenden. Gemein ist diesen Methoden, so unterschiedlich sie auf den ersten Blick auch erscheinen mögen, dass sie grundsätzlich datenbasiert arbeiten.<sup>44</sup> Jedoch ist bereits beim verwendeten Datentyp ein wesentlicher Unterschied auszumachen, der weitreichende Konsequenzen für die jeweiligen *Ziele* der unterschiedlichen methodischen Ansätze hat: Während qualitative Ansätze mit „tiefen Daten“ über nur wenige Individuen oder Sachverhalte arbeiten, greifen quantitative Forschungsansätze im weiteren Sinne auf „Oberflächendaten“ von zahlreichen Menschen oder Objekten zurück.<sup>45</sup> In einem weiteren Sinne ist selbstverständlich auch die Arbeit mit klassischen kunst- oder geisteswissenschaftlichen Methoden als datenbasiertes Vorgehen zu verstehen. Allerdings handelt es sich hierbei häufig um literatur- oder bildbasierte Daten, anhand derer beispielsweise im Sinne einer Diskursanalyse Ergebnisse und neue Inhalte

---

<sup>44</sup> Balzert/Schröder/Schäfer 2011: 74.

<sup>45</sup> Manovich 2014: 67.

---

durch vergleichendes In-Beziehung-Setzen und tiefgehende Bild- und Textinterpretationen erarbeitet werden.<sup>46</sup> Der wesentliche Unterschied ist hier, dass im Rahmen empirischer Ansätze die zu analysierenden Daten *in der Regel* erst neu erhoben werden, also ein Datenkorpus geschaffen wird, der zuvor noch nicht vorhanden oder zumindest nicht als solcher klassifiziert war.<sup>47</sup> Empirische Ansätze generieren also (meistens) gänzlich neues Forschungsmaterial, welches nach Abschluss der Untersuchung dem Feld auch weiterhin zur Verfügung steht.

Forschende, welche qualitativen Ansätzen folgen, nutzen die neu erhobenen „tiefen Daten“, um über strukturiert-interpretatives Vorgehen vorliegende Zusammenhänge zu beschreiben und zu verstehen, sowie aus ihnen Hypothesen abzuleiten, die möglicherweise über die wenigen untersuchten Individuen (oder Objekte) hinaus Gültigkeit besitzen. Forschung, die quantitative Ansätze zur Grundlage hat, setzt an diesem Punkt an: Hier werden Hypothesen aufgegriffen, die entweder auf der Basis klassischer Theoriearbeit oder aber im Rahmen qualitativer Arbeit generiert wurden. Diese Thesen werden dann – in der Regel über statistische Maße – daraufhin getestet, inwiefern sie sich als soziale und/oder sachbezogene Regelmäßigkeiten in der Praxis tatsächlich nachweisen lassen.<sup>48</sup> Zur Interpretation quantitativer Ergebnisse muss notwendigerweise wieder auf theoretische Arbeiten zurückgegriffen werden, da mittels quantitativer Methoden lediglich ein numerischer Output produziert wird, der für sich genommen keine Bedeutung besitzt. Quantitative Forschung ist also ausschließlich im Rahmen von Theoriearbeit möglich und kann nicht als eigenmächtiges Werkzeug gedacht werden. Zweck und Funktion von qualitativen Zugängen einerseits und quantitativen Zugängen andererseits unterscheiden sich deswegen deutlich: während qualitative Methodenwerkzeuge als Ergänzung zu theoriegebenden Verfahren zu verstehen sind, stellen quantitative Verfahren eine Möglichkeit dar, die Ergebnisse theoriegebender Arbeit breitflächig zu validieren oder falsifizieren.<sup>49</sup>

---

<sup>46</sup> Balzert/Schröder/Schäfer 2011: 73.

<sup>47</sup> Ebd.: 74.

<sup>48</sup> Ebd.: 76 und Grau 2015.

<sup>49</sup> Vgl. Kindt/Schmidt 1979.

---

Sowohl qualitative als auch quantitative Methoden können zwei groben Überkategorien zugeordnet werden, die das praktische Vorgehen kategorisieren: Beide Ansätze können entweder ein experimentelles oder eines nicht-experimentelles Forschungsdesign zur Grundlage haben.<sup>50</sup> Der Begriff des Experiments bezieht sich hier allerdings nicht auf die Verwendung technischer Messgeräte, sondern meint die systematische Herstellung kontrollierter Bedingungen, die es erlaubt, durch die Variation eines einzelnen Parameters bei gleichzeitigem Konstanthalten aller anderen Größen den Einfluss der veränderlichen Größe exakt zu messen.<sup>51</sup> Der Einfluss der veränderlichen Größe wird hierbei immer über einen Vergleich zweier Zustände gemessen;<sup>52</sup> entweder handelt es sich um einen Vorher-Nachher-Versuch, der innerhalb einer Person (*within-subject*) oder eines Objekts vergleicht, was die Veränderung eines bestimmten Parameters bewirkt hat; oder es handelt sich um den Vergleich zweier Gruppen (*between-subject*), wobei für die Gruppen bis auf die zu untersuchende Größe identische Rahmenbedingungen geschaffen werden.<sup>53</sup> Häufig werden die beiden Experimenttypen kombiniert: wenn in etwa sowohl Vorher-Nachher-Vergleiche angestellt werden als auch gleichzeitig zwei Gruppen verglichen werden.<sup>54</sup> Das wäre zum Beispiel der Fall, wenn sowohl eine Gruppe von Kunstexpert\_innen als auch eine Gruppe

---

<sup>50</sup> Qualitative Methoden wählen in der Regel (aber nicht zwingend) einen nicht-experimentellen Zugang.

<sup>51</sup> Balzert/Schröder/Schäfer 2011: 74.

<sup>52</sup> Dabei ist die erfolgreiche Manipulation der interessierenden Größe durch einen Manipulationscheck nachzuweisen. Der Manipulationscheck kann durch unterschiedliche Maße erbracht werden. Er belegt, dass die Veränderung auf die Manipulation zurückzuführen ist.

<sup>53</sup> Balzert/Schröder/Schäfer 2011: 74, Diekmann 2007: 723ff.

<sup>54</sup> Da selbstverständlich nicht alle intra- und interpersonellen Unterschiede konstant gehalten werden können, gilt für Experimente grundsätzlich das Randomisierungsprinzip. Das bedeutet, dass die Versuchsteilnehmer\_innen oder Objekte zufällig auf die beiden Gruppen verteilt werden müssen, da so die Wahrscheinlichkeit einer proportionalen Überrepräsentation bestimmter Merkmale für beide Gruppen gleich hoch ist und sich dadurch im statistischen Mittel „herauskürzt“ (siehe Diekmann 2007: 723ff). Außerdem müssen Experimente grundsätzlich wiederholbar sein und somit die Replizierbarkeit ihrer Ergebnisse sichern (siehe Balzert/Schröder/Schäfer 2011: 75). Da es viele Faktoren gibt, die eine Untersuchung beeinflussen können, etwa Lichtverhältnisse bei der Betrachtung eines Werkes, der herrschende Lärmpegel, die Gegenwart anderer Personen usw., finden experimentelle empirische Erhebungen häufig, wenn auch nicht zwingend, in Laborsettings statt. Hier lassen sich zumindest einige Größen verhältnismäßig gut kontrollieren und sichern damit die Wiederholbarkeit des Versuches (vgl. z. B. Brinkmann/Commare 2015, Brinkmann et al. 2014).

---

von Kunstlaien, zweimal hintereinander ein zuvor unbekanntes Bild betrachten würde. Die interessierende Größe wäre in diesem Fall die wiederholte Betrachtung bzw. der Unterschied zwischen der Erstbetrachtung und der zweiten Beschäftigung mit dem Werk – zum einen innerhalb der Personen (*within-subject*), aber auch der generelle Unterschied hinsichtlich einer wiederholter Beschäftigung mit demselben Werk zwischen diesen beiden Gruppen (*between-subject*).<sup>55</sup>

Eine nicht-experimentelle Datenerhebung (auch Feldforschung<sup>56</sup>) stellt keine solchen kontrollierten Bedingungen her. Eine Besucherbefragung im Museum, in der alle Befragten einen Fragebogen ausfüllen, wäre hierfür ein Beispiel.<sup>57</sup> Das hat natürlich den Vorteil, dass reale, „ungekünstelte“ Situationen untersucht werden können; allerdings ist hier kein Kausalschluss möglich, sondern nur die Beschreibung eines beobachteten Zusammenhangs. Über diesen kann lediglich unter Zuhilfenahme von theoretischen Positionen eine Vermutung über kausale Wirkverhältnisse getroffen werden.<sup>58</sup> In beiden Fällen ist die *Erklärung* des vorgefundenen Zusammenhangs ausschließlich durch bereits existente Theorie möglich.<sup>59</sup>

Zusammenfassend lässt sich festhalten, dass Datenformat und Erhebungsverfahren die wichtigsten Kriterien zur Klassifikation empirischer Methoden darstellen, da die jeweiligen Instrumente immer diesen Kategorien zuordenbar sein und ihren Voraussetzungen gerecht werden müssen.<sup>60</sup> Diese beiden Klassifikationsmarker setzten gleichzeitig die Eckpfeiler des

---

<sup>55</sup> Für ein konkretes Beispiel vgl. Brinkmann et al. 2014, wo der Expert\_innen-Status eine *between-subject* Kategorie ist, und die Betrachtung verschiedener Gemälde unterschiedlichen Abstraktionsgrades eine *within-subject* Kategorie.

<sup>56</sup> Ebd.: 75.

<sup>57</sup> Z. B. Thalwitzer/Brinkmann/Rosenberg 2015 und Reitstätter 2015.

<sup>58</sup> Der Kausalschluss ist aufgrund der fehlenden Randomisierung und Bedingungskontrolle nicht möglich. vgl. Diekmann 2007: 723.

<sup>59</sup> Vgl. auch Grau 2015 und Kindt/Schmidt 1979: 13f.

<sup>60</sup> Aufgrund der Vielzahl individueller empirischer Methoden und Instrumente verzichte ich an dieser Stelle auf eine gesonderte Besprechung spezifischer Beispiele. Grundsätzlich ist jedoch festzuhalten, dass für die jeweiligen Instrumente, wie auch für den grundlegenden empirischen Ansatz (z. B. quantitativ-experimentell oder qualitativ-nicht-experimentell), bestimmte Grenzen und Probleme bestehen, die vor und während der Anwendung Berücksichtigung finden müssen. So ist

---

Möglichkeitenraums der unterschiedlichen Methoden, womit sowohl die Chancen als auch die Grenzen der jeweiligen empirischen Ansätze transparent werden.

### *1.3 Chancen und Möglichkeiten empirischer Ansätze in der Kunstwissenschaft*

Wie in Kapitel 1.1 deutlich geworden ist, produzieren oder verwenden die meisten empirischen Ansätze in der Kunstwissenschaft größere Datensätze. Das hat – insbesondere im Fall von quantitativ-empirischen Untersuchungen – einen „entkanonisierenden Effekt“, wie es Hubertus Kohle formuliert hat; da hier zwangsläufig der Blick des Forschenden nicht mehr auf das Einzelwerk gerichtet ist, sondern die Breite des Materials favorisiert.<sup>61</sup> Ein solches auf die Breite gerichtetes Vorgehen ist jedoch nicht zwangsläufig eine empirische Perspektive. So merkt etwa Raphael Rosenberg hierzu an: „Quantitative art-historical inquiries were less attractive before the emergence of digital technologies, but they did exist. They serve as a corrective to qualitatively obtained hypotheses and they improve the general validity of qualitative analyses.“<sup>62</sup> Als bekannte Beispiele für die von Rosenberg konstatierte quantitative Perspektive vergangener Arbeiten wären etwa die unvollendete metaphorische Bildenzyklopädie Aby Warburgs, der *Mnemosyne-Atlas*, zu nennen, oder auch Heinrich Wölfflins stilgeschichtlicher Ansatz, der im Wesentlichen als „Kunstgeschichte ohne Namen“ operiert und der auch die so prominente Schrift *Kunstgeschichtliche Grundbegriffe*

---

beispielsweise für alle Instrumente empirischer Messungen wichtig, dass sie intersubjektiv verlässliche und replizierbare Resultate erzielen. Im Bezug auf fragebogenbasierte Forschung bedeutet das beispielsweise, dass sichergestellt sein muss, dass ein und dieselbe Frage von zwei Personen nicht unterschiedlich und bei wiederholtem Einsatz noch immer im gleichen Sinne verstanden wird. Wird nicht auf validierte Instrumente zurückgegriffen, muss ein Test vor dem Einsatz des Messinstruments erfolgen, der die Replizierbarkeit der Messung sicherstellt. Kapitel 4.2.2 wird auf die Spezifika der Methoden im Detail eingehen, welche in den nachfolgend besprochenen Studien zur Anwendung kommen (vgl. auch Diekmann 2007: 479-487).

<sup>61</sup> Kohle bzgl. Bilddatenbanken im Wortlaut: „Datenbanken als die in den Digital Humanities dominierende Organisationsform von künstlerischen Werken versammeln in fortgeschrittenen Ausbaustufen hunderttausende, zuweilen Millionen von Objekten. Recherchen in solchen Datenbanken lenken den Blick zwangsläufig von den großen Einzelwerken ab und favorisieren die Breite. Sie zeitigen einen entkanonisierenden Effekt. Gleichzeitig lassen sich die ›Meisterwerke‹ in ein Verhältnis zur breiten Produktion setzen.“ (Kohle 2016: 152.).

<sup>62</sup> Rosenberg 2016: 305f.

---

hervorgebracht hat.<sup>63</sup> Eine quantitative Perspektive ist der kunstwissenschaftlichen Forschung also weder fremd noch ist sie grundsätzlich als unangemessen zu bewerten. Dennoch muss an dieser Stelle eingewendet werden, dass der „quantitative Blick“, wie wir ihn in klassischen Forschungsätzen zuweilen finden, nicht völlig deckungsgleich mit jenem ist, was empirische Ansätze charakterisiert, die den Natur- oder Sozialwissenschaften entlehnt sind. So sind etwa die quantitativen Perspektiven Warburgs und Wölfflins trotz ihrer Systematik doch stark an der detaillierten Auseinandersetzung mit Fallbeispielen orientiert.<sup>64</sup> Der hier zu Tage tretende, quantitative Charakter ist daher im Sinne der im vorherigen Abschnitt vorgestellten Klassifikationsmerkmale empirischer Arbeiten, allenfalls ein ideeller. Das tatsächliche Vorgehen, die methodische Umsetzung der ideellen quantitativen Perspektive, folgt jedoch eigentlich einem *qualitativen Forschungsparadigma*. Was Kohle also hervorhebt, wenn er von einem Perspektivenwechsel spricht, geht über die in klassischen Ansätzen schon vorhandene Idee einer quantitativen Perspektive hinaus. Hier ist tatsächlich eine *echte Quantität* gemeint, eine Vielfalt von Daten in enormem Umfang. Die Analyse eines solchen Datenkorpus war vor der Etablierung computergestützter Techniken nicht nur – im Sinne Rosenbergs – „unattraktiv“, sondern praktisch gesehen unmöglich. Einer der wesentlichen Punkte, die Befürworter empirischer Ansätze deshalb häufig ins Feld führen, seien die jeweiligen Ansätze nun eher der *digitalen Kunstgeschichte* oder der *empirischen Bildwissenschaft* zuzuordnen,<sup>65</sup> sind die Möglichkeiten, welche sich durch eine echte empirisch-quantitative Analyse eröffnen.

---

<sup>63</sup> Siehe Brink 2000, Wölfflin 1915; vgl. auch Kohle/Brassat 2009: 51.

<sup>64</sup> Vgl. Raphael Rosenberg: “[...] art history has never taken institutional ownership of systematic research. Systematic questions are discussed in relation to historical examples by researchers with established historical knowledge.” (Rosenberg 2016: 312.).

<sup>65</sup> Darüber ob – und wenn ja, wie – Grenzen zwischen den Begriffen zu ziehen sind, herrscht derzeit kein Konsens. So ordnet etwa der Arbeitskreis für digitale Kunstgeschichte in Deutschland die empirische Bildwissenschaft der digitalen Kunstgeschichte zu (vgl. dazu die Auflistung der durch den Arbeitskreis definierten, wichtigsten Themenfelder der digitalen Kunstgeschichte: <http://www.digitale-kunstgeschichte.de/>). Vertreter der empirischen Bildwissenschaft hingegen grenzen sich begrifflich teils explizit von der Zuordnung zur digitalen Kunstgeschichte ab. Beispielweise Rosenberg: “Using those devices produces large amounts of data that have to be analyzed with specific software. The collection of psychophysiological measurements (by art historians) might therefore be considered as part of digital art history. I prefer to call it ‘cognitive

---

Die sich durch empirische Ansätze erschließenden Möglichkeiten sind vielschichtig und je nach Ansatz ganz unterschiedlicher Art:<sup>66</sup>

## I.

Zum einen sind sie in der Lage *ideell-quantitative Perspektiven mit einem adäquaten Methodeninstrumentarium zu unterstützen*. Damit tragen sie dazu bei, Einzelfallstudien ihres Podestes zu entheben, wobei sie gleichzeitig deren Gültigkeit untermauern (können). Ein Beispiel hierfür wäre etwa Lev Manovichs Studie zu Vincent van Goghs Werk in den Jahren 1886-1889, dem Zeitraum in welchem van Gogh von Paris nach Südfrankreich übersiedelte.<sup>67</sup> Manovichs Untersuchung, in dessen Rahmen van Goghs Werke auf zwei Achsen, nämlich Helligkeit und Farbsättigung, angeordnet werden, kommt zum Ergebnis, dass insbesondere die späteren Arbeiten durch eine hellere Farbgebung und eine höhere Farbsättigung charakterisiert sind. Eine Beurteilung dieses Ergebnisses, welche betont, wie die von Manovich gewählte, auf quantifizierbaren Parametern beruhende Analyse qualitative Zugänge zu van Goghs Werk unterstützen kann, äußert zum Beispiel Kohle, der dazu schreibt: „[...] hier werden Skeptiker anmerken, dass dies angesichts des intensiveren Lichtes in Südfrankreich naheliegt, aber die Feststellung, dass die Außenbedingungen empirisch messbar in die künstlerische Produktion eingeflossen sind, scheint mir doch keineswegs selbstverständlich oder trivial.“<sup>68</sup> Selbstverständlich steht die empirisch-motivierte Analyse dabei nicht für sich selbst, denn wie diverse Autor\_innen, darunter auch Hubertus Kohle oder

---

research in art history’ (or in German: empirische Bildwissenschaft), to underline the psychophysiological, experimental element that precedes digital processing.” (Rosenberg 2016: 308.).

<sup>66</sup> Siehe z. B. Bernhard Rieder und Theo Röhle zu digitalen empirischen Tools: “These computational tools hold a lot of promise: they are able to process much larger corpora than would ever be possible to do manually; they provide the ability to seamlessly zoom from micro to macro and suggest that we can reconcile breadth and depth of analysis; they can help to reveal patterns and structures which are impossible to discern with the naked eye; some of them are even probing into that most elusive of objects, meaning.” (Rieder/Röhle 2012: 67).

<sup>67</sup> Manovich 2015, Manovich 2011.

<sup>68</sup> Kohle 2013: 85.

---

der hier im Folgenden zitierte Raphael Rosenberg festhalten: „[...] no quantitative analyses in art history have ever been successful without qualitative analyses. Statistics about works of art are unobtainable in the absence of some form of prior qualitative analysis, without which the works could not be categorised in the first place. Also, statistics themselves need to be interpreted.“<sup>69</sup> Somit steht die klassische Einzelfallstudie nicht etwa neben oder abseits einer empirisch-arbeitenden Kunstwissenschaft, sondern ist vielmehr als deren Voraussetzung zu sehen. Oder um es mit Hubertus Kohles Worten zu sagen: „All dies geschieht in einem Geist, der traditionelle Verfahrensweisen nicht etwa überflüssig macht, sondern diese in ein neues, überwölbendes Paradigma einbindet.“<sup>70</sup> Empirische Zugänge erlauben es somit bestehende Fragen mit einem breiteren Instrumentarium auf breiterer Basis anzugehen und ermöglichen so die Ausweitung des gesicherten Erkenntnishorizontes.

## II.

Eine weitere Möglichkeit, die sich durch den Einsatz von empirischen Methoden eröffnet, ist das *Aufzeigen unbeantworteter Fragen*. Der meist eher auf die breite Systematik angelegte Untersuchungsrahmen ist in der Lage Regelmäßigkeiten, aber auch Unregelmäßigkeiten in der untersuchten Struktur zu identifizieren, die dem bloßen Auge häufig – aufgrund der schier unermesslichen Materialmasse – nicht zugänglich sind. So ist es etwa möglich, über Netzwerkanalysen diverse Strukturen in ihrer Gesamtheit gleichzeitig sichtbar zu machen und dadurch enorme Mengen an einzelnen Datenpunkten eingebettet in ihr wechselseitiges Bezugssystem zu betrachten. Ein Beispiel hierfür wäre Maximilian Schichs Analyse des Kunstmarktes in den Räumen Großbritannien, Belgien, Frankreich und den Niederlanden in

---

<sup>69</sup> Rosenberg 2016: 306.

<sup>70</sup> Kohle 2013: 8; Diese Haltung drückt Kohle auch an anderer Stelle in ähnlicher Form aus: „Ästhetische und historische Eigenschaften auf der Basis von Messung zu beschreiben, ist eine Domäne des Digitalen in der Kunstgeschichte, das auf wie komplexe Weise auch immer nichts anderes unternimmt als zu rechnen. Quantitative Bestimmungen haben in den hermeneutisch grundierten Geisteswissenschaften allerdings keinen guten Ruf. Die Digital Humanities mit ihrer Neigung zu statistischen Verfahren könnten zu deren Rehabilitierung beitragen. Die kennerschaftliche Perspektive wird dadurch ergänzt, nicht ersetzt.“ (Kohle 2016: 151.).



---

den Jahren 1801 bis 1820. Die Grundlage dieser Analyse waren 230.000 Datenbankeinträge des *Getty Provenance Index*. Das Ergebnis ist eine Netzwerk-Karte, die sowohl Verkäufer als auch Käufer sowie deren Umsatzstärke über den Zeitraum von 20 Jahren gleichzeitig darstellt; wodurch ein visuell-greifbares Netzwerk mit vielen Einzelknoten unterschiedlicher Relevanz entsteht.<sup>71</sup> In der Auseinandersetzung mit diesem Ergebnis ist auffällig geworden, dass der Kunsthandel deutlich internationaler ausgerichtet war als von der Forschung bisher angenommen.<sup>72</sup> Ein Hinweis auf eine Forschungslücke, die nur dank der quantitativen Perspektive der Analysemethode offenbar werden konnte. Sichtbar wird dadurch jedoch lediglich die Lücke selbst, nicht die Gründe für die noch unerklärten beobachteten Strukturen, denn, wie Kohle richtig betont: „Als ein Verfahren, das Korrelationen zwischen Phänomenen aufweisen kann, deren kausale Verbindung aber letztendlich durch eine schlussfolgernde Intelligenz zu bestimmen ist, können die Digital Humanities keine Erklärungen liefern, sondern immer nur Hinweise darauf geben, wo etwas Erklärungsbedürftiges auftritt. Sie bieten mithin Material, um Argumentationen zu unterfüttern, und können als wissenschaftliche Recommender-Systeme fungieren.“<sup>73</sup> Außerdem ist anzumerken, dass kein Datensatz, und sei er noch so groß, jemals vollständig sein kann. Deswegen hat eine auf solchen Daten basierende Analyse keinen allgemeinen Gültigkeitsanspruch, sondern bildet lediglich Näherungswerte ab. Das ist jedoch, um es abermals mit Kohles Worten zu sagen, nicht notwendigerweise problematisch, denn hinsichtlich dieser auf Grundlage breiter Daten entstandenen Näherungswerte „[...] sollte man sich klar machen, dass statistisch gesehen eine Näherung ziemlich viel wert ist – und vor allem mehr wert als eine Eindeutigkeit auf der Basis von zu wenigen Daten.“<sup>74</sup>

---

<sup>71</sup> Schich et al. 2017; ein Beispiel einer solchen Netzwerk-Karte findet sich auch Online unter: <http://www.getty.edu/research/tools/provenance/zoomify/index.html> (letzter Abruf: 27.10.17).

<sup>72</sup> Vortrag “Mapping Markets. Der Getty Provenance Index auf dem Prüfstand” von Christian Huemer am 06.12.2016 im Rahmen des Panels „Collections, Collecting and the Art Market. Digital perspectives for Art Historical Research I” auf der 3<sup>rd</sup> DHA Conference in Wien (Huemer 2016).

<sup>73</sup> Kohle 2016: 152.

<sup>74</sup> Kohle 2013: 92.

---

### III.

Die vielleicht spannendste, für die vorliegende Arbeit aber in jedem Fall wichtigste Möglichkeit, die sich durch das Heranziehen empirischer Ansätze eröffnet, ist keine der beiden bereits genannten, sondern eine weitere, dritte Möglichkeit: das *Erschließen neuer Themen- und Fragenbereiche*. Wie weiter oben bereits angesprochen wurde, sind es insbesondere rezeptionsästhetische Ansätze, die von einer empirischen Prüfung profitieren können, da somit gesicherte Aussagen über die hier zugrundeliegenden Thesen getroffen werden können, welche explizit auf tatsächliche Prozesse menschlicher Wahrnehmung und Erfahrung rekurren. Wenn auch viele der grundlegenden Gedanken des Ansatzes deutlich älter sind, so kommt die Rezeptionsästhetik doch erst mit Wolfgang Kemp als methodisches Vorgehen in Mode.<sup>75</sup> Die wesentliche Frage der hier gemeinten rezeptionsästhetischen Perspektive dreht sich dabei in der Regel immer um die emotionale und kognitive Wahrnehmung von Kunstwerken und deren Vorortung; damit ist gemeint, es stellt sich die Frage, inwiefern eine solche Wahrnehmung im Kunstwerk selbst verankert bzw. angelegt ist, also durch dieses notwendig evoziert wird, oder inwiefern die Wahrnehmung und Verarbeitung von Kunst erst im Prozess der Rezeption entsteht, also wesentlich im Betrachter angelegt ist. Die Inklusion soziologischer, semiotischer und vor allem psychologischer Theorie liegt bei dieser Art Fragen nahe, da deren Relevanz für alle in diesem Kontext erarbeitbaren Thesen offensichtlich ist. Es sind allerdings nicht ausschließlich Fragestellungen, die genuin als rezeptionsästhetische Überlegungen markiert werden können, welche im weiteren Sinne auf Thesen zur Rezeption zurückgreifen. So merkt beispielsweise Rosenberg in seiner Stellungnahme zur Lage der Kunstgeschichte an, dass das fruchtbare Wechselspiel von psychologischer Theorie und kunsthistorischen Überlegungen lange Tradition hat: “Psychological theories have been influencing the basic assumptions of art history for a long time, from empathy theory and gestalt psychology to Gombrich’s interpretation of ecological psychology.”<sup>76</sup> Hierbei ist es insbesondere die Frage nach der

---

<sup>75</sup> Vgl. Kohle/Brassat 2009, Kemp 1992, Kemp 1986.

<sup>76</sup> Rosenberg 2016: 310.

---

Wahrnehmung von Kunst und deren personeller, kultureller oder historischer Bedingtheit, die häufig Basis für weitreichende Theoriebildung war: “The notion of historical variations in perception was and still is a basic assumption for many art historians – from Moritz von Thausing’s inaugural lecture in Vienna (1873), through Alois Riegl and Heinrich Wölfflin to Michael Baxandall and Jonathan Crary.”<sup>77</sup> Im Rahmen dieser unterschiedlichen Ansätze, die doch alle in der einen oder anderen Form auf der Basis von Annahmen zur menschlichen Wahrnehmung argumentieren, zeigt sich häufig eine selbstreferenzielle Tendenz. Diese wurde von Rosenberg in besagter Stellungnahme sehr treffend beschrieben: „Assumptions such as this tend to steer historical analyses, though the analyses themselves are often used to confirm the very assumptions they make.”<sup>78</sup> Diese Selbstreferentialität kann mittels eines empirischen Methodenzugangs gebrochen werden.

Viele Aspekte menschlicher Wahrnehmung, seien sie nun emotionaler oder kognitiver Natur, sind durch aus den Natur- und Sozialwissenschaften entlehnte Techniken abbildbar. Damit wird die Praxisrelevanz von Thesen zur Rezeption, der Grad in welchem hier tatsächlich Konzepte abgebildet sind, die echte Wahrnehmungsvorgänge beschreiben, überprüfbar – die tönernen Füße kunstwissenschaftlicher Rezeptionstheorien erhalten somit ein solides Fundament.<sup>79</sup> Die mittels empirischen Untersuchungen gewonnenen Erkenntnisse bieten nicht nur die Möglichkeit, bestehende Theorien durch die Verifikation bestimmter Thesen zu stützen oder anzureichern, sondern sie können ebenfalls für die Weiterentwicklung im Theoriebildungsprozess herangezogen werden; nämlich immer dann, wenn bestehende Thesen sich empirisch nicht bewähren und damit ein Überarbeitungsbedarf augenscheinlich wird. Der Rückgriff auf empirische Methoden in einem Kontext, der Theorie- und Erkenntnisaspekte diverser Disziplinen vereint, hat dabei einen leicht „utopischen Charakter,

---

<sup>77</sup> Rosenberg 2016: 311.

<sup>78</sup> Ebd.

<sup>79</sup> Siehe hierzu auch Rosenberg: „[...] we now have sophisticated and non-invasive means of investigating the emotional and cognitive processes involved in the reception and even production of art. The new experimental approach allows us to verify basic assumptions and to elaborate new theories on the basis of empirical evidence.” (Rosenberg 2016: 310f.).

---

der die Demokratisierung von Kultur und Wissenschaft beinhaltet und gleichzeitig die Schranken zwischen Geistes-, Sozial- und Naturwissenschaften niederlegt.“<sup>80</sup> Fragen- und Teilbereiche, welche für kunstwissenschaftliche Überlegungen zwar zentral sind, von der Kunstwissenschaft in der Regel aber nicht selbst in der ihnen angemessenen (empirischen) Form behandelt wurden, können so in einen Geistes-, Natur- und Sozialwissenschaften umspannenden Wissensraum inkludiert werden und damit dem langen historischen Prozess, der sich in spezialisiertes Nischenwissen ausdifferenzierten Disziplinen entgegenwirken.<sup>81</sup> Die Kunstgeschichte, wie Hubertus Kohle bemerkt „[...] öffnet sich an dieser Stelle zwanglos in höchst spannende Fragen der Interkulturalität, der Wahrnehmungspsychologie, der Geschlechterforschung, ohne sich hier zwanghaft der politischen Forderung einer Inter- oder Transdisziplinarität hinzugeben.“<sup>82</sup>

Wie aus diesem kurzen (und keineswegs vollständigen) Abriss der Möglichkeiten empirischer Zugänge zur Kunstwissenschaft hervorgeht, eröffnet die empirisch-arbeitende Kunstwissenschaft viele unterschiedliche Chancen. Empirische Zugänge können einerseits ein unterstützendes Werkzeug und „Recommender-Tool“ für klassische Fragestellungen sein. Andererseits können sie aber auch dazu beitragen die historisch-bedingte Verkrustung und Einnischung des Faches zu überwinden und damit neue Themenbereiche erschließen bzw. es ermöglichen, Verantwortung für „neue“ Teilbereiche zu übernehmen, die strenggenommen nicht wirklich „neu“ sind. Gerade letzteres ist keineswegs trivial, denn – wie Rosenberg bemerkt – die Kunstwissenschaft hat längst kein Monopol mehr für die Analyse von Kunst: „Information scientists working on computer vision are developing algorithms for the analysis of pictures; psychologists are conducting a number of studies on the viewing of art, sometimes in interdisciplinary networks within ‘empirical aesthetics’; some neuroscientists are relating artistic regularities to the anatomy and function of the brain

---

<sup>80</sup> Kohle 2013: 7.

<sup>81</sup> Ebd.

<sup>82</sup> Ebd.: 95.

---

and trying to establish neuroaesthetics as an independent discipline.”<sup>83</sup> Angesichts dieser fachlichen Ausweitung benachbarter Disziplinen stellt sich die legitime Frage nach der Positionierung der Kunstwissenschaft in diesem Raum. Forschende, die einen empirischen Zugang zum Fach aus den hier skizzierten Gründen befürworten, sehen in der Überwindung der traditionellen Trennung von geistes- und natur- bzw. sozialwissenschaftlichen Perspektiven und Methoden keine Bedrohung, sondern vielmehr eine Chance einen besseren Blick auf das „Big Picture“ zu erhaschen. Ein Standpunkt, der von Cathy N. Davidson sehr anschaulich beschrieben wurde: “Perhaps we need a paradigm shift. Perhaps we need to see technology and the humanities not as a binary but as two sides of a necessarily interdependent, conjoined, and mutually constitutive set of intellectual, educational, social, political, and economic practices.”<sup>84</sup>

Eine Haltung, welche die Perspektiven, Methoden, und Forschungsansätze von Geistes- und Natur- bzw. Sozialwissenschaften nicht als Antipoden, sondern als zwei Seiten einer Medaille versteht, ist jedoch, wie eingangs dargelegt, nur *eine* Position im Raum der Frage nach der Positionierung der zeitgenössischen Kunstwissenschaft. Forschende, die gegenüber empirischen Zugängen zu den Kunstwissenschaften eine eher skeptische Position einnehmen, teilen diese Haltung aus diversen Gründen entweder überhaupt nicht – oder nur sehr bedingt.

---

<sup>83</sup> Rosenberg 2016: 312; Konkrete Beispiele für die von Rosenberg hier geäußerten Ansätze finden sich in den eigens für diesen Kontext ins Leben gerufenen Zeitschriften *Empirical Studies of the Arts* und *Psychology of Aesthetics, Creativity, and the Arts*. Wobei erstere die offizielle Zeitschrift der Internationalen Vereinigung für Empirische Ästhetik ist (International Association of Empirical Aesthetics) und zweitere eine offizielle Zeitschrift mit Schwerpunkt Ästhetik der Amerikanischen Vereinigung für Psychologie (American Psychological Association, APA) darstellt.

<sup>84</sup> Davidson 2008: 709.

---

#### 1.4 Kritik und Probleme empirischer Ansätze in der Kunstwissenschaft

Eine empirisch-ausgerichtete Kunstgeschichte ist streng genommen kein zeitgenössisches Phänomen, denn der sogenannte *empirical turn* ist eigentlich schon der zweite seiner Art. Um 1900, etwa zeitgleich mit der Gründungsphase des Faches Kunstgeschichte, erlebten empirische Zugänge zum Feld erstmalig eine Konjunktur. Im Kontext eines fruchtbaren Wechselverhältnisses von Psychologie und Kunstgeschichte wurde damals der Versuch unternommen eine *empirische Ästhetik* als systematisch-angelegten Forschungsbereich zu etablieren. Mit dem Ersten Weltkrieg erlebt dieser erste „*empirical turn*“ der Kunstwissenschaften allerdings eine Zäsur und wird später (zunächst) nicht wieder aufgegriffen.<sup>85</sup> Obwohl sich die während der ersten empirischen Wende vertretenen Ansätze zunächst nicht durchsetzen konnten und neueste empirische Ansätze nicht im direkten Verhältnis zu diesen historischen Zugängen stehen, besteht doch eine unbestreitbare Analogie in den grundlegenden *Zielen* beider Bewegungen. So zeigt sich etwa in beiden Bewegungen das starke Bestreben ein Korrektiv zu etablieren, welches es gestattet, die häufig ausschließlich theoretische Argumentation der Kunstgeschichte auf ihre Praxisrelevanz zu prüfen. Dies gilt insbesondere für jene Teilbereiche des Faches, welche Aussagen über die menschliche Wahrnehmung treffen oder auf dieser Grundlage arbeiten. Hier ist ohne empirische Prüfung häufig nicht zu klären, inwiefern plausibel argumentierte, aber dennoch rein theoretische Positionen tatsächlich wahr sind. Klarheit über die zugrunde gelegten Prämissen ist jedoch notwendig, wenn tiefgreifende Fehlschlüsse vermieden werden sollen. Während dies, aus einer theoretischen Perspektive betrachtet, keine kontroverse

---

<sup>85</sup> Vgl. u. a. Commare/Brinkmann 2017; Obschon von Kunsthistoriker\_innen selbständig durchgeführte empirische Arbeiten eine Seltenheit bleiben, ist das Ausmaß der *empirisch-informierten* Arbeiten beachtlich. Für eine umfassende Übersicht zum ersten „*empirical turn*“ sowie für konkrete Beispiele zur fruchtbaren Wechselbeziehung von Psychologie und Kunstwissenschaft siehe: Jarzombek 1994, Büttner 2003, Guthmüller/Klein 2006, Bushart 2007, Clausberg 2011, Vasold 2011, Lanzoni/Brain/Young 2012, Moravánszky 2012, Collenberg-Plotnikov 2013, Mitrović 2013, Müller-Tamm/Schmidgen/Wilke 2014.

---

Aussage darstellt, scheint die Forderung nach *tatsächlicher empirischer* Prüfung jedoch problematisch.

Wie zuletzt ganz besonders in einigen populär platzierten Beiträgen von Peter Geimer deutlich wurde, besteht ein tiefgreifender Zweifel an der *Möglichkeit* empirischer Forschung, die Konzepte der Geistes- bzw. Kunstwissenschaften adäquat zu adressieren. So ist etwa in Geimers Beitrag im Feuilleton der Frankfurter Allgemeinen Zeitung vom 23.11.2016 zu lesen: „Man müsste dieser Selbstaufgabe geisteswissenschaftlicher Kritik keine größere Beachtung schenken, wenn sie nicht einem aktuellen Trend zur Enthistorisierung und Essentialisierung der Kunstgeschichte entsprechen würde. So werden im Wiener Kunsthistorischen Museum in Zusammenarbeit mit einem Forschungsprojekt der dortigen Universität Experimente zur Messung der Blickbewegungen der Besucher durchgeführt, die den komplexen Vorgang der Wahrnehmung von Kunst auf basale Körperfunktionen reduzieren. Solche Versuche beruhen auf einer verbreiteten Verwechslung von Wirklichkeit und Messbarkeit: als käme allem, was messbar ist, schon von sich aus ein gehobener Grad an Wirklichkeit zu, ganz gleich wie unterkomplex die Ergebnisse sind.“<sup>86</sup> An anderer Stelle schreibt Geimer dazu: „Die Suche nach der >biologischen Verankerung< der Kunst beruht auf dem Vorurteil, dass Verfahren der Naturwissenschaften (wie Messung und Experiment) per se solides und generalisierbares Wissen erzeugen und auch für die Forschungsgegenstände der Kunstgeschichte zuständig sind – eine Vorstellung, die durch wissenschaftliche Arbeiten längst relativiert ist. Sie erinnert aber auch an eine Praxis, die Roland Barthes bereits von Jahrzehnten zu recht als >Mythologisierung< kritisiert hat: die Verwandlung kulturell und historisch verfasster Phänomene in vermeintliche >Natur<, um ihnen den Anschein unhinterfragbarer Evidenz zu geben.“<sup>87</sup>

Jedoch zweifeln die kritischen Stimmen nicht nur an der *Möglichkeit* empirischer Forschung. Die Absage an die Empirie ist häufig deutlich radikaler. Nicht selten wird im Zuge der

---

<sup>86</sup> Geimer 2016: 9.

<sup>87</sup> Geimer 2015: 340.

---

kritischen Reflexion auch ihre *Sinnhaftigkeit* grundsätzlich in Frage gestellt. Peter Geimer etwa – um bei demselben Beispiel zu bleiben – schreibt: „[...] dieses Unternehmen beruht auf der unbegründeten Annahme, dass sich mit der Erhebung empirischer Daten im Bereich der Kunst schon von sich aus wissenschaftliche Seriosität erzeugen lässt. Solche Datenerhebungen tragen jedoch in der Regel nichts zum besseren Verständnis der Kunstwerke bei, da sie diese als bloße Reizauslöser betrachten und ihr Augenmerk auf die Quantifizierung von Körperfunktionen lenken. Ob das ein Erkenntnisgewinn für die Beschäftigung mit Kunst ist, kann bezweifelt werden: Wenn jemand beim Anblick eines niederländischen Stilllebens Appetit auf Meeresfrüchte bekommt und sich diese Reaktion physiologisch messen ließe, wäre das Resultat noch kein Beitrag zur Kunstgeschichte der Niederlande.“<sup>88</sup> Etwas defensiver fährt er fort: „Solche Skepsis an der geisteswissenschaftlichen Mimikry der Laborwissenschaften hat nichts mit Technikfeindlichkeit, mangelnder Neugierde oder dem Festhalten an etablierten Positionen zu tun. Wenn man die Studien aber nicht nach an ihrer szientistischen Rhetorik, sondern an ihren Resultaten misst, zeigt sich in der Regel, dass hier entweder mit hohem finanziellen und apparativen Aufwand experimentell ermittelt wurde, was man auch vorher schon wusste (etwa dass Experten Gemälde anders betrachten als Laien), oder aber Aussagen formuliert werden, deren wissenschaftliche Subtilität bezweifelt werden kann (etwa dass Männer in Ausstellungen Unterhaltung suchen, Frauen gerne die Begleittexte lesen und ansonsten emotional angesprochen werden möchten).“<sup>89</sup> Kritik in noch verschärfterem, nahezu polemischen Tenor äußert Kia Vahland im Feuilleton der Süddeutschen Zeitung. Am 17.05.2010 erscheint dort ihre Besprechung von Martin Tröndles groß angelegtem Projekt, welches zum Ziel hat die Erfahrung der Besucher\_innen des St. Galler Museums möglichst umfassend empirisch zu erfassen.<sup>90</sup> In Vahlands Bewertung dieses Unterfangens scheint nicht nur die Frage nach der *Sinnhaftigkeit* eines solches Vorgehens auf, sondern überdies ein Unbehagen mit den *von ihr* gedeuteten Konsequenzen empirischer Forschung im Bereich

---

<sup>88</sup> Geimer 2016: 9.

<sup>89</sup> Ebd.: 9.

<sup>90</sup> Für einen umfassenden Einblick in Tröndles Arbeit siehe: Kirchberg/Tröndle 2015, Tröndle/Tschacher 2012, Tröndle et al. 2014, Tschacher et al. 2012, Wintzerith/van den Berg/Tröndle 2011.



---

Museums- und Ausstellungswesen: „Wohin aber mag es führen, wenn jede Körperregung auch der heutigen Museumsbesucher durchleuchtet wird wie in den Warenhäusern und Supermärkten, für die solche Messmethoden einst erfunden wurde? Dann müssen Kuratoren bald nach dem Prinzip Käsergal hängen – demnächst abgelaufene Ware immer in Blickrichtung, die guten Stücke etwas versteckt. Keine Texte mehr auf Ausstellungswände, weil die sicher keiner liest. Sponsorengelder gäbe es nur noch nach vorigem Probelauf – anhand einer Ausstellung von Kopien vor repräsentativ ausgewähltem Publikum müsste sich die Kandinskyschau erst einmal beweisen. Und Kopfschmerzkunst oder Klaustrophobisches von Gregor Schneider würde in separate Gruselkabinette verbannt. [...] Auch wenn die Forscher aus St. Gallen dies nicht wollen: Ihre Besuchermessung verabsolutiert momentane Befindlichkeit und riskiert mit dieser schon ignoranten Selbstbezüglichkeit der Heutigen, alles Gemachte, historisch Gebundene der Kunst zu verdrängen. [...] Doch die Kunst wird auch Mediziner überleben, die die Wirkung des Ästhetischen nur aushalten, wenn sie sie vermessen.“<sup>91</sup>

Diesen sehr knapp umrissenen, kritischen Positionen sind – jenseits ihrer Polemik – *drei* wesentliche Punkte zu entnehmen:

## I.

Erstens zeigt sich, dass das notwendig *reduktionistische Vorgehen empirischer Forschungsansätze* dem geisteswissenschaftlichen Arbeiten als unangemessen empfunden wird. Die meist eher holistisch verstandenen Konzepte geisteswissenschaftlichen Denkens basieren auf einer notwendigen Gleichzeitigkeit ihrer Teilelemente. Eine Reduktion auf einzelne Teilaspekte, bzw. deren gesonderte Betrachtung, läuft somit dem Grundgedanken der hier zur Diskussion stehenden Inhalte oder Konzepte zuwider.<sup>92</sup> Wenn Peter Geimer also in der Anwendung empirischer Methoden eine „Verwechslung von Wirklichkeit und

---

<sup>91</sup> Vahland 2010: Online, ohne Seitenzahl.

<sup>92</sup> Vgl. auch Kohle 2018: 17.

---

Messbarkeit“ vermutet, dann ist damit ausgedrückt, dass reduktionistisch-operierende Ansätze im Wesentlichen ein anderes, von holistisch-gedachten geisteswissenschaftlichen Konzepten verschiedenes, Wirklichkeitsbild zugrunde liegt. Die für die Geisteswissenschaften eigentlich interessanten Aspekte werden somit durch die Anwendung empirischer Methoden verfälscht bzw. überhaupt nicht wiedergegeben. Das heißt, dass der *Möglichkeit* empirischer Forschung auf geisteswissenschaftlich relevante Konzepte einzugehen, eine grundsätzliche Absage erteilt wird.

## II.

Neben der Frage nach den spezifischen Möglichkeiten wird in den kritischen Positionen eine zweite Skepsis deutlich: Die Ergebnisse empirischer Forschungsarbeit im Kontext der Geistes- und Kunstwissenschaften gelten als redundant. Diese Behauptung einer unnötigen Wiederholung betrifft zwei Aspekte: a) die Praxis der *Überprüfung von theoretisch etablierten und somit als fundiert empfundenen Annahmen* und b) die Relevanz zeitgenössischer Messungen für *kulturhistorische Fragestellungen*.

### a.

In der Grundsatzkritik an der Praxis der empirischen Prüfung von theoretischen Annahmen werden die Forschungsperspektiven der Geisteswissenschaft und Naturwissenschaft in Opposition gestellt. Hermeneutisch-orientierte Ansätze erachten Beobachtbarkeit und somit Messbarkeit nicht als notwendige Voraussetzung für den Nachweis von Praxisrelevanz. Empirisch-analytische Ansätze hingegen nehmen theoretische Annahmen niemals ungeprüft hin, da sonst Gefahr bestünde, dass sie lediglich einer (falschen) subjektiven Perspektive geschuldet sein könnten. Der Vorwurf, dass eine empirisch-arbeitende Kunstwissenschaft „mit hohem finanziellen und apparativen Aufwand experimentell ermittelt [...], was man

---

auch vorher schon wusste“<sup>93</sup> ist demnach den unterschiedlichen zugrundeliegenden erkenntnistheoretischen Paradigmen geschuldet.<sup>94</sup>

b.

Ein Nebeneinander beider Perspektiven, in welchem sie einander ergänzende und komplementierende Teile wären, ist hierbei nicht ohne weiteres möglich. Das liegt daran, dass das zeitgenössische Messen für *kulturhistorische* Fragestellungen als irrelevant erachtet wird.<sup>95</sup> Das gilt natürlich in besonderem Maße für die Untersuchung und Befragung von Menschen und weniger im Falle einer empirischen Analyse von Objekten. Die Arbeit mit zeitgenössischen Betrachtenden, so die Annahme, erlaubt uns allenfalls Rückschlüsse über Zeitgenossen, nicht aber über Betrachtende vergangener Zeiten. Die Bewertung von empirischen Ansätzen als ahistorisch, gekoppelt mit einem hermeneutisch-geprägten Forschungsideal, führt also zur Überzeugung, dass der Rückgriff auf empirische Methoden für *kunsthistorische* Fragestellungen kaum als sinnvoll zu erachten sein kann.

Die Punkte II.a und II.b zusammengenommen heißt das, dass die kritischen Stimmen nicht nur die *praktische Möglichkeit* empirischer Ansätze anzweifeln, auf geisteswissenschaftliche Konzepte einzugehen, sondern auch die grundsätzliche *Sinnhaftigkeit* von empirischer Forschung vor dem Horizont klassischer Fragestellungen hinterfragen.

III.

Der dritte Punkt, der sich deutlich in den Positionen der Kritik abzeichnet, ist ein Unbehagen mit den vermeintlichen Konsequenzen empirischer Arbeit im geisteswissenschaftlichen Sektor. Damit ist nicht nur die von Kia Vahland sehr anschaulich beschriebene Angst

---

<sup>93</sup> Geimer 2016: 9.

<sup>94</sup> Vgl. auch Floridi 2015, Pollock 2014, Hayles 2012, Anderson 2009.

<sup>95</sup> Vgl. auch Rushkoff 2014.

---

gemeint, dass der Kulturbetrieb einer Ökonomisierung anheimfallen könnte, sondern vor allem die sich dahinter verbergende Annahme, dass empirische Arbeiten ihre Messergebnisse in den absoluten Vordergrund ihrer Argumentation stellen. Damit würde die Messung an sich mehr Gewicht erhalten als die theoretische Rahmung. Die Konsequenz wäre eine, wie Peter Geimer es beschrieben hat, „Essentialisierung der Kunstgeschichte“.<sup>96</sup> Wertvolle Kompetenzen und Aufgabengebiete, wie etwa die hermeneutische Aufarbeitung klassischer Fragen- und Problemstellungen oder auch das geschulte Sehen des Connaisseurs, würden durch eine solche „Essentialisierung“ verdrängt; der Primat kunstwissenschaftlicher Forschung würde sich auf die numerischen Ergebnisse messbarer Konzepte verlagern, die *Erforschung von Kultur und Geschichte würde einer Theoriearmut anheimfallen*.<sup>97</sup> Dieses Unbehagen mit der vermeintlichen Schwerpunktsetzung empirischer Forschung impliziert die Infragestellung des Erklärungspotenzials empirischer Messungen und ihren Ergebnissen. Mit einem Fokus auf Empirie und die hier häufig generierten numerischen Ergebnisse, so die Annahme, würden die Gegenstände kunstwissenschaftlicher Forschung nicht mehr vor dem Horizont ihrer sozialen, kulturellen und historischen Einbettung erklärt, sondern lediglich gemessen werden. Die Kunstwissenschaft würde sich demnach mit der Zuwendung zu empirischen Ansätzen selbst in ihrer Relevanz einschränken.<sup>98</sup>

---

<sup>96</sup> Kontext des Kurzzitates: „Man müsste dieser Selbstaufgabe geisteswissenschaftlicher Kritik keine größere Beachtung schenken, wenn sie nicht einem aktuellen Trend zur Enthistorisierung und Essentialisierung der Kunstgeschichte entsprechen würde. [...]“ (Geimer 2016: 9; siehe auch S. 30).

<sup>97</sup> Vgl. auch Anderson 2009, und N. Katherine Hayles, die – Timothy Lenoir zitierend – dessen provokante Forderung wiederholt: „forget meaning and follow the data streams“ (Hayles 2012: 33).

<sup>98</sup> Vgl. auch Elton Barker und Kollegen, die fordern: “as humanists, we need technologies that allow open-endedness, that preserve the ambiguity and nuance of our work, rather than render the data flat and colourless.” (Barker et al. 2012: 190).

---

### 1.5 Zusammenschau der positiven und negativen Kritik

Legt man allein die Positionen der kritischen Stimmen zugrunde, entsteht leicht der Eindruck, empirische Zugänge zu den Kunstwissenschaften seien nicht nur weitestgehend unreflektiert, sondern würden überdies tatsächlich „mit hohem finanziellen und apparativen Aufwand“ lediglich irrelevanten oder im schlimmsten Fall sogar „gefährlichen“ Output produzieren. Für die Prüfung theoretischer Positionen wären empirische Zugänge zur Kunstwissenschaft dementsprechend nicht als geeignet zu verstehen. In der Zusammenschau der positiven und negativen Kritik ergibt sich jedoch ein differenzierteres Bild. Zum einen wird deutlich, dass bestimmte Aspekte der negativen Kritik auf einem Missverständnis hinsichtlich des Selbstverständnisses und der Arbeitspraxis empirisch-arbeitender Forschender beruhen. Zum anderen zeigt sich, dass die Argumente der positiven und der negativen Kritik weitestgehend ohne wechselseitigen Bezug bestehen. Weder positive noch negative Stimmen berufen sich in ihrer Argumentation *direkt* auf Aspekte, die im jeweils anderen „Lager“ betont werden. Eine Anschlussfähigkeit ist allenfalls im indirekten Bezug der Argumente gegeben. Eine voreilige Beurteilung der Diskurspositionen ist daher problematisch.

Der in diesem Kontext am einfachsten zu adressierende Aspekt sind die bestehenden Missverständnisse zwischen den Positionen. Hier ist insbesondere der im vorangegangenen Kapitel zuletzt dargestellte Punkt hervorzuheben: die Frage nach dem Verhältnis von Theorie und Messung. Während in den kritischen Positionen greifbar wird, dass ein Unbehagen mit der vermeintlichen Theoriearmut empirischer Zugänge besteht (Kapitel 1.4, III), bemühen sich Befürworter\_innen empirischer Zugänge ausdrücklich darum, zu betonen, welche fundamentale Rolle Theorie in der empirischen Forschung spielt. Nicht nur wird hier wiederholt unterstrichen, dass Empirie immer einen theoretischen Startpunkt benötigt, von welchem ausgehend eine empirisch-quantitative Analyse überhaupt erst angestoßen werden kann;<sup>99</sup> auch ist man sich der Tatsache durchaus bewusst, dass die Ergebnisse empirischer

---

<sup>99</sup> Vgl. Rosenberg 2016, Kohle 2016.

---

Messungen numerischer Natur sind und somit nicht für sich selbst sprechen können (Kapitel 1.3, I).<sup>100</sup> Die zwingende Notwendigkeit, empirische Ergebnisse mit Rückbezug zur theoretischen Rahmung der Forschungsfrage zu interpretieren, sie also gewissermaßen zu „re-theoretisieren“, gehört daher zum Selbstverständnis empirischer Forschung. Dies geht schon aus den allgemeinen Definitions- und Klassifikationsmerkmalen für empirische Forschungsansätze hervor, insbesondere für quantitative Forschungsansätze.<sup>101</sup> Es liegt also offensichtlich ein Missverständnis vor, wenn Kritiker empirischen Forschungsansätzen unterstellen, dass hier der Primat auf die numerischen Ergebnisse von Messungen gelegt würde.

Weit weniger einfach zu klären, sind diejenigen Aspekte der positiven und negativen Kritik, die sich nicht in ein direktes Bezugsverhältnis stellen lassen. Indirekte Bezüge lassen sich hier zwar herstellen, allerdings nicht zwischen den jeweiligen spezifischen Aspekten, sondern nur zwischen der ihnen zuordenbaren groben Überkategorie: den Möglichkeiten empirischer Zugänge zum Feld der Kunstwissenschaften. Als spezifische Aspekte lassen sich hier seitens der *positiven* Kritik vor allem hervorheben: die Überwindung von subjektiven Forschungsperspektiven (Kapitel 1.3, I) und die Etablierung einer auf regelmäßige Strukturen bedachten Herangehensweise (Kapitel 1.3, III). Beides zusammen führt zu einer Entkanonisierung der Forschung, da sich der Blick der Forschenden zwangsweise nicht mehr auf einzelne Meisterwerke oder Case Studies richtet, sondern Strukturen in den Blick nimmt, welche zwar im einzelnen Objekte realisiert sind, aber dennoch über dieses hinausgehen. Der Fokus auf Strukturen gestattet eine breitflächige Prüfung theoretischer Annahmen, die an Einzelfallstudien entwickelt wurden und das Aufzeigen noch bestehender Forschungslücken. Ein Fokus auf Strukturen trägt außerdem dazu bei, das „Big Picture“ bei konkreten Fragestellungen hervorzuheben, und hat somit sogar das Potenzial Theorie anzureichern und weiterzuentwickeln. Dies ist insbesondere für Themenbereiche relevant, deren Grundannahmen noch nicht gesichert sind. Die Chancen empirischer Zugänge zum Fach

---

<sup>100</sup> Ebd.

<sup>101</sup> Siehe Kapitel 2.2

---

betreffen also hauptsächlich die Erweiterung des kunstwissenschaftlichen Blickwinkels und die Sicherung praxisrelevanter Erkenntnisse.

Die negative Kritik an den Möglichkeiten empirischer Zugänge zu den Kunstwissenschaften richtet sich nun *nicht* gegen die von der positiven Kritik hervorgehobenen praktischen Vorzüge. Stattdessen zielen die hier zum Tragen kommenden Argumente auf die *Prämissen* der positiven Kritik ab. Die Überwindung von subjektiven Forscherperspektiven durch Empirie wird als unmöglich empfunden, da empirische Zugänge zwingend einen reduktionistischen Ansatz verfolgen (Kapitel 1.4, I). Hierzu wird die Komplexität holistischer, geisteswissenschaftlicher Konzepte betont, wohingegen empirische Untersuchungen in der Regel das Gesamtkonzept auf vereinzelte, testbare Parameter reduzieren.<sup>102</sup> Den kritischen Stimmen zufolge wird in empirischen Untersuchungen demnach nicht das eigentlich interessierende Konzept gemessen, sondern etwas Anderes. Dieses Andere ist bestenfalls eine Abstraktion der ursprünglichen Idee, schlimmstenfalls aber ein unvollständiges Bruchstück des Ganzen, welches in seiner ausgelösten Form nicht für sich sprechen kann. Es besitzt somit strenggenommen überhaupt keine Aussagekraft. Die Möglichkeit, durch empirische Untersuchungen praxisrelevante Erkenntnisse abzusichern, sehen kritische Positionen demnach nicht als gegeben. Diesem allgemeinen Kritikpunkt lässt sich auch das spezifische Argument der negativen Kritik zuordnen, dass empirische Arbeiten mit zeitgenössischen Betrachtenden notwendigerweise ahistorisch sind und somit eigentlich keine Rückschlüsse über *kultur- und kunsthistorische* Fragestellungen erlauben (Kapitel 1.4, II.b). Auch in diesem sehr spezifischen Fall wären die Ergebnisse einer empirischen Untersuchung nicht als praxisrelevante Prüfung bestehender theoretischer Annahmen zu verstehen.

Ein weiterer wesentlicher Aspekt der negativen Kritik, ist die Frage nach der Relevanz der Ergebnisse, die mittels eines auf Regelmäßigkeiten bedachten Blicks erzeugt werden können

---

<sup>102</sup> Siehe auch Kohle 2018: 14.

---

(Kapitel 1.4, II.a). In den Positionen der negativen Kritik wird sehr deutlich, dass die empirische Prüfung von Thesen, welche allgemeine Strukturen betreffen, als redundant erachtet wird. Das wird überaus deutlich in Peter Geimers Worten, der dazu schreibt: Eine empirische Untersuchung zeigt lediglich auf „was man auch vorher schon wusste“<sup>103</sup>. Die von der positiven Kritik so prominent betonte Möglichkeit empirischer Zugänge, die Blickwinkel der kunstwissenschaftlichen Forschung auszuweiten, indem eine stärkere auf Strukturen bedachte Analyse erfolgt, wird von der negativen Kritik somit ebenfalls zurückgewiesen. Eine solche Ausweitung der Blickwinkel ist den Positionen der negativen Kritik zufolge nicht notwendig, da relevante Strukturen durch das bestehende Methodenrepertoire hinreichend präzise antizipiert werden können. Auch hier liegt das eigentliche Problem in der unterschiedlichen Bewertung der zugrundeliegenden Prämissen. Anders als bei der Frage nach der Überwindbarkeit von subjektiven Forschungsperspektiven sind es hier die Positionen der *positiven* Kritik, die die Prämissen der Argumentation der Gegenseite in Frage stellen. Eine Antizipation von generellen Strukturen, ausgehend von einer subjektiven Forschungsperspektive oder von einzelnen Vergleichsbeispielen, ist von einem empirisch-analytischen erkenntnistheoretischen Standpunkt aus betrachtet nicht zulässig. Ganz allgemein werden all jene Generalisierungen als problematisch betrachtet, deren Verankerung in allgemeinen Strukturen nicht nachgewiesen ist. Ein solches Vorgehen wird als Quelle vieler Fehlschlüsse angenommen und ist deswegen grundsätzlich abzulehnen. Die unterschiedlichen Haltungen der positiven und negativen Kritik zu diesem Thema können daher im Wesentlichen als Ausdruck unterschiedlicher erkenntnistheoretischer Standpunkte verstanden werden.<sup>104</sup>

---

<sup>103</sup> Einbettung des Zitats in Geimers Text: „[...]an ihren Resultaten misst, zeigt sich in der Regel, dass hier entweder mit hohem finanziellen und apparativen Aufwand experimentell ermittelt wurde, was man auch vorher schon wusste (etwa dass Experten Gemälde anders betrachten als Laien) [...]“ (Geimer 2016: 9).

<sup>104</sup> Diese Problematik wurde auch an anderer Stelle explizit als sich unterscheidende „epistemic cultures“ beschrieben. (Knorr-Cetina 1999).



---

Die Tatsache, dass die Ziele und Möglichkeiten empirischer Zugänge zu den Kunstwissenschaften von der positiven und negativen Kritik auf der Basis jeweils anderer Prämissen bewertet werden, zeigt auf, wie komplex die Frage nach der Sinnhaftigkeit einer empirisch-arbeitenden Kunstwissenschaft ist. Die eigentlichen Kernaspekte des Diskurses, nämlich die zugrunde gelegten Prämissen, werden durch die jeweiligen Positionen nur einseitig adressiert. Fragen nach den Chancen einer Konsolidierung der Positionen stellen sich somit gar nicht erst. Der Diskurs läuft aneinander vorbei.

Festhalten lässt sich allerdings, dass es Fragen nach den praktischen Möglichkeiten empirischer Forschung sind, die sowohl die Argumentationslinien der positiven als auch der negativen Kritik charakterisieren. Hierbei zielt die *positive Kritik* tendenziell eher auf die *inhaltlichen Möglichkeiten* ab, die sich durch die Anwendung von empirischen Ansätzen für die Kunstwissenschaft ergeben. Die *negative Kritik* nimmt hingegen hauptsächlich die *methodischen Möglichkeiten* empirischer Forschung in den Blick. Es lässt sich daher ebenfalls festhalten, dass nach einer ersten Gegenüberstellung der Positionen offenbleibt, ob und inwiefern eine empirische Untersuchung in der Lage ist, kunstwissenschaftliche Konzepte zu adressieren. Die von der negativen Kritik herausgestrichene Problematik der reduktionistischen Perspektive ist zunächst nicht von der Hand zu weisen. Das holistische Denken geisteswissenschaftlicher Untersuchungen, dessen Grundsatz die *Nicht-Reduzierbarkeit* des Ganzen auf die Summe seiner Teile ist, wird in empirischen Untersuchungen häufig nicht sonderlich ernst genommen. Aus forschungspraktischen Gründen greift hier oft ein pragmatischer Ansatz, dessen Maxime zunächst die Reduktion von Komplexität ist. Durch das allmähliche Testen von Hypothesen setzt sich so, quasi additiv, nach und nach ein Gesamtbild zusammen. Es kollidieren hier also zwei Welt- und Wissenschaftsbilder, die zunächst unvereinbar scheinen.<sup>105</sup> Ähnliches gilt für die Frage nach dem auf Strukturen ausgerichteten Blick. Ob und inwiefern der Schluss auf generalisierbare

---

<sup>105</sup> Vgl. auch Knorr-Cetina 1999.

---

Konzepte ohne empirische Prüfung im Kontext der Kunstwissenschaften möglich ist, bleibt zunächst offen.

Vor dem Hintergrund dieser Problematik ist es interessant, dass der Diskurs häufig hauptsächlich abstrakt-theoretischer Natur bleibt. Vergleichsbeispiele für die angeführten Thesen werden allenfalls genannt, nicht aber im Detail diskutiert. Dies gilt insbesondere für die negative Kritik. Inwiefern empirische Untersuchungen in der Kunstwissenschaft also *de facto* reduktionistisch sind, inwiefern sie ahistorisch sind, und ob die hier gewonnenen Erkenntnisse auch ohne einen empirischen, auf Strukturen abzielenden Zugang erzielt hätten werden können, bleibt demnach unklar. In letzter Konsequenz sind dies Fragen nach der Praxis empirischer Zugänge zu kunstwissenschaftlichen Fragestellungen, Fragen nach den *tatsächlichen* Vorgehensweisen und den *konkreten* Interpretationen der Ergebnisse. Ohne Kenntnisse über diese ganz konkreten Praxen empirischer Forschung bleiben sowohl die Argumente der positiven, als auch die Argumente der negativen Kritik bloße Spekulation.

Für die weitere Analyse der – durch positive und negative Kritik – aufgeworfenen Fragen ist somit der Bezug auf konkrete Fallbeispiele sinnvoll und notwendig. Im Folgenden werde ich daher drei empirische Studien zur Rezeptionsästhetik im Detail vorstellen. Die Domäne der Rezeptionsästhetik ist, wie zuvor dargelegt, insofern zweckmäßig und sinnvoll als hier einerseits besonders viele empirische Zugänge in der Forschungspraxis zum Tragen kommen und andererseits, weil die hier postulierten Annahmen zur menschlichen Wahrnehmung und Erfahrung immer auch Thesen über eine tatsächliche, realweltliche Praxis darstellen. Die publizierten Journalartikel zu diesen Studien sind dem Fließtext des folgenden Kapitels in ihrem ursprünglichen Publikationsformat in englischer Sprache beigelegt. Das Beibehalten des ursprünglichen Publikationsformates hat den Vorteil, dass auch die Form der Kommunikation offengelegt wird, also der durch die Publikation realisierte Diskursbeitrag zum Thema. Die Diskussion wird die theoretische Einbettung sowie relevanten Thesen und Argumente aufzeigen und für alle drei Fälle darlegen, inwiefern durch den gewählten empirischen Zugang ein Mehrwert für die gesetzten Fragestellungen nachweisbar ist. Eine

---

Methodendiskussion, welche die Einschränkungen der in den Studien generierten Ergebnisse aufzeigt und somit auch die Beleuchtung der negativen Kritik, folgt in einem gesonderten Kapitel (Kapitel 3). Auf der Basis von Kapitel 2 und 3 werden in der abschließenden Diskussion schließlich die tatsächlichen Vor- und Nachteile der empirischen Herangehensweise(n) gegenübergestellt und einer finalen Bewertung unterzogen.

---

*Kapitel* **2**

**Studien**

---

Die *positive* Kritik zielt hauptsächlich auf die Chancen empirischer Methoden ab, einen inhaltlichen Beitrag zu leisten; die negative Kritik hingegen stellt die Realisierbarkeit eines solchen Beitrages qua Methode in Frage. Im Folgenden werde ich beide Seiten anhand dreier empirischer Studien am konkreten Beispiel nachvollziehen.

Die Studien, die ich zu diesem Zweck besprechen werde, sind zwar, wie bereits erwähnt, allesamt der Rezeptionsästhetik zuzuordnen, behandeln aber innerhalb dieser unterschiedliche Themen. Die erste beschäftigt sich mit dem Konzept der Komplexität in der Malerei (Kapitel 2.1). Im Kreise der drei diskutierten Studien ist diese erste Studie die umfangreichste und vielschichtigste, weswegen ihre Vorstellung und Diskussion im Folgenden den größten Raum einnimmt. Die beiden weiteren Studien haben sich mit Fragen zur Betrachtung von, und Auseinandersetzung mit, abstrakter Kunst beschäftigt. Hier ging es zum einen um die von Werner Haftmann geäußerte These, abstrakte Kunst sei als eine Art Weltsprache zu verstehen (Kapitel 2.2); und zum anderen um die Frage, ob, und wenn ja wie, die, z. B., in Jackson Pollocks Arbeiten Ausdruck findende Dynamik sich in den Prozess der Auseinandersetzung mit den Kunstwerken einschreibt (Kapitel 2.3). In den drei Studien wurden unterschiedliche empirische Methoden verwendet – sie bilden somit einen guten Ausgangspunkt für eine breite Methodendiskussion. Darüber hinaus verfolgen die Studien, im Abstrakten gesprochen, unterschiedliche Zwecke: Während es das Ziel der ersten Studie ist, ein Konzept zu definieren, und sie daher im weiteren Sinne als theoriegebend zu verstehen ist, stellen die Studien zwei und drei Beispiele dar, die bereits bestehende Thesen aufgreifen und testen. Hier handelt es sich also nicht um theoriegebende, sondern um theorie-testende Ansätze. Die publizierten Journalartikel zu diesen Studien sind, wie zuvor angekündigt, dem nachfolgenden Fließtext in ihrem ursprünglichen Publikationsformat in englischer Sprache beigelegt.

Die Diskussion der in Kapitel 1 herausgearbeiteten Chancen und Probleme empirischer Forschung erfolgt in zwei gesonderten Abschnitten. Ziel *des ersten* Abschnittes ist es, für jede Studie konkret aufzuzeigen, welcher spezifische Nutzen und Erkenntnisgewinn jeweils

---

verzeichnet werden konnte und inwiefern dieser durch ihren empirischen Zugang bedingt ist. Ziel des *zweiten* Abschnittes ist es, methodische Probleme und Einschränkungen der vorgestellten Studien offenzulegen.

Die Trennung der Diskussion in einen inhaltlichen und einen methodischen Block ist dabei nicht unproblematisch. Nicht nur ist, wie in Kapitel 1.2 und 1.3 beschrieben, die Anwendung von empirischen Ansätzen notwendigerweise inhaltlich motiviert und gerahmt, auch lassen sich bestimmte Ergebnisse inhaltlich nur mit Blick auf die angewandte Methode fassen. Die schwerpunktmäßige Eingrenzung der beiden folgenden Diskussionsabschnitte scheint mir – gegeben der unterschiedlichen Fokusse von positiver und negativer Kritik – dennoch sinnvoll. Trotzdem kann die folgende Diskussion nicht völlig trennscharf in einen inhaltlichen und einen methodischen Abschnitt geteilt werden. Beide Diskussionsabschnitte werden immer wieder Bezüge zu entweder inhaltlichen oder methodischen Aspekten aufweisen, auch wenn diese im jeweiligen Abschnitt nicht schwerpunktmäßig behandelt werden. Diese „Grenzüberschreitungen“ zwischen den beiden Diskussionsräumen sind dem Anliegen dabei jedoch nicht abträglich. Ganz im Gegenteil, hier realisieren sich im Konkreten Brücken zwischen der positiven und der negativen Kritik, die einen wesentlichen Ansatzpunkt für die Konsolidierung der beiden Argumentationslinien bieten.

---

# Studie

## More Than the Sum of Its Parts. Perceiving Complexity in Painting

LITERATURANGABE

---

Laura Commare/ Raphael Rosenberg/ Helmut Leder, More Than the Sum of Its Parts.  
Perceiving Complexity in Painting, in: *Psychology of Aesthetics, Creativity and the Arts*,  
2018, Advance online publication: [10.1037/aca0000186](https://doi.org/10.1037/aca0000186).

---

## 2.1 *More Than the Sum of Its Parts. Perceiving Complexity in Painting*

Die hier besprochene Studie zielt auf die Definition des Konzeptes Komplexität ab. Sie stellt einen Brückenschlag zwischen psychologischer Ästhetik-Forschung und kunsttheoretischen Positionen dar. Das Grundanliegen der Untersuchung war es, das Komplexitätskonzept in einer für psychologische und kunsttheoretische Theoriepositionen anschlussfähigen Form zu definieren. Der Text mit dem Titel *More Than the Sum of Its Parts. Perceiving Complexity in Painting* ist in Zusammenarbeit mit Raphael Rosenberg und Helmut Leder entstanden und beim Journal *Psychology of Aesthetic, Creativity and the Arts* (PACA) publiziert.<sup>106</sup> Da es sich bei PACA um eine psychologische Zeitschrift handelt, ist die Diskussion im nachfolgend-abgedruckten Originaltext stark an einer psychologischen Perspektive auf das Problem der Komplexität in der Kunst orientiert. Eine kunstwissenschaftliche Einbettung hingegen fehlt dem Aufsatz. Um auch die kunstwissenschaftliche Relevanz der Fragestellung deutlich zu machen, sowie die Notwendigkeit aufzuzeigen, dem gesetzten Problem mit empirischen Mitteln zu begegnen, ist dem eigentlichen Studientext eine etwas ausführlichere Einleitung vorangestellt. Ein der Studie nachgereihtes Kapitel erläutert und bespricht die im Originaltext angeführten statistischen Ergebnisse und diskutiert anschließend, inwiefern diese empirischen Ergebnisse für die kunstwissenschaftliche Fragestellung relevant sind und inwiefern die empirische Herangehensweise zur Erlangung dieser Ergebnisse zwingend erforderlich war, bzw. inwiefern durch die empirische Herangehensweise Vorteile realisiert wurden, die anderweitig nicht hätten erzielt werden können.

---

<sup>106</sup> Die Mehrfachautor\_innenschaft des Textes ist durch das arbeitsteilige Vorgehen in der Vorbereitung, Durchführung und Nachbereitung der empirischeren Untersuchung bedingt. Die unterschiedlichen Arbeitsschritte sind dabei schwerpunktmäßig von jeweils bestimmten Teilen des Teams bearbeitet worden. Eine ausführliche Aufstellung der jeweiligen Aufgaben und Einzelleistungen findet sich im Anhang (Anhang *Stellungnahme zur Arbeitssteilung I*).



---

## Einleitung

Die kunsthistorische Praxis greift zur Beschreibung, Klassifikation und Bewertung ihrer Untersuchungsgegenstände auf eine Vielzahl von Begriffen, Konzepten und Theoriewerkzeugen zurück. Einer dieser Begriffe ist *Komplexität*. Es ist ein vielfältiger Begriff, der immer wieder und in einer ganzen Reihe von unterschiedlichen Untersuchungskontexten Anwendung findet. Es handelt sich um einen intuitiven Begriff, einen Begriff, den wohl beinahe jeder, der über Kunst schreibt, schon einmal verwendet hat. Der Begriff impliziert eine „Bipolarität“, die die Annahme eines linearen Raumes erlaubt, das heißt eines Raumes, der sich geradlinig vom Einfachen zum Komplexen erstreckt. Gerade zum Zweck der Differenzierung von Kunstwerken, Stilen etc. ist Komplexität also ein dankbarer Begriff. In diesem Sinne bezeichnet beispielsweise Hofmann das künstlerische Profil der Brücke, in Bezug auf farbgestalterische sowie auf graphische Darstellungsmittel, als "komplexer als das der Fauves".<sup>107</sup> In ähnlicher Funktion, das heißt zu Unterscheidungszwecken, verwenden auch Panofsky, Gombrich und Locher den Begriff. Panofsky argumentiert die Güte von Dürers *Melancholia I* sei von Kopisten niemals erreicht worden und beschreibt diesen Eindruck mit den folgenden Worten: "Just as Dürer's Melencolia I was copied by numerous subsequent Artists but grasped in its full complexity by none, so the engraving B.79 was appropriated by sixteenth century, but with only fractional understanding".<sup>108</sup> Gombrich spricht in *Norm and Form* von einer auftretenden, unklassischen Komplexität in der Formensprache der Renaissance<sup>109</sup> und auch bei Locher zeigt sich eine beschreibend-unterscheidende Verwendung des Begriffs, der sich hier ebenfalls, wie in den vorangegangenen Beispielen, hauptsächlich auf Formaspekte bezieht.<sup>110</sup> Der Begriff der Komplexität wird allerdings keineswegs ausschließlich in der Auseinandersetzung mit Formaspekten herangezogen, sondern gleichermaßen auch in Bezug auf semantische Strukturen im Bild angewendet. Im Dahl beispielsweise schreibt über die

---

<sup>107</sup> Hofmann 1987: 266.

<sup>108</sup> Panofsky 1983: 308.

<sup>109</sup> Gombrich 1966: 96.

<sup>110</sup> Z. B. Locher 2001: 226.

---

szenische Organisation von Giotto's Fresken in der Arenakapelle: "[...] überliegt zugleich die szenische, semantische Komplexität."<sup>111</sup> Ähnliches findet sich auch bei Pächt, der in Bezug auf die szenische Entfaltung bei Michael Pacher von einer inhaltlichen Komplexität spricht.<sup>112</sup>

Bereits anhand dieser wenigen Beispiele, lassen sich recht deutlich zwei wesentliche Punkte aufzeigen, die die Verwendung des Begriffs rahmen. Erstens zeigt sich, dass es sich nicht um eine lediglich umgangssprachliche Beschreibung handelt. Die Verwendung des Begriffes in der Literatur legt das etwa nahe, wenn von semantischer Komplexität die Rede ist, oder von Komplexität, die nicht von jedermann erfasst und reproduziert werden kann. Beispiele, die illustrieren, dass sich hinter dem Wort ein Begriff verbirgt, welcher auf mehr referiert als schlicht die objektive Beschreibung von Form oder Inhalt. An diesen Beispielen wird deutlich, dass Wahrnehmung, kognitive Verarbeitung und Verstehen für Komplexitätsurteile eine Rolle spielen. Sie zeigen, dass der Begriff Komplexität eine Bewertung ausdrückt. Zweitens lässt sich anhand der Verwendung des Begriffs aufzeigen, dass Komplexität in der Regel als positiv konnotiertes Merkmal oder gar als Gütekriterium verwendet wird. Das zeigt sich etwa, wenn Panofsky die Qualität von Dürers *Melancholia I* an der im Kupferstich angelegten Komplexität festmacht. Eine Begriffsverwendung, die eine „starke“ subjektive Perspektive eröffnet.

Wie anhand dieser ausgewählten Verwendungen des Begriffes deutlich wird, verweist der Ausdruck *Komplexität* auf ein nicht-triviales theoretisches Konstrukt. Damit verlangt dessen Verwendung notwendig eine Definition und Erläuterung. Dennoch bleibt das Konzept Komplexität, trotz seiner Allgegenwärtigkeit im wissenschaftlichen Diskurs zumeist unterdefiniert – eine Blackbox, die zahlreiche Fragen aufwirft. Nicht nur bleibt unklar, warum und wie genau bestimmte Aspekte im Bild, seien sie nun formaler oder semantischer Natur, Komplexität bedingen. Unbeantwortet bleibt auch die Frage, welche Konsequenzen

---

<sup>111</sup> Imdahl 1980: 94.

<sup>112</sup> Pächt 1977: 105.

---

der Umstand größerer oder kleinerer Komplexität für die Kunst und ihre Betrachter\_innen tatsächlich hat. Insbesondere die positiv-bewertende Verwendung des Begriffs, die ebenfalls ohne Auflösung oder Motivation verbleibt, wird so zur subjektiv-suggestiven Form.

Um eine qualifizierte Verwendung des Begriffes als Arbeitswerkzeug in den Kunstwissenschaften zu gewährleisten, muss demnach eine Untersuchung erfolgen. Fragen nach der Begriffsbedeutung, dessen Konstituenten und der Objektivierbarkeit des Begriffes müssen eindeutig geklärt werden. Auch muss offengelegt werden, welche Rückschlüsse das Prädikat ‚komplex‘ erlaubt oder erfordert; das heißt die Frage nach dem Maß, in welchem Komplexität tatsächlich als Eigenschaft von Kunstwerken selbst verstanden werden darf und in welchem Maß die subjektive Perspektive der Betrachtenden zum Entstehen eines Komplexitätseindrucks beiträgt, bedarf der Klärung.

Trotz dieser offenen Fragen, ist das Konzept *Komplexität* bisher nicht Gegenstand der kunstwissenschaftlichen Forschung gewesen. Anders in der psychologischen Ästhetikforschung – hier finden sich zahlreiche Publikationen zum Thema. Da aus der kunstwissenschaftlichen Praxis deutlich wird, dass Aspekte der menschlichen Wahrnehmung und kognitiven Verarbeitung notwendig Teil einer Antwort auf die Frage nach dem Wesen, dem Logos der Komplexität in der Kunst sein müssen, sind die im Rahmen psychologischer Untersuchungen generierten Ergebnisse durchaus relevant für die Erarbeitung eines fundierten Komplexitätskonzeptes.

Den verschiedenen psychologischen Erklärungsansätzen zur ästhetischen Wahrnehmung ist die Annahme gemein, dass die visuelle Wahrnehmung zwei Momente hat: einerseits ist sie reizgeleitet („bottom-up“), andererseits kognitionsgeleitet („top-down“). Während sich reizgeleitete Vorgänge bei der Wahrnehmung auf bestimmte Merkmale des dargebotenen Reizes beziehen, wie etwa Farbwerte, Helligkeit oder Kontrastkanten, ist mit kognitionsgeleiteten Einflüssen gemeint, dass gewissermaßen „geistige“ Prozesse, wie

---

subjektive Erinnerungen, Assoziationsvermögen oder schlicht die Reflektion des Gesehenen, den Vorgang der Betrachtung beeinflussen können.<sup>113</sup>

In der psychologischen Literatur findet eine Auseinandersetzung mit Komplexität meistens nur auf der „objektiven“, das heißt formalen, Ebene statt. Seit der Publikation von Daniel Berlynes einflussreichen Arbeiten zur ästhetischen Erfahrung wird Komplexität häufig allein über die Anzahl der im Stimulus präsenten Elemente, der Art ihrer Anordnung (Symmetrie, Rhythmus) sowie ihrer Heterogenität definiert.<sup>114</sup> Allen diesen Untersuchungen ist gemein, dass ihr Kernanliegen die ästhetischen Präferenzurteile sind, und nicht Komplexitätswahrnehmung selbst. Gemeinhin konnte gezeigt werden, dass eine mittlere Komplexität die positivsten Bewertungen hervorruft, da sie die Studienteilnehmer\_innen weder unter- noch überfordert und somit optimal stimuliert. Die in den meisten Versuchen verwendeten Bildbeispiele bestehen allerdings ausschließlich aus inhaltsleeren oder zumindest überwiegend bedeutungslosen Formenzusammenstellungen.

Einen der ersten Versuche, Komplexität an echten Kunstwerken zu untersuchen, hat Wohlwill unternommen.<sup>115</sup> Wohlwill zeigte in seinen Studien nicht mehr nur inhaltsleeres Formenmaterial, sondern Kunstwerke, weswegen erstmalig auch Bildeigenschaften wie Farbigkeit oder Textur untersucht wurden. In Wohlwills Folge wurden eine ganze Reihe von Versuchen mit echten oder leicht modifizierten Kunstwerken durchgeführt.<sup>116</sup> Auch hier war das Kerninteresse jedoch nicht die Komplexitätswahrnehmung selbst, sondern die Frage nach den Determinanten der ästhetischen Präferenzbildung. Bestrebungen, eine allgemeine Definition des Komplexitätsbegriffs zu finden, blieben jedoch aus, sodass seitdem eine große

---

<sup>113</sup> Für eine Übersicht siehe Zimbardo 1995.

<sup>114</sup> Z. B. Berlyne 1963, Munsinger/Kessen 1964, Day 1967, Eisenman 1967, Berlyne et al. 1968, Eisenman/Gillens 1968, Berlyne 1970, Nicki 1972, Kreidler et al. 1974.

<sup>115</sup> Wohlwill 1968.

<sup>116</sup> Z. B. Osborne/Farley 1970, Nicki/Moss 1975, Saklofske 1975, Farley/Weinstock 1980, Nicki/Lee/Moss 1981, Krupinski/Locher 1988, Neperud/Marschalek 1988, Messinger 1998.

---

Anzahl unterschiedlicher Konzepte geprägt wurden, welche bis heute nebeneinander bestehen.

Ein Versuch, die verschiedenen Komplexitätskonzepte zu kondensieren, findet sich erst 2010 in den Arbeiten von Marcos Nadal und Kollegen.<sup>117</sup> Auf der Basis einer umfangreichen empirischen Untersuchung konnten die Autor\_innen zeigen, dass die Komplexitätswahrnehmung drei Hauptkomponenten unterliegt: a) der Anzahl der im Werk vorhanden Elemente, b) der Übersichtlichkeit ihrer Anordnung und c) der Symmetrie. Diese Untersuchung, wie auch die bisher referierten, adressiert jedoch ausschließlich objektive, also im Bild angelegte, Faktoren.

Dass Komplexitätswahrnehmung oder -erfahrung jedoch nicht allein vom betrachteten Objekt, sondern auch vom betrachtenden Subjekt abhängig ist, konnte in einigen weiteren psychologischen Studien ebenfalls gezeigt werden. So konnten etwa Pablo Tinio und Helmut Leder zeigen, dass die Vertrautheit („familiarity“) mit einem bestimmten Objekt die wahrgenommene Komplexität reduziert.<sup>118</sup> Des Weiteren konnten auch Einflüsse auf die Bewertung von Komplexität durch Wissens- oder Erfahrungswerte nachgewiesen werden.<sup>119</sup> Der Komplexitätseindruck, den Betrachtende von einem Kunstwerk gewinnen, ist demnach sowohl durch formal-visuelle Faktoren geprägt, als auch Abhängig von den Voraussetzungen und Bedingungen der betrachtenden Personen.

Studien zur *inhaltlichen* Komplexität von Kunstwerken sind im Vergleich zu Arbeiten, deren Schwerpunkt auf der Analyse von formal-visuellen Aspekten liegt, kaum unternommen worden. Tatsächlich finden sich hierzu nur einige wenige Beispiele, die zudem meist den Begriff *Inhalt* unterschiedlich auffassen, weswegen die bisher erlangten Ergebnissen kein klares Bild zeichnen.<sup>120</sup> Es lassen sich jedoch die zwei wesentliche Erkenntnisse festhalten:

---

<sup>117</sup> Nadal 2007, Nadal et al. 2010.

<sup>118</sup> Tinio/Leder 2009.

<sup>119</sup> Thompson 1973.

<sup>120</sup> Vitz 1966, Aitken 1974, Nicki/Moss 1975, Nicki/Lee/Moss 1981, Jakesch/Leder 2016.

---

einerseits dass Abweichungen zwischen Personen in der Beurteilung der Komplexität bestimmter Werke zumindest zum Teil dadurch erklärt werden können, dass die inhaltliche Verarbeitung, und damit die Bildsemantik, eine *indirekte* Rolle spielt;<sup>121</sup> andererseits dass semantische Strukturen das Komplexitätserleben tatsächlich auch *direkt* beeinflussen. Martina Jakesch und Helmut Leder konnten zum Beispiel zeigen, dass Ambiguität, ähnlich wie auch Familiarität, das Erleben von Komplexität beeinflusst.<sup>122</sup>

Eine Untersuchung, die objektive und subjektive Faktoren gleichermaßen in vollem Umfang berücksichtigt, die Ergebnisse zu formal-visueller Komplexität und semantischer Komplexität integriert und die nach wechselseitigen Verschränkungen der Einflüsse fragt, ist bisher im Rahmen der psychologischen Ästhetikforschung jedoch nicht erfolgt.

Solch ein integriertes Komplexitätsmodell für Kunstwerke ist aus kunsttheoretischer Sicht anderen Theorieparametern unterworfen als in der psychologischen Ästhetikforschung. Diese sind jedoch bei genauerer Betrachtung miteinander vereinbar. Werfen wir, bevor wir die Frage nach der grundsätzlichen Überführbarkeit der beiden Theoriekontexte stellen, zunächst noch einen kurzen Blick auf die Etymologie des Begriffs Komplexität. Diese Rückbesinnung auf die „eigentliche“ Wortbedeutung ist hier insofern hilfreich, als der Begriff Komplexität, wie zuvor dargestellt, im Kontext kunstwissenschaftlicher Forschung hauptsächlich intuitiv-beschreibend verwendet wird. Die eigentliche Wortbedeutung des Terminus ist daher für die Rahmung jeder diesbezüglichen theoretischen Überlegung als zentral zu verstehen.

Dem Duden ist zu entnehmen, dass ‚komplex‘, sich von dem lateinischen Wort *complexum* ableitet. *Complexum* ist das Partizip Perfekt von *complecti*, das wiederum ein Kompositum aus der Präposition *cum* und *plectere*, also ‚mit‘ und ‚flechten‘ bzw. ‚ineinanderfügen‘, darstellt. Übersetzen lässt sich *complecti* mit ‚umschlingen‘, ‚umfassen‘ oder auch

---

<sup>121</sup> Vitz 1966, Aitken 1974.

<sup>122</sup> Jakesch/Leder 2016.

---

„zusammenfassen“.<sup>123</sup> Hinsichtlich seiner lexikalischen Bedeutung ist der Begriff Komplexität also mit Konzepten wie Verflochtenheit, Vielschichtigkeit oder Beziehungsreichtum verbunden. Er verweist auf reiche, vielfältige, multidimensionale und eng zusammenhängende Systeme oder Sachverhalte. Der Eindruck von Komplexität entsteht dann, wenn das Ganze mehr ist als die schlichte Summe seiner Teile, wenn die Gesamtheit nicht reduziert werden kann auf einzelne Komponenten die sukzessive abgehandelt werden, wenn durch Verflechtung die Gleichzeitigkeit vieler Aspekte wesentlich wird.<sup>124</sup>

Die bisherigen Ergebnisse der psychologischen Ästhetik greifen diese etymologischen Merkmale insofern sinnvoll auf, als sie sich etwa in den von Nadal identifizierten „formal-visuellen Haupteinflusskomponenten“ der Komplexitätswahrnehmung widerspiegeln. Für eine aus kunsttheoretischer Sicht anschlussfähige Definition sind die Ergebnisse dennoch nicht zufriedenstellend, denn aus dem Korpus der bereits vorhandenen empirischen Ergebnisse zur Komplexitätserfahrung ist die Verflechtung formal-visueller Faktoren mit semantischen Aspekten von Kunst nicht ableitbar. Dies ist nicht nur aufgrund der bereits erbrachten empirischen Hinweise auf das Vorhandensein einer solchen Verschränkung problematisch, sondern auch aufgrund der kunstwissenschaftlichen Praxis der Verwendung des Begriffes Komplexität. Dieser kann sich, wie bereits dargelegt, sowohl auf formal-visuelle als auch auf semantische Faktoren im Kunstwerk beziehen. Darüber hinaus ist die strikte Trennung von visueller Form und semantischen Inhalt aus der Perspektive kunsttheoretischer Überlegungen nicht unproblematisch.

Sollen das Komplexitätskonzept und vor allem seine Konstituenten kunsttheoretisch eingebettet werden, ist es zentral jene Merkmale festzustellen, welche aus kunsttheoretischer Perspektive für Bilder charakteristisch sind. Dazu ist es hilfreich Bilder im Sinne semiotischer Überlegungen als Objekte mit kommunikativem Potenzial, also als Zeichensysteme, zu begreifen. Eine *semiotische Perspektive* einzunehmen, ist insofern

---

<sup>123</sup> Siehe Duden ‚komplex‘, Online abrufbar unter: <http://www.duden.de/rechtschreibung/komplex>.

<sup>124</sup> Vgl. Latour/Hermant 1998: 30.

---

sinnvoll, als Bilder hier bereits als integrierte Systeme verschiedener sowohl formaler als auch inhaltlicher Ebenen verstanden werden, welche grundsätzlich in direktem Bezug zum Betrachtenden stehen.<sup>125</sup> Als Zeichen verstanden, ist ein Bild *in sich* unterteilt in Form und Bedeutung. Sinnlich-wahrnehmbar ist dabei jedoch lediglich die Form. Die hinter, oder in, der Form liegende Bedeutung zu entschlüsseln oder zuzuschreiben, obliegt den Rezipierenden und muss von diesen unter Zuhilfenahme von Wissens- und Erfahrungswerten aktiv hergestellt werden. Bildzeichen sind daher wesentlich als Dreiheit zu verstehen, als Kopplung von Form, Bedeutung und Subjekt. Diese drei Momente des Zeichens bedürfen dabei notwendig einer Gleichzeitigkeit, um ihre gemeinsame Eigenschaft, der des Symbolisierens, nicht einzubüßen. Die Semiose, also jener Prozess in welchem sich die Welt dem Menschen anhand von Zeichen vermittelt, lässt sich mit Pierce beschreiben als "[...] action, or influence, which is, or involves, a cooperation of three subjects, such as a sign, its object, and its interpretant, this tri-relative influence not being in any way resolvable into actions between pairs."<sup>126</sup>

Legt man einen semiotischen Bildzeichenbegriff zugrunde, ist evident, dass Komplexität als Konzept holistischer gedacht werden muss, als die bisherigen Studien der psychologischen Ästhetik zu formal-visueller Komplexität vorgeben. Es ist notwendig, eine Kopplung von Form, Inhalt und betrachtendem Subjekt anzunehmen und den gemeinsamen Einfluss dieser Teilelemente auf die Komplexitätserfahrung zu beschreiben. Nur mittels einer Untersuchung, die dies berücksichtigt, können letztlich die für die Kunstwissenschaften interessanten Fragen nach der Begriffsbedeutung, den Konstituenten des Konzeptes und der Objektivierbarkeit des Begriffes Komplexität praxisrelevant beantwortet werden. Eine Reduktion der Untersuchung auf die Teilelemente von Bildzeichen würde bedeuten, dass die in einem solchen Rahmen generierten Ergebnisse wesentlichen Prämissen geisteswissenschaftlicher Konzeptbildung widersprechen – und damit nur von beschränkter Relevanz für die kunstwissenschaftliche Forschung wären. Es sind jedoch nicht nur

---

<sup>125</sup> Z. B. Saussure 2001, Pierce 1998, Morris 1939, Morris 1938, Eco 1977.

<sup>126</sup> Pierce 1998: 411.



---

geisteswissenschaftliche Konzepte, welche einen holistischen Blick auf das Komplexitätskonzept nahelegen, sondern, wie aus dem kurzen Überblick zum Forschungsstand der psychologischen Ästhetik deutlich wurde, auch die Ergebnisse empirischer Untersuchungen.

### *Zusammenfassung und Bemerkungen zur Studie*

Wie bereits dargelegt, erfordert das Konzept der Komplexität eine Definition, um dessen gerechtfertigte Verwendung als Arbeitswerkzeug zu gewährleisten. Die augenscheinlichsten Fragen, die diesbezüglich gestellt werden müssen, beziehen sich auf die Klärung der Begriffsbedeutung. Hierbei sind insbesondere die Konstituenten des Konzeptes von Interesse, denn sie sind es, die den Wahrnehmungsinhalt wesentlich bestimmen. Wenn wir also die Frage „Was ist Komplexität in Kunst?“ stellen, ist damit eigentlich die Frage „Was führt dazu, dass wir ein Kunstwerk als komplex wahrnehmen?“ gemeint. Die Notwendigkeit hier auf der Basis menschlicher Wahrnehmung zu argumentieren, zieht einige Probleme nach sich: Zum einen stellt sich die Frage, ob Komplexität überhaupt als Eigenschaft von Kunstwerken verstanden werden kann oder ob, bzw. inwiefern, es sich hier im Wesentlichen um eine subjektive Bewertung der Betrachtenden handelt. Daran anschließend stellt sich auch die Frage nach der Objektivierbarkeit: Mit steigendem Einfluss von subjektiven Bewertungskomponenten, sinkt die Wahrscheinlichkeit, dass Komplexität als intersubjektiv geteiltes Konzept fassbar wird. Über die Klärung der Begriffsbedeutung hinaus müssen auch die Konsequenzen aufgezeigt werden, die durch die Erfahrung von Komplexität für die Auseinandersetzung mit einem Werk bestehen. Was bedeutet es ganz konkret, wenn ein Kunstwerk als komplex verstanden werden kann? Die alltägliche Verwendung des Begriffes zeigt bereits deutlich, dass er häufig mit positiver Konnotation, quasi als Gütekriterium, geführt wird. Damit ist aufgezeigt, dass die Wahrnehmung von Komplexität andere Bewertungen beeinflusst, die einem Kunstwerk gegenüber geäußert werden können. Die Erfahrung von Komplexität greift demnach in den Prozess der Wahrnehmung ein. In welcher Form sich dieser Eingriff im Konkreten realisiert, ist ebenfalls zu klären. Eine solche Klärung

---

ist nicht nur aus methodischen Gründen sinnvoll, da sie Forschenden Anhaltspunkte zur Hinterfragung der eigenen subjektiven Forschungsperspektive bieten kann, sondern auch aus inhaltlichen Gründen: Nur mittels einer Offenlegung der durch Komplexität induzierten Wahrnehmungskonsequenzen, lässt sich ermitteln welche Rückschlüsse das Prädikat „komplex“ tatsächlich erlaubt.

Die erste Studie adressiert nun im Kern die Frage nach der Definition von Komplexität in der Malerei. Hierzu werden – auf der Basis von psychologischen und kunsttheoretischen Positionen – Bild- und Betrachtendenmerkmale identifiziert, welche für die Wahrnehmung von Komplexität als wesentliche Determinanten begriffen werden können. Anschließend werden diese Merkmale in einer empirischen Untersuchung hinsichtlich ihres Praxisgehalts geprüft. Das Hauptaugenmerk der Untersuchung liegt auf dem Einfluss von semantischen Bildaspekten auf die Komplexitätserfahrung, sowie auf der Verschränkung von semantischen und formalen Bildstrukturen in der Wahrnehmung von Komplexität. Im Konkreten wird außerdem auch die Frage nach der Subjektivität des Komplexitätseindrucks gestellt. Teilanliegen der empirischen Analyse ist es, explizit zu machen, inwiefern Komplexitätsurteile in Bezug auf Kunstwerke als intersubjektiv-geteilte Aspekte von Werken verstanden werden können, bzw. wie groß die Schwankungsbreite des Komplexitätseindrucks zwischen den verschiedenen Personen ist.

Methodisch greift die Studie auf ein eigens für diesen Untersuchungskontext entwickeltes Fragebogeninstrument zurück, welches Urteile über das Erleben jener formalen und semantischen Bildaspekte abfragt, welche mit Komplexität in Verbindung gebracht werden können. Die Analyse der so erhobenen Daten erfolgt mittels einer Mehrebenen-Modellierung. Diese ermöglicht die Unterscheidung von auf unterschiedlichen kognitiven Verarbeitungsebenen operierenden Einflussfaktoren. Darüber hinaus ist in dieser Analyseform auch die Möglichkeit gegeben, die diversen Verschränkungen der untersuchten Merkmale offenzulegen.

---

Trotz der gewählten *empirischen* Herangehensweise an die skizzierten Probleme ist die Fragestellung der folgenden Studie – im Allgemeinen gesprochen – *kunsttheoretischer* Natur. Sie verlangt nach einer diskursiven Auseinandersetzung mit dem Wesen und den sachimmanenten Gesetzmäßigkeiten von Kunst und deren Wahrnehmung. Alle in diesem Kontext entstehenden Ergebnisse sind damit als Theoriebausteine von Kunsttheorie im weiteren Sinne zu verstehen. Im Konkreten lässt sich der Fragenraum, wie bereits aufgezeigt, rezeptionsästhetischen Überlegungen zuordnen, da hier die Wahrnehmung von Kunstwerken und der Verankerung dieser Wahrnehmung im Zentrum des Interesses steht. Der Fragenraum lässt sich darüber hinaus auch als methodologischer Problemraum fassen. Theoretische Methodenfragen sind hier insofern präsent, als die untersuchten kunsttheoretischen Konzepte Arbeitswerkzeuge in der Praxis kunstwissenschaftlicher Forschung darstellen. Diese Praxis wird durch die inhaltliche Auseinandersetzung mit dem Problemraum indirekt adressiert. Diesbezügliche Ergebnisse haben somit auch das Potenzial zur Dekonstruktion und Aktualisierung kunstwissenschaftlicher Praxis beizutragen.



© 2018 American Psychological Association  
1931-3896/18/\$12.00

tapraid5/aca-aca/aca-aca/aca99918/aca0615d18z xppws S=1 5/20/18 9:02 Art: 2016-0347

Psychology of Aesthetics, Creativity, and the Arts

<http://dx.doi.org/10.1037/aca0000186>

AQ: 1 **More Than the Sum of Its Parts: Perceiving Complexity in Painting**

AQ: au  
AQ: 2

Laura Commare, Raphael Rosenberg, and Helmut Leder  
University of Vienna

To date the complexity of an image is most often equated with its visual complexity. However, recent studies have shown that when it comes to artworks, complexity ratings are also influenced by semantic processes. The present study investigates the relationship between perceived image complexity and content-related processes in paintings by comparing ratings of perceived complexity from experts (art historians) and nonexperts. We show that the processing of the different dimensions of an artwork's content (preiconographic, iconographic, and iconological contents) all independently contribute to both groups' perceptions of an artwork's complexity. However, experts rated artworks to be less complex than nonexperts and considered different dimensions regarding the artwork's content.

*Keywords:* complexity, semantics, expertise, genre

Past research in empirical aesthetics has identified perceived image complexity as a strong predictor for liking of and interest in artworks. Most studies have equated complexity with *visual complexity* (Nadal, Munar, Marty, & Cela-Conde, 2010; Gartus & Leder, 2013) conceived as measurable visual features such as number, variety, and organization of elements (Berlyne, 1963; Nadal et al., 2010).

This conceptualization of image complexity is challenged by recent studies in psychology of the arts, where it was found that in the case of artworks, objective measures of visual complexity do not capture perceived image complexity adequately. Marin and Leder (2013) report only moderate correlations between different objective measures of visual complexity (Cavalcante et al., 2011; Chikhman, Bondarko, Danilova, Goluzina, & Shelepin, 2012; Forsythe, Nadal, Sheehy, Cela-Conde, & Sawey, 2011) and perceived image complexity. Jakesch and Leder (2015) provided a possible explanation for this discrepancy: they found a significant effect of *content ambiguity* on perceived image complexity. Artworks with unambiguous content (e.g., a bottle) were perceived as less complex than a version of them with more ambiguous content (e.g., a bottle morphed into a carrot). This is clear evidence that perceived complexity of artworks can be dependent on content-related processes. However, it remains an open question as to what *kind* of

semantic content is particularly relevant. The field of study in which semantic content was theoretically and systematically studied is art history. Thus, in the present study, we propose a semantic conceptualization of the perception of an artwork's complexity based on art historical classifications, operationalized and tested with psychological, empirical methods.

### A Semantic Conceptualization of Perceived Complexity

Content is a relevant dimension for the engagement with art (Leder, Belke, Oeberst, & Augustin, 2004; Millis, 2001; Shigeto & Nittono, 2010; Swami, 2013). To this end, Leder et al. (2004) put style and content as central dimensions in their explicit classification stage, and Augustin, Leder, Hutzler, and Carbon (2008) demonstrated that variations of content are perceived much faster than variations of style. Since it is almost impossible to consciously perceive form as meaningless (Albright & Stoner, 2002) and Jakesch and Leder (2015) have shown that content-related processes are related to complexity perception, we propose extending the concept of perceived artwork complexity (Nadal et al., 2010) to also encompass the various dimensions of an artwork's content. Beyond the relation of visual properties and perceived complexity (Nadal et al., 2010), we discuss the role of *semantic complexity*. To this end, we address the following three questions: (a) "How can we characterize the content of artworks?"; (b) "How does a perceiver process the content of artworks?"; and (c) "What consequences arise thereof for (overall) perceived image complexity?"

In principle, the term *content* refers to the topic of an artwork, which can be given as a description of the "matter" dealt with in the image (Leder et al., 2004). In some approaches, content was understood as a reference to objects and situations depicted in the artwork, such as a group of trees or a bottle (Augustin et al., 2008; Cupchik & Gebotys, 1988; Jakesch & Leder, 2015), but, content may also refer to the artwork's "meaning" which is *expressed* in the depicted situation (Russell & Milne, 1997; Swami, 2013). In

Laura Commare, Research Platform Cognitive Science, Department of Art History, University of Vienna; Raphael Rosenberg, Department of Art History, University of Vienna; Helmut Leder, Department of Psychological Basic Research, University of Vienna.

This research was supported by a grant from WWTF CS 11-023-LedRos to Helmut Leder.

Correspondence concerning this article should be addressed to Laura Commare, Research Platform Cognitive Science, Department of Art History, University of Vienna, Spitalgasse 2 (Hof 9), Vienna 1090, Austria. E-mail: [laura.commare@univie.ac.at](mailto:laura.commare@univie.ac.at)



Leder et al. (2004) this was referred to in the processing stage of cognitive mastering, which includes finding meaning and gaining understanding of what was perceived and recognized.

In art history, Panofsky (1994) offered a definition of content which differentiates these different aspects (Panofsky, 1994). Panofsky claimed that the content of artworks has *three main aspects* that are relevant for the perceiver at all times: (a) preiconographic content, that is, the depicted objects and their semantic interrelation akin to everyday situations (e.g., 13 male persons seated at a table), (b) iconographic content, that is, a classification of the preiconographic contents according to cultural interpretation conventions (e.g., the biblical story of the *Last Supper*), and (c) iconological content, that is, the broader *interpretation* of the artwork (e.g., the reason why *this specific Last Supper* painting was commissioned and painted at this time, in this specific place, in this specific manner). A description of the preiconographic content of an artwork would answer the question “What is depicted?” A description of an artwork’s iconographic content would answer the question “What does it mean?” (e.g., recognizing the *Last Supper* implies knowing about the biblical story; cf., Vygotsky, 1925/1974). A description of the iconological content would answer the question, “What does it mean within a general (historical) context?”

Panofsky’s (1994) three aspects of content correspond to different *processing stages* during the aesthetic engagement with works of art. In the model of aesthetic appreciation and judgment by Leder et al. (2004), image processing is described as comprising both early automated and later deliberate stages. Early processes include, i.e., perceptual analyses of form, object recognition, and explicit classification of content. Based on prior experience, objects are first recognized and evaluated in terms of familiarity or prototypicality. These processes establish the preiconographic content of the artwork. Then, on a second processing stage, the integration of preiconographic contents with declarative knowledge permits the classification of iconographic contents (cf., Gauthier, James, Curby, & Tarr, 2003). These analyses of preiconographic and iconographic contents are quite fast and, from a subjective perspective, can feel instantaneous. For instance, Augustin et al. (2008) found that some aspects of content, such as the preiconographic aspects of an artwork, can be processed within presentation times of only 10 ms (see also Biederman, 1981; Grill-Spector & Kanwisher, 2005; Thorpe, Fize, & Marlot, 1996). The analysis of iconological content, however, requires a form of deliberate processing, and therefore much longer times (Leder, Carbon, & Ripsas, 2006). During the deliberate processing phase, the perceiver *may* engage in an interpretative process by explicitly employing declarative knowledge to attribute iconological contents to the artwork (for a discussion on the situated interaction of individual and aesthetic object and its conditions, see Jacobsen, 2006, 2013).

According to the above mapping of Panofsky’s (1994) notion of multidimensional content and Leder et al.’s (2004) process model of aesthetic experience, it is apparent that declarative knowledge plays a crucial role throughout semantic engagement with artworks, and previous approaches explicitly addressing semantic complexity took this into account. They usually defined semantic complexity as a lack of understanding (Berlyne, 1960, 1963; Evans, 1970). Hence, content which perceivers rate as semantically complex is content which they have a hard time understand-

ing. As the semantic complexity of an artwork is thus equated with the *difficulty of semantic processing*, the difficulty a perceiver experiences when engaged with an artwork is conceived as relative to the amount of relevant declarative knowledge they have about the artwork. The more declarative knowledge someone possesses about the various contents of an artwork, the easier it should be (and feel) to recognize and interpret its contents (cf., Russell & Milne, 1997). However, such a definition of perceived semantic complexity does not do justice to the multidimensionality of an artwork’s content. In fact, declarative knowledge might not only ease the understanding of an artwork’s contents, but it also affects the extent to which a person can understand its meaning (Leder et al., 2006). Declarative knowledge enables the attribution and interpretation of iconographical and iconological content. Hence, the availability of declarative knowledge about a specific artwork might also imply the anticipation of the multidimensionality of the depicted content and, accordingly, a feeling of *semantic depth* (see also Kesner, 2014). Yet, the more semantic depth a perceiver is able to reveal within an artwork, the more information they have to deal with while engaging with this particular artwork. Thus, while higher ease with semantic processing should lead to a *decrease* of perceived complexity, the perception of semantic depth, in turn, should lead to an *increase* of perceived complexity.

In principle, it is conceivable that both the felt ease of semantic processing and semantic depth together constrain a perceiver’s perception of semantic complexity in an artwork. However, while the availability of relevant declarative knowledge is necessary and sufficient for the perception of semantic depth, it is not clear whether this is the case for semantic processing ease as well. The availability of declarative knowledge might be a prerequisite for the processing of iconographic and iconological aspects of content, but the mere availability of declarative knowledge might not necessarily render the processing easier. It may well be that semantic processing ease stems from well-established heuristics that facilitate the application of declarative knowledge (Ericsson & Kintsch, 1995).

Along this line of reasoning, we can expect systematic differences in the way art experts and nonexperts process artworks and judge their complexity. By definition, both groups differ systematically in the amount, structure, and availability of declarative knowledge about artworks (Augustin & Leder, 2006; Glaser, 1994; Honeck, Firment, & Case, 1987). Regarding the ease of semantic processing, experts—due to their extensive declarative knowledge as well as their well-established heuristics—should find it easier than nonexperts to understand the meaning of artworks (Axelsson, 2007; Kozbelt & Seeley, 2007; Müller, Höfel, Brattico, & Jacobsen, 2010; Silvia, 2006, 2013). Hence, art experts should rate the complexity of the same artworks lower than nonexperts. Furthermore, if semantic processing ease is indeed caused by the availability of application heuristics instead of declarative knowledge only, we should find this difference between experts and nonexperts to be robust even when both groups know the depicted contents equally well. Likewise, experts and nonexperts can be expected to differ in how they perceive semantic depth. Prior experience and expertise with artworks sets the baseline against which semantic depth is experienced and, hence, determines what qualifies as semantic depth in the first place (cf., Axelsson, 2007; Kesner, 2006; Pihko et al., 2011; Siegel, 1968). More specifically, it has been shown that art experts and nonexperts differ system-

atically in how they access content in art (Cupchik & Gebotys, 1988; Hekkert, 1995). Nonexperts are mostly concerned with preiconographical contents, whereas art experts focus on iconographical and iconological contents. Thus, when nonexperts attribute iconographical contents to an artwork, this imbues the artwork with meaning, potentially resulting in an experience of semantic depth (Cupchik & Gebotys, 1988; Shafto & Coley, 2003). Experts, on the contrary, focus on iconographical and, in particular, iconological contents when viewing art (Bauer, 2014; Nodine, Locher, & Krupinski, 1993; Winston & Cupchik, 1992). To them, attributing iconographical contents to a piece of art will likely feel relatively ordinary. Instead, they might only have a significant perception of semantic depth when they can attribute more abstract contents on an iconological level (Bauer, 2014). Hence, depending on the expertise of an individual, the perception of semantic depth might be mediated by different dimensions of content.

Moreover, the type and genre of an artwork might have an impact on perceived complexity as some genres genuinely convey more semantic complexity than others (cf., Swami, 2013). Here, the term “genre” is not an expression of style but rather refers to a classification of the topics dealt with, such as portrait, still life, landscape, abstract, or narrative-figurative painting (Duro, 2007). For instance, narrative-figurative paintings (e.g., the *Last Supper*), as well as most still life or landscape paintings, often contain a broad variety of topics in the same piece (e.g., communion of the Apostles and “ennoblement” of the client by including him in the depiction; Heehs, 1995). Consequently, the perception of semantic complexity may also differ between genres.

### The Present Study

The main aim of the present study was to test for content-related perception of complexity in artworks. Content is defined as a multidimensional concept, whose different aspects are accessed by declarative knowledge. We expect declarative knowledge about an artwork to increase perceived complexity if it enables the experience of semantic depth—but, we expect differences for experts and nonexperts, with respect to the particular relevance of the various dimensions of an artwork’s content in the perception of semantic depth. We further expect that perceived complexity is lower when semantic processing ease is high. We assume semantic processing ease to be connected to application heuristics rather than by declarative knowledge itself. If this is confirmed, then the level of perceived complexity should be lower in experts than in nonexperts. Finally, we test whether the genre of paintings (e.g., mythological vs. still life) is a modulating factor for both processing ease and perceived semantic depth.

To test these assumptions, we employed a between-subjects design that compared multifaceted self-reports of experts and nonexperts. To capture how perceived complexity varies with different aspects of content-related processing, we used a comprehensive questionnaire that asked participants to engage with a series of 90 artworks (over the course of several hours). Participants rated the overall perceived complexity as well as various content-related processes and the processing of formal art aspects (in accordance with the categories established by Nadal et al., 2010). To control for effects of art genre, the stimulus set comprised artworks from a broad variety of different genres and topics

(see the Stimuli section for details). The obtained data was analyzed with hierarchical mixed effect models, which allowed us to disentangle effects on different levels (e.g., artwork specific effects of declarative knowledge vs. general expertise effects).

### Method

#### Ethics Statement

The present study was carried out in accordance with the ethical guidelines of the Faculty of Psychology and the Department of Art History at the University of Vienna, and treated in accordance with the Declaration of Helsinki. All participants provided written informed consent prior to their participation and could withdraw at any time during the study without further consequences.

#### Participants

Ten art experts (7 female,  $M_{age} = 34.6$ ,  $SD = 9.1$ ) and 13 nonexperts (9 female,  $M_{age} = 32.7$ ,  $SD = 14.4$ ) participated in this study. Experts at least held a master’s degree in art history and had at least two years working experience in the field of art historical research or art education. The nonexpert participants were master students and PhD candidates from other disciplines who affirmed in the questionnaire that they had no formal training in the domain of art.

#### Stimuli

A small committee of art historian professionals chose the set of 90 digital high quality reproductions of paintings used in this study. The set covers a broad variety of styles from the late Middle Ages to contemporary art and can be roughly divided into three categories: abstract art ( $N = 13$ ); still life, landscape, and portrait painting ( $N = 28$ ); and narrative-figurative painting ( $N = 49$ ). These three genre categories differ in the predominant aspect of content: still life, landscape, and portrait paintings disclose quite obviously what they depict on a preiconographic level. Iconographic or iconological reasoning is not necessary to recognize a landscape, for example. This is different for narrative-figurative paintings, where iconographic content needs to be established to allow the viewer a meaningful classification of the content (e.g., 13 male persons seated at a table = *Last Supper*). Per definition, abstract art does not allow the automated classification of content at all, as it is not possible to recognize elements by their analogy to everyday life or to attribute iconographic content. Content in abstract art is therefore predominantly of an iconological nature.

The number of artworks in the respective genre categories is unbalanced because we balanced style and topics rather than genre. For historical reasons the variation of existing styles, as well as topics, in the respective genre categories is quite different. Abstract art varies less distinctively in styles and topics compared with still life, landscape, and portrait painting. Still life, landscape, and portrait painting cover less clearly distinct styles and topics than narrative-figurative painting. Hence, to assure that we may generalize effects across style differences and across topic categories we balanced both style and topics throughout the sample. A complete list of paintings ordered by genre category can be found in the Appendix.



## Procedure

The study was conducted in six separate identical approximately one-hour sessions, separated by short breaks. In each session, the participants rated 15 artworks in two blocks. The first block only asked for perceived complexity. Participants were requested to spontaneously rate the artwork's complexity on a 7-point scale (*According to you, how complex is this painting?*). The second block addressed form and content perception. Participants saw the same 15 paintings again and were asked to rate them on a 7-point scale according to a series of items, either referring to formal aspects or to content-related aspects. Items associated with formal aspects were chosen according to Nadal et al. (2010), who identified the following categories to be relevant for perceived visual complexity: amount of depicted elements, variety of depicted elements, intelligibility of elements, organization, symmetry, and variety of colors (the scales were: *I perceive the artwork to be rich in detail; I perceive the depicted elements to be variegated; The individual elements are readily identifiable; I perceive the artwork to be clearly arranged; I perceive the artwork to be symmetrical; I perceive the artwork to be colorful/varicoloured/I perceive the depicted colors to be natural*). As an additional formal aspect, we also included the perceived consistency of the depicted perspective (Andersen, 2007; *The artwork is depicted in a consistent perspective.*).

Items associated with content aspects refer either to the content aspects defined by Panofsky (1994) or to felt processing ease. Whether participants relied solely on preiconographic or also on iconographic reasoning was measured by their ability to classify the depicted content according to conventions (*Are you familiar with the depicted motif/topic? and If yes, please specify*). Iconological reasoning was more challenging to operationalize, as the measure would need to be an expression of perceived semantic depth achieved by a broader classification of the artwork for both experts and nonexperts. The perception of stylistic features or art historical relevance was thus not suited for this purpose, as nonexperts would not have been able to give meaningful answers to this question. Instead, we used a proxy item that asked for the general notoriety of the depicted content (*Taking a superordinate perspective, altogether this content can be classified as common/generally understandable*). This ensured that participants had to judge the depicted content in relation to other topics, and therefore perform a broader classification of that content. Felt processing ease was measured both directly and indirectly by two items referring to participant's experience with the depicted content (*The depicted motif/topic was easy to identify/The depiction of the motif/topic deviates from how I had expected it to be depicted*).

In addition, we also tested participants' familiarity with each artwork, and asked them to provide a liking rating. In all six sessions, paintings were presented in random order.

## Task Analysis

Several studies demonstrated a clear-cut dissociation between affective and cognitive modes of aesthetic contemplation and aesthetic judgments (Brattico et al., 2016; Jacobsen & Höfel, 2003). Some descriptive aesthetic judgments, for example, of symmetry, seem to be related to automated and spontaneous processes (Jacobsen & Höfel, 2003). Aesthetic appreciation on the contrary, was shown to require either intrinsic or (task) induced

intention (Höfel & Jacobsen, 2007; Istók, Brattico, Jacobsen, Ritter, & Tervaniemi, 2013). Task demands can therefore essentially structure how participants engage with the presented artworks. In the present study participants were always asked to engage with both descriptive aesthetic judgments and evaluative aesthetic judgments (liking). This should ensure that all participants, regardless of differences in intrinsic intention, would maintain an aesthetic stance during the task and therefore engage in consistent and equivalent manner with the artworks.

## Data Analysis

Despite being cross-sectional, the collected data showed substantial clustering, as the single rating observations were nested within subjects. Such a clustered design provides rich information about processes operating at different levels (Rabe-Hesketh & Skrondal, 2012); it is possible to disentangle determinants of single judgments like declarative knowledge about specific content from general personal abilities that are not exclusively linked to a single artwork, such as expertise. Despite its explanatory benefits, the nested data-structure violates the standard assumption of independent observations of ordinary least squares regression models, since the observations within the same individual are not autonomous. To handle the dependence of observations and exploit the richness of the data, we employed a multilevel approach in our analysis (cf., also Silvia, 2007).

Multilevel or hierarchical models contain fixed effects and random effects. The fixed effects can be interpreted as analogous to standard regression coefficients, and are estimated directly. The random effects can be interpreted as heterogeneity within the main effects and are not directly estimated, but summarized according to their estimated variances and covariances. Random effects may take the form of either random intercepts or random coefficients.

Following Gelman and Hill (2007) we used the following general hierarchical model specification:

$$y_i \sim N(X_i^0 \beta^0 + X_i B_{j \Gamma \Gamma}, \sigma_\epsilon^2), \quad \text{for } i = 1, \dots, n$$

$$B_j \sim N(U_j G, \Sigma_B), \quad \text{for } j = 1, \dots, J$$

where  $X^0$  is the  $n \times R$  matrix of all individual predictors, and  $\beta^0$  is the vector of their unmodeled regression coefficients.  $X$  is the  $n \times K$  matrix of individual predictors that have coefficients varying by groups, including a varying intercept.  $B$  is the  $J \times K$  matrix of regression coefficients.  $U$  is the  $J \times L$  matrix of group-level predictors and  $G$  is the  $L \times K$  matrix of coefficients for the group-level regression.  $\Sigma_B$  is the covariance of the varying intercepts and slopes.

We estimated three separate sets of mixed-linear models. The first two sets include all 90 trials (artworks) per participant, and served to investigate the role of expertise, as well as iconographic classification (via declarative knowledge) for perceived complexity. They differ only in the number of included predictors. The third set of models extended the first set by including genre categories. The third set of models limited the analysis to artworks that both experts and nonexperts, could relate to on an iconographic level. Consequently, the number of trials included varied across participants. This was done to (a) more specifically inspect the effect of expertise, and (b) to provide a meaningful subset of trials for the inspection of iconological reasoning. Iconological

reasoning requires that the iconographic content is known to the observer. We also restricted the analysis of subjective feelings of processing ease to this subset of artworks. This was done to ensure that the effect can indeed be related to the application of declarative knowledge. As Lüdtké et al. (2008) reported no effect of varying cluster size in mixed linear models, we did not correct for the unbalanced trial numbers.

All models include controls for familiarity with the respective artworks, as well as visual complexity. To ameliorate estimation power as well as the readability of the results, we included the items related to visual complexity as single additive index ranging from 7 to 49 (Cronbach's alpha = 0.7028). We also accounted for individual variance in the data which is related to unobserved factors such as taste (Jacobsen, 2004) by modeling random intercepts. Additionally, we also modeled random slopes for the respective main predictor variables to account and correct for effect heterogeneity. To interpret the respective effect heterogeneities, we provide intraclass correlation coefficients (ICCs) for each random parameter.

**Results**

The first block of analyses investigated how expertise and the attribution of iconographic content via declarative knowledge influence perceived complexity (Table 1). First, we ran a mixed linear model with random intercept and random slopes for the included predictors for all participants, both experts and nonexperts (Table 1, Model 1). Besides declarative knowledge, the model includes the additive index for visual complexity and familiarity ratings as additional Level 1 predictors, and expertise as a Level 2 predictor.

Our main hypothesis was confirmed; perceived complexity was not only significantly influenced by visual complexity ( $\beta = 0.154, p < .001$ ) but also by content-related processing. As expected we found declarative knowledge to have a significant positive effect on perceived complexity ( $\beta = 0.231, p < .05$ ). Expertise on the other hand has a significant negative effect on perceived complex-

ity ( $\beta = -0.521, p < .01$ ). The very low intraclass correlation (ICC <0.001) for the effect of expertise suggested a very consistent effect among the sample. This held true for visual complexity as well, where effect heterogeneity only accounted for a variance of below 0.1% (ICC <0.001). The effect of declarative knowledge, however, was characterized by considerable variance between individuals (ICC = 0.18).

We therefore ran separate models for experts (Table 1, Model 2) and nonexperts (Table 1, Model 3). A comparison of Model 2 and 3 revealed that, as expected, declarative knowledge had a significant positive effect on perceived complexity only for the group of nonexperts ( $\beta = 0.259, p < .05$ ). However, the effect heterogeneity was still substantial (ICC = 0.15), and may stem from other factors than expertise, such as genre characteristics.

The second block of analysis took up the reported effect heterogeneities and investigated the mediating role of genre categories (Table 2). When including genre categories in the analysis, we found that genre did indeed affect perceived complexity. Narrative-figurative art was judged to be more complex when compared with the reference category landscape, still life, and portrait painting. This was true for the model including both experts and nonexperts (Table 2, Model 1;  $\beta = 0.351, p < .001$ ), as well as for the group-specific models (Table 2, Model 2;  $\beta = 0.535, p < .001$ , Table 2, Model 3.1;  $\beta = 0.218, p < .01$ ). However, differences between the groups were found regarding abstract art, while experts showed no significant difference in their complexity judgments of landscape, still life, and portrait painting and abstract art (Table 2, Model 2;  $\beta = -0.234, p > .05$ ), nonexperts perceived abstract art to be the least complex genre category (Table 2, Model 3.1;  $\beta = -0.264, p < .05$ ). This was only the case when not taking the interindividual variances into account. When accounting for effect heterogeneity, the main effect was not significant (Table 2, Model 3.2;  $\beta = -0.268, p > .05$ ). We also found that nonexperts showed considerable variance regarding their evaluation of narrative-figurative art (ICC = 0.21).

Table 1  
Complexity Ratings Prediction Based on Declarative Knowledge and Expertise (Mixed Linear Models With Random Intercepts and Random Slopes for Selected Predictors)

	Model 1 (all)		Model 2 (experts)		Model 3 (nonexperts)	
<b>Fixed effects</b>						
Expertise		-.521**				
Declarative knowledge		.231*	.217		.259*	
Visual complexity		.154***	.162***		.146***	
Familiarity		.073	.091		.129	
Intercept		.413	-.262		.555	
	Estimate	ICC	Estimate	ICC	Estimate	ICC
<b>Random effects</b>						
Expertise	5.09 × 10 <sup>-6</sup>	<.001				
Declarative knowledge	.258	.18	.118	.08	.204	.15
Visual complexity	.008	<.001	.008	<.001	.273	<.001
Intercept	.311		7.58 × 10 <sup>-6</sup>		.879	
Residual	1.201		1.228		1.164	

Note. ICC = intraclass correlation coefficient.  
\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .



Table 2  
Complexity Ratings Prediction Based on Declarative Knowledge and Genre (Mixed Linear Models With Random Intercepts and Random Slopes for Selected Predictors)

	Model 1 (all)	Model 2 (experts)	Model 3.1 (nonexperts)	Model 3.2 (nonexperts)
<b>Fixed effects</b>				
Expertise	-.484**			
Declarative knowledge	.160*	.163	.210*	.170
Abstract art	-.254	-.234	-.264*	-.268
Narrative-figurative art	.352***	.535***	.218**	.217*
Visual complexity	.155***	.163***	.144***	.144***
Familiarity	.051	.045	.184	.152
Intercept	.237	-.504	.550	.554
	Estimate	ICC	Estimate	ICC
<b>Random effects</b>				
Abstract art	.675	.37	$1.96 \times 10^{-8}$	.11
Narrative-figurative art	.225	.16	.081	.06
Intercept	.391		$7.58 \times 10^{-6}$	
Residual	1.158		1.191	

Note. ICC = intraclass correlation coefficient.  
\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Experts, by comparison, were more consistent in this regard (ICC = 0.06).

Besides influencing perceived complexity directly, genre categories also mediated the effects of declarative knowledge and expertise. Although we found the effect of expertise to be robust when including genre, we found that it was overestimated in the first analysis block (Table 2, Model 1;  $\beta = -0.433, p < .01$  vs. Table 1, Model 1;  $\beta = -0.521, p < .01$ ). The same applied to the effect of declarative knowledge: the effect was slightly overestimated in the first analysis (Table 2, Model 3;  $\beta = 0.210, p < .05$  vs. Table 1, Model 3;  $\beta = 0.259, p < .05$ ). Additionally, when accounting for genre the main effect of declarative knowledge was only significant when not considering the effect heterogeneity of the genre effect (Table 2, Model 3.2;  $\beta = 0.170, p > .05$ ). Hence, the effects of declarative knowledge and expertise on perceived complexity were not independent of genre.

The inclusion of genre did not alter the effect visual complexity on perceived complexity (Table 2, Model 1;  $\beta = 0.155, p < .001$  vs. Table 1, Model 1;  $\beta = 0.154, p < .001$ ). Comparing the effects of visual complexity obtained from the reported models with a reference model, which only includes visual complexity and a random intercept (Table 3), revealed that the effect size was consistent across all conditions. The effect of visual complexity on overall perceived complexity thus was independent of the investigated aspects of semantic complexity.

As the effects of visual complexity and aspects of semantic complexity proved to be independent, it is not surprising that the full model (Table 2, Model 1) showed a significantly better data fit than the reference model (see Table 3; Akaike information criterion [AIC]:  $AIC_{\text{reference}} = 6,704.475, AIC_{\text{Table 2 Model 1}} = 6,570.32$ ; Bayesian information criterion [BIC]:  $BIC_{\text{reference}} = 6,726.997, BIC_{\text{Table 2 Model 1}} = 6,632.244$ ). Hence, visual and semantic complexity both contributed significantly to overall perceived complexity.

The last block of analyses addressed how iconological reasoning, that is to say, building semantic depth by performing a broader classification of the artwork, and subjective feelings of semantic

processing ease contribute to perceived complexity (Table 4). Besides a proxy measure for iconological reasoning (*universal notoriety*), and both a direct and indirect measure for subjective feelings of processing ease (*processing ease* and *deviation from expectations*), all models included the additive index for visual complexity and familiarity ratings as additional Level 1 predictors. Only artworks (trials) participants could relate declarative knowledge to were included.

Model 1 included no random intercepts, and only served to prove that the effect of expertise was consistent in this subset of trials ( $\beta = -0.573, p < .05$ ): Expertise still led to lower complexity ratings even when the artwork's iconographic content was equally known to both experts and nonexperts, and when visual complexity as well as the degree of familiarity with the artworks were controlled for.

A comparison of the group-specific models revealed, as expected, that broader (iconological) classification of the content (universal notoriety) only influenced perceived complexity significantly in the group of experts (Table 4, Model 2;  $\beta = -0.145, p < .001$ ); experts rated artworks with more commonly known

Table 3  
Complexity Ratings Prediction Based on Visual Complexity (Mixed Linear Model With Random Intercept)

	Reference model (all)
<b>Fixed effects</b>	
Visual complexity	.153***
Intercept	.329
	Estimate
<b>Random effects</b>	
Intercept	.360
Residual	1.214

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

T3, AQ: 4

AQ: 5

Table 4  
Complexity Ratings Prediction Based on the Iconological Reasoning and Felt Processing Ease  
(Mixed Linear Models With Random Intercepts and Random Slopes for Selected Predictors)

	Model 1 (all)	Model 2 (experts)		Model 3.1 (nonexperts)	Model 3.2 (nonexperts)	
Fixed effects						
Expertise	-.566*					
Processing ease	-.016	-.017		-.021	-.009	
Deviation from expectations	-.007	-.030		.132*	.127	
Universal notoriety	-.135***	-.145***		.125	-.019	
Visual complexity	.145***	.148***		.125***	.122***	
Familiarity	-.038	.013		.007	.066	
Intercept	1.719	1.206		.974	.969	
	Estimate	Estimate	ICC	Estimate	Estimate	ICC
Random effects						
Processing ease		$3.47 \times 10^{-11}$	<.001		.060	.05
Deviation from expectations		.039	.03		.089	.07
Intercept	.455	.172		.598	$1.30 \times 10^{-6}$	
Residual	1.160	1.124		1.205	1.118	

Note. ICC = intraclass correlation coefficient.  
\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

contents to be less complex. This was not the case for the group of nonexperts (Table 4, Model 3.1;  $\beta = 0.125, p > .05$ ).

There was no significant effect of subjectively felt processing ease on experts' complexity ratings (Table 4, Model 2; processing ease:  $\beta = -0.030, p > .05$ ; Table 4, Model 2; deviation from expectations:  $\beta = -0.017, p > .05$ ). Nonexperts' complexity ratings were influenced significantly only by the indirect measure of processing ease (Table 4, Model 3.1; deviation from expectations:  $\beta = 0.132, p < .05$ ) but not by the direct measure (Table 4, Model 3.1; processing ease:  $\beta = -0.021, p > .05$ ). This, however, only held true when not accounting for effect heterogeneity. When accounting for interindividual variances (ICC = 0.07) the main effect did not reach significance (Table 4, Model 3.2;  $\beta = 0.127, p > .05$ ). Hence, we did not find any systematic effect of subjectively felt processing ease on perceived complexity.

Finally, in none of the reported models did we find familiarity to significantly influence perceived complexity.

### Discussion and Conclusion

We studied the role of content-related processing for perceived complexity in paintings. In an empirical study that compared various self-reported evaluations from art experts and nonexperts, the main hypothesis was confirmed: The perceived complexity of paintings was significantly influenced by aspects of content-related processing in both experts and nonexperts. Furthermore, content-related effects were independent of the paintings' visual complexity (cf., also Shigeto & Nittono, 2010). Hence, content-related and form related processing each contributed autonomously to overall perceived complexity.

Specifically, we investigated two aspects of content-related processing—the role of semantic processing ease, and the role of (perceived) semantic depth. Both aspects were important for participants' overall judgments of an artwork's complexity.

### Semantic Processing Ease

Semantic processing ease influenced participants' complexity judgments negatively; perceived complexity decreased when processing the artwork's content was easy. This effect, however, was not phenomenologically accessible to the participants; subjective feelings of semantic processing ease did not contribute significantly to perceived complexity. Yet, expertise did. Experts consistently judged artworks to be less complex than did nonexperts (cf., Lundy, 2010). They did so even when visual complexity, artwork-specific declarative knowledge, and familiarity was controlled for. The effect was also found when the content/topic of the studied artworks was equally known to both experts and nonexperts. This suggests that semantic processing ease stems from expert characteristics other than extensive declarative knowledge, such as interpretation or classification heuristics.

At first glance, it may seem like a bold claim to interpret the expertise effect only with respect to semantic processing ease. Certainly, experts develop elaborate automated classification heuristics and processes of making meaning which ease their semantic processing (Axelsson, 2007; Kozbelt & Seeley, 2007). But, as Blazhenkova and Kozhevnikov (2010) have shown, experts also develop a greater visual-object ability which aids their processing of form. However, we did not find substantial changes in the effect of visual complexity when comparing a model that predicts perceived complexity based on visual complexity only to a model that also includes expertise. Moreover, we did not find differences in how experts and nonexperts were influenced by visual complexity in their overall complexity judgments when investigating the groups separately. Therefore, we may conclude that the reported expertise effect is at least, to a great extent, independent of form-related processing. Additionally, we find genre, which by definition is a semantic category, to influence the expertise effect; the expertise effect decreased when controlling for genre. This sug-



gests that the expertise effect depends on genre, and is indeed related to semantic processing.

### Perceived Semantic Depth

The second aspect of content-related processing we focused on was the perception of semantic depth. Perception of semantic depth affected complexity judgments positively; perceived complexity increased when participants perceived semantic depth. Yet, semantic depth is a concept relative to declarative knowledge, and its application heuristics. The baseline against which semantic depth is experienced depends on previous experience (Pihko et al., 2011). Experts are used to and thus able to perceive (or attribute) multiple layers of content. Nonexperts on the contrary, lack comparable declarative knowledge and cannot relate to the same amount of content aspects as experts. Experts and nonexperts thus have different baselines for the evaluation of semantic depth (Bauer, 2014).

Moreover, nonexperts were shown to be mostly concerned with personal feelings and building analogies to everyday life when engaging with art (Cupchik & Gebotys, 1988; Shafto & Coley, 2003). They focus on what is depicted rather than on what the depiction means. It has, however, been demonstrated that additional information about the depiction's content and meaning significantly enhanced participants' aesthetic appreciation and interest (Belke, Leder, & Carbon, 2015; Leder et al., 2004; Millis, 2001; Swami, 2013). To this end, we found a corresponding effect; when nonexpert participants possessed declarative knowledge about the depicted content they judged the artworks to be significantly more complex.

Experts were shown to be mostly concerned with questions of style and art historical significance (Nodine et al., 1993; Winston & Cupchik, 1992); they generally do not focus on what is depicted, but are instead occupied with the artwork's meaning in a broader sense. Yet, the kind of meaning experts are mostly interested in clearly differs from the kind of meaning nonexperts are concerned about: Nonexperts are shown to be more interested in meaning that explains the depiction; experts, on the contrary, are shown to be interested in meaning that transcends the single depiction, and explains it against the background of a greater historical or art historical context (Bauer, 2014; Silvia, 2013). Following the suggestions of the art historian Erwin Panofsky, we refer to the first as iconographic content and to the latter as iconological content (Panofsky, 1994).

As outlined above, we found that nonexperts judged art to be more complex, when their declarative knowledge allowed them to attribute iconographic content. However, they were not affected in their complexity judgments by a broader classification of the artwork, which would correspond to iconological reasoning. Experts, on the contrary, showed a reverse pattern. Their complexity judgments were not influenced by the attribution of iconographic content, but they judged artworks which, on an iconological level, revealed semantic depth to them to be more complex.

This difference is not just an expression of diverging interest focus. Iconographic content, which is important to nonexperts, is a classification of content according to conventions. Such content classification happens mostly during automated phases of processing (Leder et al., 2004; Leder & Nadal, 2014). Iconological content, which is important to experts, is the consequence of more

elaborate cognitive mastering. It should thus be attributed to an artwork during deliberate stages of processing (Leder et al., 2006). Hence, the experience of semantic depth in the respective groups unfolds on different time scales. As the experience of semantic depth nevertheless influences perceived complexity in both groups, this may have consequences for the measurement of perceived complexity—especially with short presentation times. To ensure that complexity judgments rely on the same parameters in all participants, and hence measure the same concept, future studies should account for participants' expertise status, and, if necessary, adapt presentation times.

### Genre

We also found effects of genre in our study. Different genre types directly affected perceived complexity. We found that narrative-figurative art was generally judged to be more complex than landscape, still life, and portrait painting. This direct effect may be the result of some topics being conventionally more complex than others; for example, there may be implicit social agreement that a depiction of human interaction is per se semantically more complex than a landscape (cf., Vygotsky, 1925/1974). Abstract art marks a special case in this regard; while experts do not judge it considerably different from landscape, still life, and portrait painting, nonexperts consider abstract art to be the least complex art genre (see also Pihko et al., 2011)—even though visual complexity was systematically controlled for. The reason for this may be that abstract art per definition hinders the recognition of depicted elements and iconographic classification. Content in abstract art is therefore predominantly accessed via iconological reasoning. Nonexperts, however, lack the declarative knowledge necessary to attribute iconological content, and are thus unable to experience semantic depth in abstract art.

However, not all nonexpert participants consistently perceived abstract art to be less complex than any other genre. There was substantial variability within this effect (cf., Schepman, Rodway, & Pullen, 2015). This applied to genre effects in general (nonexperts as well as experts). Nevertheless, the variance was much more pronounced in nonexperts than in experts. This may be due to various reasons, such as differing personality traits (Chamorro-Premuzic, Burke, Hsu, & Swami, 2010; Feist & Brady, 2004; Furnham & Walker, 2001).

In the nonexpert group, these differences in sensibility to genre also affected the effect of declarative knowledge; when we accounted for the variance of genre effects, the main effect of declarative knowledge on perceived complexity was not significant. This suggests that participants reverted to different strategies when judging semantic aspects of complexity in art. While some stuck to the vague content classifications that genre categories offer, others relied on iconographic classification of content via their declarative knowledge.

As mentioned above, genre also mediated the effect of expertise; we found the expertise effect to decrease when controlling for genre. As some genres can be assumed to be more complex than others, we may relate this effect to general genre characteristics; the simpler the semantic structure of a given artwork, the easier is the semantic processing *in general*. Consequently, we found the difference between experts and nonexperts to be smaller when semantic processing was easy for everyone.

### Concluding Remarks

In summary, we have shown that content-related processing influenced perceived complexity independent of a painting's visual complexity. We identified two main sources of influence: semantic processing ease and perceived semantic depth. While semantic processing ease seems to influence complexity judgments implicitly, the perception of semantic depth could be directly linked to complexity judgments. Furthermore, we found that experts and nonexperts differ in how they perceive complexity in art. Moreover, we can state that nonexperts are far less consistent than experts in their complexity judgments and the respective strategies they apply. While traditional concepts of (perceived) complexity have indeed often overseen the influence of content-related processing, the present study only marks the first step in understanding the complex mechanisms underlying perceived complexity in the domain of art. Future research must shed more light on the details of the processes involved.

### References

- Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience*, 25, 339–379. <http://dx.doi.org/10.1146/annurev.neuro.25.112701.142900>
- Andersen, K. (2007). *The geometry of art—The history of the mathematical theory of perspective from Alberti to Monge*. New York, NY: Springer.
- Augustin, M. D., & Leder, H. (2006). Art expertise: A study of concepts and conceptual spaces. *Psychology Science*, 48, 135–156.
- Augustin, M. D., Leder, H., Hutzler, F., & Carbon, C.-C. (2008). Style follows content: On the microgenesis of art perception. *Acta Psychologica*, 128, 127–138. <http://dx.doi.org/10.1016/j.actpsy.2007.11.006>
- Axelsson, Ö. (2007). Individual differences in preferences for photographs. *Psychology of Aesthetics, Creativity, and the Arts*, 1, 61–72. <http://dx.doi.org/10.1037/1931-3896.1.2.61>
- Bauer, D. (2014). *Meaning making with art: Expert and lay perspectives in understanding artworks and exhibition concepts* (Unpublished doctoral thesis). Eberhard Karls University, Tübingen, E-Thesis. Retrieved from [https://publikationen.uni-tuebingen.de/xmlui/bitstream/handle/10900/56823/Dissertationsschrift\\_Daniela%20Bauer.pdf?sequence=1](https://publikationen.uni-tuebingen.de/xmlui/bitstream/handle/10900/56823/Dissertationsschrift_Daniela%20Bauer.pdf?sequence=1)
- Belke, B., Leder, H., & Carbon, C. C. (2015). When challenging art gets liked: Evidences for a dual preference formation process for fluent and non-fluent portraits. *PLoS ONE*, 10, e0131796. <http://dx.doi.org/10.1371/journal.pone.0131796>
- Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. Toronto, Canada: McGraw-Hill. <http://dx.doi.org/10.1037/11164-000>
- Berlyne, D. E. (1963). Complexity and incongruity variables as determinants of exploratory choice and evaluative ratings. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 17, 274–290. <http://dx.doi.org/10.1037/h0092883>
- Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy, & J. R. Pomerantz, (Eds.), *Perceptual organization* (pp. 213–253). Hillsdale, NJ: Erlbaum.
- Blazhenkova, O., & Kozhevnikov, M. (2010). Visual-object ability: A new dimension of non-verbal intelligence. *Cognition*, 117, 276–301. <http://dx.doi.org/10.1016/j.cognition.2010.08.021>
- Brattico, E., Bogert, B., Alluri, V., Tervaniemi, M., Eerola, T., & Jacobsen, T. (2016). It's sad but i like it: The neural dissociation between musical emotions and liking in experts and laypersons. *Frontiers in Human Neuroscience*, 9, 676. <http://dx.doi.org/10.3389/fnhum.2015.00676>
- Cavalcante, A., Mansouri, A., Mansouri, L. K. E., Kardec, B., Takeuchi, Y., Matsumoto, N., & Ohnishi, N. (2011). *On the relation between contrast and the perception of visual complexity*. Paper presented at the 21st Annual Conference of the Japanese Neural Network Society, Okinawa, Japan.
- Chamorro-Premuzic, T., Burke, C., Hsu, A., & Swami, V. (2010). Personality predictors of artistic preferences as a function of the emotional valence and perceived complexity of paintings. *Psychology of Aesthetics, Creativity, and the Arts*, 4, 196–204. <http://dx.doi.org/10.1037/a0019211>
- Chikhman, V., Bondarko, V., Danilova, M., Goluzina, A., & Shelepin, Y. (2012). Complexity of images: Experimental and computational estimates compared. *Perception*, 41, 631–647. <http://dx.doi.org/10.1068/p6987>
- Cupchik, G. C., & Gebotys, R. J. (1988). The search for meaning in art: Interpretive styles and judgments of quality. *Visual Arts Research*, 14, 38–50.
- Duro, P. (2007). Imitation and authority: The creation of the academic canon in French art, 1648–1870. A. Brzyski, (Ed.), *Partisan canons* (pp. 95–114). Durham, NC: Duke University Press.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211–245. <http://dx.doi.org/10.1037/0033-295X.102.2.211>
- Evans, D. R. (1970). Conceptual complexity, arousal, and epistemic behaviour. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 24, 249–260. <http://dx.doi.org/10.1037/h0082861>
- Feist, G. J., & Brady, T. R. (2004). Openness to experience, non-conformity, and the preference for abstract art. *Empirical Studies of the Arts*, 22, 77–89. <http://dx.doi.org/10.2190/Y7CA-TBY6-V7LR-76GK>
- Forsythe, A., Nadal, M., Sheehy, N., Cela-Conde, C. J., & Sawey, M. (2011). Predicting beauty: Fractal dimension and visual complexity in art. *British Journal of Psychology*, 102, 49–70. <http://dx.doi.org/10.1348/000712610X498958>
- Furnham, A., & Walker, J. (2001). Personality and judgements of abstract, pop art, and representational paintings. *European Journal of Personality*, 15, 57–72. <http://dx.doi.org/10.1002/per.340>
- Gartus, A., & Leder, H. (2013). The small step toward asymmetry: Aesthetic judgment of broken symmetries. *i-Perception*, 4, 361–364. <http://dx.doi.org/10.1068/i0588sas>
- Gauthier, I., James, T. W., Curby, K. M., & Tarr, M. J. (2003). The influence of conceptual knowledge on visual discrimination. *Cognitive Neuropsychology*, 20, 507–523. <http://dx.doi.org/10.1080/02643290244000275>
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multi-level/hierarchical models*. New York, NY: Cambridge University Press.
- Glaser, R. (1994). Expertise. In M. W. Eysenck (Eds.), *The Blackwell dictionary of cognitive psychology* (pp. 139–142). Cambridge, UK: Blackwell.
- Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition. *Psychological Science*, 16, 152–160. <http://dx.doi.org/10.1111/j.0956-7976.2005.00796.x>
- Heehs, P. (1995). Narrative painting and narratives about paintings: Pousin among the philosophers. *Narrative*, 3, 211–231.
- Hekkert, P. (1995). *Artful judgements. A psychological inquiry into aesthetic preference for visual patterns* (Unpublished doctoral thesis, Delft University of Technology, the Netherlands).
- Höfel, L., & Jacobsen, T. (2007). Electrophysiological indices of processing aesthetics: Spontaneous or intentional processes? *International Journal of Psychophysiology*, 65, 20–31. <http://dx.doi.org/10.1016/j.ijpsycho.2007.02.007>
- Honeck, R. P., Firment, M., & Case, T. J. (1987). Expertise and categorization. *Bulletin of the Psychonomic Society*, 25, 431–434. <http://dx.doi.org/10.3758/BF03334732>
- Istók, E., Brattico, E., Jacobsen, T., Ritter, A., & Tervaniemi, M. (2013). 'I love rock 'n' roll'—music genre preference modulates brain responses



- to music. *Biological Psychology*, 92, 142–151. <http://dx.doi.org/10.1016/j.biopsycho.2012.11.005>
- Jacobsen, T. (2004). Individual and group modelling of aesthetic judgment strategies. *British Journal of Psychology*, 95, 41–56. <http://dx.doi.org/10.1348/000712604322779451>
- Jacobsen, T. (2006). Bridging the arts and sciences: A framework for the psychology of aesthetics. *Leonardo*, 39, 155–162. <http://dx.doi.org/10.1162/leon.2006.39.2.155>
- Jacobsen, T. (2013). On the electrophysiology of aesthetic processing. *Progress in Brain Research*, 204, 159–168. <http://dx.doi.org/10.1016/B978-0-444-63287-6.00008-7>
- Jacobsen, T., & Höfel, L. (2003). Descriptive and evaluative judgment processes: Behavioral and electrophysiological indices of processing symmetry and aesthetics. *Cognitive, Affective & Behavioral Neuroscience*, 3, 289–299. <http://dx.doi.org/10.3758/CABN.3.4.289>
- Jakesch, M., & Leder, H. (2015). The qualitative side of complexity: Testing effects of ambiguity on complexity judgments. *Psychology of Aesthetics, Creativity, and the Arts*, 9, 200–205. <http://dx.doi.org/10.1037/a0039350>
- Kesner, L. (2006). The role of cognitive competence in the art museum experience. *Museum Management and Curatorship*, 21, 4–19. <http://dx.doi.org/10.1080/09647770600302101>
- Kesner, L. (2014). The predictive mind and the experience of visual art work. *Frontiers in Psychology*, 5, 1417. <http://dx.doi.org/10.3389/fpsyg.2014.01417>
- Kozbelt, A., & Seeley, W. P. (2007). Integrating art historical, psychological, and neuroscientific explanations of artists' advantaged in drawing and perception. *Psychology of Aesthetics, Creativity, and the Arts*, 1, 80–90. <http://dx.doi.org/10.1037/1931-3896.1.2.80>
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95, 489–508. <http://dx.doi.org/10.1348/0007126042369811>
- Leder, H., Carbon, C. C., & Ripsas, A. L. (2006). Entitling art: Influence of title information on understanding and appreciation of paintings. *Acta Psychologica*, 121, 176–198. <http://dx.doi.org/10.1016/j.actpsy.2005.08.005>
- 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, 443–464. <http://dx.doi.org/10.1111/bjop.12084>
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13, 203–229. <http://dx.doi.org/10.1037/a0012869>
- Lundy, D. E. (2010). A test of consensus in aesthetic evaluation among professional critics of modern music. *Empirical Studies of the Arts*, 28, 243–258. <http://dx.doi.org/10.2190/EM.28.2.h>
- 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, e72412. <http://dx.doi.org/10.1371/journal.pone.0072412>
- Millis, K. (2001). Making meaning brings pleasure: The influence of titles on aesthetic experiences. *Emotion*, 1, 320–329. <http://dx.doi.org/10.1037/1528-3542.1.3.320>
- Müller, M., Höfel, L., Brattico, E., & Jacobsen, T. (2010). Aesthetic judgments of music in experts and laypersons—An ERP study. *International Journal of Psychophysiology*, 76, 40–51. <http://dx.doi.org/10.1016/j.ijpsycho.2010.02.002>
- Nadal, M., Munar, E., Marty, G., & Cela-Conde, C. J. (2010). Visual complexity and beauty appreciation: Explaining the divergence of results. *Empirical Studies of the Arts*, 28, 173–191. <http://dx.doi.org/10.2190/EM.28.2.d>
- Nodine, C. F., Locher, P. J., & Krupinski, E. A. (1993). The role of formal art training on perception and aesthetic judgment of art compositions. *Leonardo*, 26, 219–227. <http://dx.doi.org/10.2307/1575815>
- Panofsky, E. (1994). *Ikonographie und Ikonologie*. In E. Kaemmerling (Ed.), *Bildende Kunst als Zeichensystem. Ikonographie und Ikonologie. Band 1: Theorien—Entwicklung—Probleme* (pp. 207–225). Cologne, Germany: DuMont.
- Pihko, E., Virtanen, A., Saarinen, V. M., Pannasch, S., Hirvenkari, L., Tossavainen, T., . . . Hari, R. (2011). Experiencing art: The influence of expertise and painting abstraction level. *Frontiers in Human Neuroscience*, 5, 94. <http://dx.doi.org/10.3389/fnhum.2011.00094>
- Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modelling using Stata*. College Station, TX: Stata Press.
- Russell, P. A., & Milne, S. (1997). Meaningfulness and hedonic value of paintings: Effects of titles. *Empirical Studies of the Arts*, 15, 61–73. <http://dx.doi.org/10.2190/EHT3-HWVM-52CB-8QHJ>
- Schepman, A., Rodway, P., & Pullen, S. J. (2015). Greater cross-viewer similarity of semantic associations for representational than for abstract artworks. *Journal of Vision*, 15, 12. <http://dx.doi.org/10.1167/15.14.12>
- Shafto, P., & Coley, J. D. (2003). Development of categorization and reasoning in the natural world: Novices to experts, naive similarity to ecological knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 641–649. <http://dx.doi.org/10.1037/0278-7393.29.4.641>
- Shiget, H., & Nittono, H. (2010). *Stimulus complexity and meaningfulness differently affect the viewing duration, recognition performance, and subjective interest of the viewer*. Paper presented at the Third International Workshop on Kansei, Fukuoka, Japan.
- Siegel, L. S. (1968). The role of experience in judgments of complexity of random shapes. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 22, 474–483. <http://dx.doi.org/10.1037/h0082787>
- Silvia, P. J. (2006). Artistic training and interest in visual art: Applying the appraisal model of aesthetic emotions. *Empirical Studies of the Arts*, 24, 139–161. <http://dx.doi.org/10.2190/DX8K-6WEA-6WPA-FM84>
- Silvia, P. J. (2007). An introduction to multilevel modeling for research on the psychology of art and creativity. *Empirical Studies of the Arts*, 25, 1–20. <http://dx.doi.org/10.2190/6780-361T-3J83-04L1>
- Silvia, P. J. (2013). Interested experts, confused novices: Art expertise and the knowledge emotions. *Empirical Studies of the Arts*, 31, 107–115. <http://dx.doi.org/10.2190/EM.31.1.f>
- Swami, V. (2013). Context matters: Investigating the impact of contextual information on the aesthetic appreciation of paintings by Max Ernst and Pablo Picasso. *Psychology of Aesthetics, Creativity, and the Arts*, 7, 285–295. <http://dx.doi.org/10.1037/a0030965>
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522. <http://dx.doi.org/10.1038/381520a0>
- Vygotsky, L. S. (1974). *The Psychology of Art*. Cambridge, MA: MIT Press. (Original work published 1925)
- Winston, A. S., & Cupchik, G. C. (1992). The evaluation of high art and popular art by naive and experienced viewers. *Visual Arts Research*, 18, 1–14.

AQ: 9

(Appendix follows)

Appendix

AQ: 11

List of Artworks

List of Artworks Ordered by Genre Category Including Mean Overall Perceived Complexity (Range 1–7) and Mean Familiarity (Range 0–1) for Experts and Nonexperts

Artworks	Mean complexity ( <i>SD</i> )		Mean familiarity	
	Nonexpert group	Expert group	Nonexpert group	Expert group
<b>Landscape, still life, and portrait painting</b>				
de' Barbari, Jacopo (1504). <i>Still Life with Partridge and Iron Gloves</i> . Alte Pinakothek, Munich.	3.31 (.99)	3.2 (1.33)	0	.78
Bellini, Giovanni (1515). <i>Naked Young Woman in Front of the Mirror</i> . Kunsthistorisches Museum, Vienna.	3.54 (1.28)	2.3 (.9)	.08	.78
Caravaggio, Michelangelo Merisi da (1599). <i>Basket of Fruit</i> . Pinacoteca Ambrosiana, Milan.	3.85 (1.46)	2.7 (1.19)	0	.78
Bosschaert, Ambrosius (1606). <i>Still Life of Flowers</i> . The Cleveland Museum of Art, Cleveland.	4.31 (1.73)	3.4 (.8)	0	.56
van Goyen, Jan (1630). <i>Flat Landscape</i> . Erich Lessing Culture and Fine Arts Archives, New York.	3.46 (1.08)	2.3 (1.55)	0	.67
Claesz, Pieter (1634). <i>Breakfast Still Life with Goblet</i> . Kunsthistorisches Museum, Vienna.	3.85 (1.29)	3.5 (1.02)	.08	.89
Snyders, Frans (1637). <i>Still Life with Fruit, Vegetables and Dead Game</i> . Detroit Institute of Arts, Detroit.	5.15 (1.41)	5.2 (1.54)	0	.56
Ruisdael, Jacob van (1660). <i>The Great Oak</i> . Metropolitan Museum of Art, New York.	4.15 (1.17)	4.10 (1.14)	0	.67
Zurbarán, Francisco de (1660). <i>Still Life</i> . Prado, Madrid.	2.54 (1.45)	1.9 (.94)	0	.44
de Heem, Jan Davidszoon (1665). <i>Breakfast Still Life</i> . Kunsthistorisches Museum, Vienna.	5.15 (1.1)	4.7 (.78)	.08	.89
Prost, Frans Jansz (1665). <i>Brazilian Landscape</i> . Detroit Institute of Arts, Detroit.	5.54 (1.39)	4 (1.18)	0	.44
van Oosterwyck, Maria (1668). <i>Still Life</i> . Kunsthistorisches Museum, Vienna.	5.62 (1.21)	5.6 (.49)	0	.67
Ruisdael, Jacob van (1670). <i>Wheat Fields</i> . Kunsthistorisches Museum, Vienna.	4.15 (1.29)	3.9 (1.51)	0	.44
Stubbs, George (1762). <i>Foxhound in a Landscape</i> . University of California, Berkeley.	3.46 (1.39)	2.4 (1.43)	0	.44
Reynolds, Joshua (1765). <i>Portrait of a Young Woman</i> . Kunsthistorisches Museum, Vienna.	3.15 (1.35)	2.3 (1.1)	0	.44
Koekkoek, Barend Cornelis (19th century). <i>Winter Landscape</i> . Rijksmuseum, Amsterdam.	4.15 (1.29)	2.9 (1.3)	0	.33
Goya, Francisco de (1812). <i>Still Life with Sheep's Ribs and Head</i> . Erich Lessing Culture and Fine Arts Archives, New York.	4.23 (.97)	2.2 (.87)	0	.67
Friedrich, Caspar David (1825). <i>Der Watzmann</i> . Nationalgalerie, Berlin.	3.77 (1.48)	2.7 (1.0)	0	.78
Moran, Thomas (1884). <i>The Much Resounding Sea</i> . National Gallery of Art, Washington DC.	4.15 (1.46)	2.3 (1.1)	.08	.44
Cézanne, Paul (1882). <i>Still Life with Carafe, Milk Can, Bowl, and Oranges</i> . Dallas Museum of Art, Dallas.	3.69 (1.38)	2.9 (1.14)	0	.56
van Gogh, Vincent (1889). <i>Portrait of a Young Peasant with Straw Hat</i> . Collection E. Sformi, Florence.	3.31 (.91)	1.6 (.49)	.31	1
Monet, Claude (1897). <i>Morning on the Seine near Giverny</i> . Metropolitan Museum of Art, New York.	3.69 (1.32)	1.8 (.6)	.15	.89
Picasso, Pablo (1908). <i>Fruit-bowl with Pears and Apples</i> . Museum Berggruen, Berlin.	2.54 (.63)	1.9 (.7)	.15	.78
Fry, Roger (1914). <i>Winter Landscape</i> . Cornell Fine Arts Museum at Rollings College, Winter Park.	4.31 (1.07)	3.5 (1.2)	0	.56
Schiele, Egon (1914). <i>Reclining Female Nude with Legs Spread Apart</i> . Graphische Sammlung Albertina, Vienna.	3.69 (1.14)	2.4 (1.02)	.31	1
Schiele, Egon (1917). <i>House with Drying Laundry</i> . Leopold Museum, Vienna.	4.46 (.93)	4.4 (1.02)	0	.67
Wesselman, Tom (1963). <i>Still Life #25</i> . University of California, San Diego.	4.23 (1.31)	4 (1.55)	.15	.44
Warhol, Andy (1976). <i>Skull</i> . Baltimore Museum of Art, Baltimore.	3.54 (1.39)	2.6 (1.62)	0	.56
<b>Abstract art</b>				
Legér, Fernand (1912). <i>Woman in Blue</i> . Kunstmuseum, Basel.	5.62 (.74)	5.7 (1.49)	0	.67
Boccioni, Umberto (1913). <i>Dynamism of a Soccer Player</i> . Museum of Modern Art, New York.	5.85 (1.03)	5.6 (.8)	.08	.56
Malevich, Kazimir (1913). <i>The Knife Grinder</i> . Yale University Art Gallery, New Haven.	5.85 (.77)	5.8 (1.47)	.15	.78
Marc, Franz (1914). <i>Kämpfende Formen (Fighting Forms)</i> . Pinakothek der Moderne, Munich.	5.38 (1.21)	4.7 (1.35)	.08	1
Malevich, Kazimir (1923). <i>Black Square</i> . State Russian Museum, St. Petersburg.	1.62 (.92)	1.8 (1.47)	0	.89
Kandinsky, Wassily (1926). <i>Circles</i> . Guggenheim Museum, New York.	5.15 (1.41)	5.2 (1.54)	.23	.67
Mondrian, Piet (1929). <i>Composition in Yellow and Blue</i> . Museum Boijmans Van Beuningen, Rotterdam.	2.15 (1.03)	2 (.89)	.31	1
Kandinsky, Wassily (1935). <i>Accompanied Contrast</i> . Guggenheim Museum, New York.	3.92 (1.54)	3.1 (1.04)	.15	.67
Motherwell, Robert (1946). <i>Western Air</i> . Museum of Modern Art, New York.	3.08 (1.27)	3.5 (1.43)	0	.33
de Kooning, Willem (1950). <i>Painting</i> . The Willem de Kooning Foundation, New York.	5.62 (1.0)	5.2 (1.25)	.08	.33
Pollock, Jackson (1952). <i>Convergence</i> . Albright-Knox Art Gallery, Buffalo.	5.23 (1.67)	4.9 (1.87)	.38	.78
Götz, Karl Otto (1953). <i>Bild vom 5.2.1953</i> . Nationalgalerie, Berlin.	4.46 (1.08)	4 (1.48)	.15	.67
Kline, Franz (1956). <i>Mahoning</i> . Whitney Museum of Art, New York.	3.54 (1.74)	2.5 (.92)	.08	.78

(Appendix continues)



Appendix (continued)

Artworks	Mean complexity (SD)		Mean familiarity	
	Nonexpert group	Expert group	Nonexpert group	Expert group
<b>Narrative-figurative painting</b>				
Giotto di Bondone (1306). <i>The Arrest of Christ (Kiss of Judas)</i> . Scrovegni Chapel, Padua.	4.23 (1.19)	4.8 (1.25)	0	.1
Duccio di Buoninsegna (1311). <i>Last Supper</i> . Opera della Metropolitana di Siena, Siena.	4.38 (1.39)	4.2 (1.08)	0	.56
Lochner, Stephan (first half 15th century). <i>Crucifixion</i> . Germanisches Nationalmuseum, Nuremberg.	3.77 (1.42)	2.8 (1.25)	0	.33
Upper Rhenish Master (1410). <i>Paradiesgärtlein</i> . Städel Museum, Frankfurt am Main.	4.85 (1.23)	5.70 (1.1)	.08	.89
Fra Angelico (1445). <i>Annunciation</i> . Museo di San Marco, Florenz.	2.69 (.99)	2.2 (1.25)	0	.89
Lippi, Fra Filippo Tommaso (1450). <i>Annunciation</i> . Alte Pinakothek, Munich.	4.77 (1.31)	5.5 (.67)	.08	.89
Memling, Hans (1475). <i>The Martyrdom of St. Sebastian</i> . Musées Royaux des Beaux-Arts, Brussels.	5.08 (.92)	4.6 (1.43)	.08	.44
Bosch, Hieronymus (1490). <i>St. John the Baptist in the Wilderness</i> . Museo Lázaro Galdiano, Madrid.	5.38 (.74)	6.2 (.75)	0	.67
Raffael (1507). <i>La Belle Jardinière</i> . Louvre, Paris.	3.54 (1.28)	3 (1.41)	.08	.89
Titian (1515). <i>Sacred and Profane Love</i> . Galleria Borghese, Rome.	4.54 (1.01)	5.2 (.87)	.15	.1
Cranach the Elder, Lucas (1530). <i>Paradise</i> . Kunsthistorisches Museum, Vienna.	5.15 (.66)	4.6 (1.28)	.08	1
Titian (1530). <i>The Madonna of the Rabbit</i> . Louvre, Paris.	4.08 (1.0)	4.4 (1.11)	0	.67
Cranach the Elder, Lucas (1537). <i>The Fall of Man</i> . Kunsthistorisches Museum, Vienna.	3.46 (1.15)	2.8 (.87)	.15	.89
Bronzino, Angnolo (1550). <i>Holy Family with St. Anne and St. John the Baptist</i> . Louvre, Paris.	4.38 (1.27)	4.5 (1.12)	0	1
Aertsen, Pieter (1552). <i>Still Life (Christ with Mary and Martha)</i> . Kunsthistorisches Museum, Vienna.	6.15 (.86)	6.5 (.81)	.08	.67
Bruegel the Elder, Pieter (1567). <i>Conversion of Paul</i> . Kunsthistorisches Museum, Vienna.	5.77 (.8)	5.9 (.94)	0	.78
Veronese, Paolo (1570). <i>Venus and Mars</i> . Metropolitan Museum, New York.	5.15 (1.03)	5.4 (1.2)	0	.67
El Greco (1580). <i>St. Veronica</i> . Museo de Santa Cruz, Toledo.	4.15 (1.29)	3.4 (1.28)	0	.56
Tintoretto, Jacopo (1592). <i>Last Supper</i> . S. Giorgio Maggiore, Venice.	6.23 (.58)	5.5 (1.69)	.15	.89
Mitsuyoshi, Tosa (17th century). <i>The Tale of Ganji Monogatari</i> . Seattle Art Museum, Seattle.	5 (1.24)	5.1 (1.3)	0	.56
van Balen I, Hendrick (17th century). <i>The Rape of Europa</i> . Musée des Beaux-Arts de Valenciennes, Valenciennes.	5.23 (1.12)	5 (.77)	0	.78
Caravaggio, Michelangelo Merisi da (1601). <i>Conversion of Paul</i> . Santa Maria del Popolo, Rome.	5 (1.18)	5 (.45)	0	.89
Caravaggio, Michelangelo Merisi da (1604). <i>The Crowning with Thorns</i> . Kunsthistorisches Museum, Vienna.	4.46 (1.39)	4.2 (.87)	0	.89
Gentileschi, Orazio (1607). <i>Christ Carrying the Cross</i> . Kunsthistorisches Museum, Vienna.	3.85 (1.1)	4 (1.1)	0	.89
Rubens, Peter Paul (1618). <i>The Head of Medusa</i> . Kunsthistorisches Museum, Vienna.	5.38 (1.27)	4.2 (1.54)	.15	.78
Velázquez, Diego (1618). <i>Old Woman Cooking Eggs</i> . National Gallery Scotland, Edingburgh.	4.23 (1.53)	4.9 (.3)	0	.78
Rombouts, Theodoor (1630). <i>Card Players</i> . Koninklijk Museum voor Schone Kunsten, Antwerp.	4.08 (1.44)	4.6 (1.11)	.15	.56
Rembrandt van Rijn (1632). <i>The Anatomy Lesson of Dr. Nicolaes Tulp</i> . Mauritshuis, The Hage.	4.77 (1.25)	3.8 (1.33)	.31	1
Crespi, Giuseppe Maria (1700). <i>Centaur Chiron teaches young Achill archery</i> . Kunsthistorisches Museum, Vienna.	5.08 (.83)	4 (1.34)	0	.67
Bianchi, Pietro (1734). <i>Pyramus and Thisbe</i> . Private Collection, Florence.	5 (1.24)	4.4 (1.02)	0	.33
Chardin, Jean Siméon (1737). <i>The House of Cards</i> . National Gallery, London.	2.92 (.83)	2.9 (1.04)	0	1
Greuze, Jean-Baptiste (1756). <i>Broken Eggs</i> . Metropolitan Museum, New York.	4.62 (1.39)	4.8 (1.17)	0	.56
Vien, Joseph-Marie (1767). <i>St. Denis preaching in Gaul</i> . Eglise Saint-Roch, Paris.	4.85 (1.17)	5.4 (.8)	.08	.67
David, Jacques-Louis (1789). <i>The Loves of Paris and Helen</i> . Louvre, Paris.	4.31 (1.43)	3.9 (.83)	0	.78
Boucher, François (1742). <i>Leda and the Swan</i> . National Museum, Stockholm.	4.23 (.97)	3.7 (.9)	.08	.67
Hokushu, Shunkōsai (1832). <i>Nakamura Utaemon III as Ishikawa Goemon</i> . Fine Arts Museum of San Francisco, San Francisco.	4.23 (1.42)	4 (1.67)	0	.67
Turner, William (1834). <i>The Burning of the Houses of Parliament</i> . Tate Britain, London.	5.31 (1.38)	5.4 (1.28)	.08	.44
Delacroix, Eugène (1837). <i>The Battle of Taillebourg</i> . Musée hiytorique (Versailles), Paris.	6.54 (.63)	5.9 (1.22)	.08	.67
Hokusai Katsushika (1868). <i>The Lake of Hakone in Sagami Province</i> . Tokyo National University of Fine Arts and Music.	2.38 (.62)	2.1 (.7)	0	.89
Hōgai, Kanō (1885). <i>Two Dragons (in Clouds)</i> . Philadelphia Museum of Art, Philadelphia.	4.38 (1.15)	3.1 (1.58)	0	.67
Corinth, Lovis (1899). <i>Salome with the Head of John the Baptist</i> . Museum der bildenen Künste, Leipzig.	5.69 (.99)	5.3 (1.49)	0	.78
Klimt, Gustav (not dated). <i>Fairy Tale</i> . Wien Museum, Vienna.	5.15 (.53)	4.8 (2.04)	0	.44
Hunt, William Holman (1905). <i>Lady of Shalott</i> . Wadsworth Atheneum, Hartford.	6.23 (.89)	6.4 (.8)	0	.78
Nolde, Emil (1910). <i>Dancer</i> . Kunsthalle, Emden.	3.38 (1.08)	1.7 (.64)	.08	.78
Klimt, Gustav (1915). <i>Death and Life</i> . Leopold Museum, Vienna.	5.62 (.74)	4.7 (1.68)	.69	1
Ernst, Max (1929). <i>Et les femmes volcaniques relévent et agitent, d' un air menacant</i> Stiftung Sammlung Diete, Berlin.	6.31 (.99)	6.3 (1.19)	0	.56
Ernst, Max (1945). <i>The Temptation of St. Anthony</i> . Wilhelm-Lehmbruck Museum, Duisburg.	6.23 (1.37)	5.9 (1.76)	0	.67
Rauch, Neo (2002). <i>Späher</i> . Private Collection, Osnabrück.	4.15 (1.41)	2.8 (1.66)	0	.33
Lassnig, Maria (2005). <i>Du oder ich</i> . Friedrich Christian Flick Collection.	3.62 (1.27)	2.4 (1.02)	.23	.89

Received October 11, 2016  
 Revision received February 6, 2018  
 Accepted April 19, 2018 ■

---

## *Diskussion*

Ziel der Studie war es, Komplexität als Begriff genauer zu fassen. Die Argumentation der Studie ist dabei darauf bedacht, die Psychologie der menschlichen Wahrnehmung zu berücksichtigen. Folglich handelt es sich um den Versuch, einen Brückenschlag zwischen psychologischer Ästhetikforschung und kunsttheoretischen Positionen zu leisten. Jene psychologische Ästhetikforschung, deren Fokus auf der Beschreibung und Erklärung von unterschiedlichen Aspekten der ästhetischen Erfahrung bei der Auseinandersetzung mit Kunstwerken liegt, bildet hierbei die Basis der Argumentation. Diese Verschränkung zweier sehr spezifischer Disziplinen ist nicht nur aufgrund der Natur des Fragenraums nachvollziehbar, sondern auch aufgrund des Forschungsstandes in den jeweiligen Fächern. Während im Kontext kunsttheoretischer Überlegungen das Konzept der Komplexität bisher kaum eine Rolle gespielt hat, sind im Rahmen der psychologischen Ästhetikforschung hierzu bereits umfassende Ergebnisse vorgelegt worden. Wie sich bei der Auseinandersetzung mit diesen Ergebnissen allerdings zeigt, sind die hier angebotenen Definitionen für den Kontext kunstwissenschaftlicher Argumentationsräume von geringem Nutzen.

Die Einschränkung des Komplexitätskonzeptes in der psychologischen Ästhetikforschung auf formale Bildaspekte sowie deren kognitive Verarbeitung, ist, wie schon aus dem Verwendungszusammenhang des Begriffes deutlich wird, für die Kunstwissenschaften problematisch. Darüber hinaus widerspricht eine solche Eingrenzung oder Reduktion einigen grundlegenden Prinzipien kunsttheoretischer Überlegungen, wie zum Beispiel einem holistisch verstandenen Bildkonzept, dass Form, Inhalt und Betrachter in ein unauflösbares Wechselverhältnis setzt. Ein solches, semiotisches Bildkonzept liegt vielen kunstwissenschaftlichen und kunsttheoretischen Argumentationskontexten zugrunde. In der vorgestellten Studie bildet die psychologische Perspektive den Ausgangspunkt, welcher mit holistischeren Bildkonzepten aus der Kunsttheorie integriert wird. Im Konkreten werden dazu die bisherigen Standards der psychologischen Ästhetikforschung um die von Erwin Panofsky benannten semantischen Bedeutungsschichten von Kunstwerken angereichert.



---

Panofsky eignet sich für einen Brückenschlag zwischen psychologischer Ästhetikforschung und Kunsttheorie, weil seine Konzepte einerseits bestehenden theoretischen Überlegungen der Psychologie zugeordnet werden können und weil Panofskys Annahmen andererseits einem semiotischen Bildverständnis entspringen.<sup>127</sup>

In ihrer Gesamtheit weisen die Ergebnisse der Studie darauf hin, dass eine Definition von Komplexität nach semiotischen Kriterien zutreffend ist: Die Wahrnehmung von Komplexität in Kunstwerken ist wesentlich von sowohl formalen, inhaltlichen und persönlichen Kriterien abhängig. Komplexität liegt somit weder ausschließlich im Bild, noch ausschließlich in den Betrachtenden verankert. Der Eindruck von Komplexität entsteht vielmehr erst im Wechselspiel, in der Interaktion, von Betrachter\_in und Kunstobjekt. Hierbei sind sowohl formale Bildkriterien von Relevanz, etwa die Art und Übersichtlichkeit, in welcher die visuellen Elemente in der Bildfläche angeordnet sind, wie auch ikonographische und ikonologische Überlegungen zum Inhalt der Darstellung. Darüber hinaus hat sich gezeigt, dass auch persönliche Kriterien, wie z.B. Expertise, einen Einfluss auf den Komplexitätseindruck der Betrachtenden nehmen. Komplexität soll und muss also holistischer gefasst werden, als bisher in der psychologischen Ästhetikforschung angenommen.

Um die detaillierte Diskussion dieser Studienergebnisse übersichtlich zu gliedern, ist der nun folgende Abschnitt in drei Blöcke unterteilt. In diesen werden die Studienergebnisse schwerpunktmäßig hinsichtlich der folgenden Aspekte diskutiert: a) die grundsätzlichen Determinanten der Komplexitätserfahrung, b) die durch Expertise motivierten interpersonellen *Unterschiede* in der Komplexitätserfahrung und c) die durch Expertise motivierten interpersonellen *Gemeinsamkeiten* in der Komplexitätserfahrung. Hierbei werde ich insbesondere auf den letzten Punkt in Tiefe eingehen, da die Spezifika der

---

<sup>127</sup> Vgl. z. B. Bal/Bryson 1991, Hasenmueller 1978.

---

Expert\_innenerfahrung für die Theorie und Praxis der Kunstwissenschaft eines der relevantesten Studienergebnisse darstellen.

*a. Grundsätzliche Determinanten der Komplexitätserfahrung*

Dem Gesamtergebnis der Studie ist zu entnehmen, dass Komplexität zwar einerseits eine Bewertung, also Ausdruck eines subjektiven Zugangs zum Bild ist; andererseits steht diese Bewertung aber in direkter Relation zum Gegenstand und seinen *formalen* Eigenschaften. Diese formalen Eigenschaften beeinflussen die Bewertung von Komplexität bei allen Versuchsteilnehmenden *gleichermaßen* bzw. auf die gleiche Art und Weise. Das zeigt sich etwa daran, dass visuelle Komplexität in der Stärke ihres Einflusses auf den gesamten Komplexitätseindruck kaum Varianz zwischen den befragten Personen und den jeweils von ihnen bewerteten Bildbeispielen aufweist. Die Untersuchung legt sogar nahe, dass Menschen – ganz gleich ob Kunstlaien oder Kunstexpert\_innen – in ihrem Komplexitätseindruck immer in der gleichen Weise von formalen Kriterien beeinflusst werden. Eine als unübersichtlich empfundene Anordnung von visuellen Elementen auf der Bildfläche führt also unabhängig von persönlichen Merkmalen, wie etwa Expertise, dazu, dass ein bestimmtes Werk tendenziell eher als komplex bewertet wird.<sup>128</sup> Ein Beispiel hierfür wäre etwa die durchschnittliche *formale* Komplexitätseinschätzung von Max Ernsts *Die Versuchung des Hl. Antonius* (1945). Laien haben für das Werk auf einer 7er-Skala einen durchschnittlichen Komplexitätswert von 5,27 vergeben, Expert\_innen haben dasselbe Werk durchschnittlich mit 5,36 bewertet.<sup>129</sup>

---

<sup>128</sup> Formale Komplexität umfasst, wie zuvor demonstriert, eine ganze Reihe formaler Bildeigenschaften. Da diese bereits hinreichend gut durch die psychologische Ästhetikforschung beleuchtet wurden, werde ich im Folgenden auf die einzelnen Aspekte formaler Komplexität nicht im Detail eingehen. Für eine Übersicht der diesbezüglichen Ergebnisse verweise ich auf Nadal 2007.

<sup>129</sup> Die an dieser Stelle genannten Werte zur durchschnittlichen Bewertung der *formalen* Komplexität sind in der Studie nicht explizit angeführt. Zwischen formaler und semantischer Komplexität wird hier nicht explizit unterscheiden. Eine isolierte Betrachtung dieser beiden Konzepte ist, wie der bisherigen Argumentation zu entnehmen ist, nicht unbedingt sinnvoll. Zur besseren Veranschaulichung der Ergebnisse führe ich die Werte an dieser Stelle dennoch an. Sie speisen sich aus demselben Datenkorpus, der auch der Studie zugrunde liegt. Sie sollten über die hier verfolgten Veranschaulichungszwecke hinaus nicht als für sich stehend verstanden werden.



Abbildung 1 | Ernst, Max (1945). *Die Versuchung des Hl. Antonius* (108 x 128 cm), Foto: Artstor, © Wilhelm-Lehmbruck Museum, Duisburg.

Robert Motherwells *Western Air* (1946), dass im Vergleich zu Max Ernsts *Die Versuchung des Hl. Antonius* einen übersichtlicher gestalteten Bildraum aufweist (vgl. Abbildung 1 und 2), hat von sowohl Laien als auch Expert\_innen einen niedrigeren Durchschnittswert hinsichtlich formaler Komplexität erhalten. Auch liegen die Werte nahe beieinander, ähnlich wie es auch bei Max Ernst der Fall war: Laien haben für *Western Air* eine durchschnittliche formale Komplexität von 3,89 vergeben und Expert\_innen haben das Werk im Durchschnitt mit 3,97 bewertet.



Abbildung 2 | Motherwell, Robert (1946). *Western Air* (182,9 x 137,2 cm), Foto: Artstor, © Museum of Modern Art, New York.

Von einer Objektivierbarkeit des Komplexitätskonzeptes zu sprechen, ist allerdings trotz der starken intersubjektiven Übereinstimmung des formalen Komplexitätseindrucks nur bedingt möglich. Zwar gibt es hinsichtlich der Einflüsse von formalen Bildaspekten auf die Komplexitätserfahrung kaum Varianz zwischen den Personen und den Bildbeispielen, hinsichtlich der *inhaltlichen* Aspekte und ihrem Beitrag zum Komplexitätserleben sind jedoch deutliche Abweichungen zwischen den Personen und auch zwischen den Bildbeispielen zu verzeichnen. Da die Ergebnisse der Studie nahelegen, dass sowohl formale als auch inhaltliche Kriterien gleichermaßen zum Entstehen eines Komplexitätseindrucks beitragen, lässt sich die augenscheinliche Varianz im Einfluss von inhaltlicher Auseinandersetzung auf das Erleben von Komplexität nicht einfach von der Hand weisen.

---

b. *Unterschiede zwischen dem Komplexitätseindruck von Laien und Expert\_innen*

Die Erfahrung von Komplexität ist je nach persönlichen Voraussetzungen der Betrachtenden von anderen Kriterien abhängig. Das zeigt sich etwa daran, dass Kunstlaien und Kunstexpert\_innen von unterschiedlichen inhaltlichen Bildaspekten in ihrer Komplexitätserfahrung beeinflusst werden: Während ein gesteigerter Komplexitätseindruck bei Kunstlaien im Wesentlichen auf eine ikonografische Auseinandersetzung mit dem Werk zurückzuführen ist, ist das Erleben von Komplexität bei Kunstexpert\_innen von ikonologischen Überlegungen abhängig. Dieser Unterschied lässt sich auf grundlegende theoretische Annahmen zurückführen, die der Wahrnehmung von semantischer Komplexität zugrunde gelegt werden können. Semantische Komplexität, wie in der Studie definiert wurde, speist sich im Wesentlichen aus zwei Kanälen: einerseits der Leichtigkeit, mit welcher semantische Inhalte von den Betrachtenden erfasst und sinnstiftend in der Auseinandersetzung mit dem Werk (weiter-)verarbeitet werden können; und andererseits der im Werk wahrgenommenen semantischen Tiefe (wobei sich ‚semantische Tiefe‘ hier auf die Vielschichtigkeit der an einem Werk wahrgenommene Bedeutung bezieht). Beide Aspekte variieren dabei stark mit dem personenbezogenen Merkmal Expertise. Die Leichtigkeit, mit welcher bestimmte Inhalte erfasst und weiterverarbeitet werden können, ist wesentlich vom Wissens-, Kenntnis- und Erfahrungsstand der Betrachtenden abhängig; und auch die Erfahrung semantischer Tiefe steht in Relation zu eben diesen Aspekten. Expert\_innen, die in der Regel in der Lage sind in jedem Werk in der ein oder anderen Form semantische Inhalte oder Aspekte zu erkennen, haben eine andere Bewertungsgrundlage als Laien, gegen welche sie abwägen, ob die in einem Werk wahrgenommene semantische Struktur Tiefe besitzt oder nicht. Es ist anzunehmen, dass diese Bewertungsgrundlage im Falle von Expert\_innen deutlich voraussetzungsreicher ist als es für Laien der Fall sein sollte. Laien können in vielen Werken – aufgrund des ihnen *fehlenden* Wissens- und Erfahrungshintergrunds – häufig überhaupt keine sinnstiftende inhaltliche Interpretation vollziehen. Im Fall von Laien kann also schon die schlichte Tatsache, dass einem bestimmten Werk *überhaupt* ein bestimmter Inhalt zugeordnet werden kann, dazu führen, dass die semantische Struktur dieses Werkes

---

im Vergleich zu anderen Werken als „tief“ erlebt wird. Führend wir dieses Argument rück, zu den inhaltlichen Ebenen die einem Werk zugeschrieben werden können, indem wir uns auf Panofskys Klassifikationsschema beziehen, so sind die Inhalte, die von Laien in der Regel erfasst werden können, hauptsächlich vorikonografischer Natur. Das Erkennen einer ikonografischen inhaltlichen Ebene stellt hingegen bereits eine tiefere Auseinandersetzung mit der semantischen Struktur des Werkes dar. Im Fall von Expert\_innen führt das Erkennen der ikonografischen Inhalte eines Werkes in den meisten Fällen noch *nicht* zu einer Komplexitätserfahrung, da ikonografische Inhalte von Expert\_innen in der Regel mühelos dechiffriert werden können. Hier erfordert eine Erfahrung von semantischer Tiefe demnach eine weitergehende inhaltliche Auseinandersetzung, wie sie etwa dann gegeben ist, wenn ikonologische Aspekte am Werk verhandelt werden. Trotz dieser Unterschiede, die für die subjektive Bewertung und Erfahrung eines Werkes gravierend sein können, lässt sich festhalten, dass den jeweiligen subjektiven Erfahrungen derselbe kognitive Mechanismus zugrunde liegt. Sowohl Laien als auch Expert\_innen unterliegen in ihrem Komplexitätsurteil der Erfahrung von semantischer Tiefe – es ist nur schlicht so, dass die Wahrnehmung semantischer Tiefe nicht unabhängig vom Expertisegrad des betrachtenden Subjektes ist.

Obwohl, wie aufgezeigt, in beiden Fällen derselbe Ausgangspunkt zugrunde liegt, nämlich die Frage nach der Erfahrung von inhaltlicher Tiefe, ist der hier aufscheinende Unterschied zwischen den Gruppen nicht trivial. Das liegt insbesondere daran, dass in der tatsächlichen Praxis menschlicher Wahrnehmung die Verarbeitung von ikonografischen und ikonologischen Bildinhalten anderen *Zeitskalen* unterworfen ist. Ikonografische Inhalte *können* quasi-automatisch erkannt werden, eine Auseinandersetzung mit ikonologischen Aspekten oder Fragen *muss* bewusst erfolgen.<sup>130</sup> Damit ist deutlich, dass die *Komplexitätsbewertung* in Abhängigkeit von den persönlichen Voraussetzungen der Betrachtenden unterschiedlichen Parametern unterworfen ist. Komplexität ist für Kunstlaien und Kunstexpert\_innen im Groben also etwas sehr Ähnliches, im Detail aber dennoch

---

<sup>130</sup> Vgl. Leder et al. 2004, Leder/Nadal 2014.

---

deutliche Verschiedenes: während sich der Komplexitätseindruck für Kunstlaien direkt aus der unmittelbaren Erfahrung von einem Kunstgegenstand ableitet, inkludiert er bei Kunstexpert\_innen die bewusste Auseinandersetzung mit dem Werk.

Diese Nicht-Übereinstimmung von Kunstlaien und Kunstexpert\_innen in ihrer Komplexitätswahrnehmung, wird durch einen weiteren Aspekt der menschlichen Wahrnehmung und kognitiven Verarbeitung weiter verstärkt. Die Ergebnisse der vorgestellten Studie legen nahe, dass Expert\_innen im Durchschnitt dazu neigen, Kunstwerke *insgesamt* als weniger komplex zu bewerten als Laien. Dieser Zusammenhang ist im Fließtext der Studie statistisch aufbereitet. Eine diesbezügliche Differenzierung auf Bildebene ist den im Anhang der Studie angefügten Tabellen zu entnehmen. Um eine bessere Lesbarkeit dieser Maßzahlen zu gewährleisten sind sie im Folgenden grafisch aufbereitet. Vergleichen wir die durchschnittlichen Komplexitätsbewertungen, die von Laien und Expert\_innen für die 90 Bilder der Studie abgegeben wurden, so wird deutlich, dass die beiden Gruppen unterschiedliche Verteilungen aufweisen (Abbildung 3). Während die Komplexitätseinschätzungen der Expert\_innen eine verhältnismäßig weite Spannbreite aufweisen und damit jeweils ähnlich viele Werke als wenig-, mittel- und hoch-komplex einstufen, zeigte sich, dass Laien eher dazu neigen, die präsentierten Werke als mittel- oder hoch-komplex zu bewerten. Der Anteil an Werken, die von Laien als kaum- oder nur wenig-komplex eingestuft wurden, ist im Vergleich zu diesen beiden anderen Kategorien deutlich geringer.

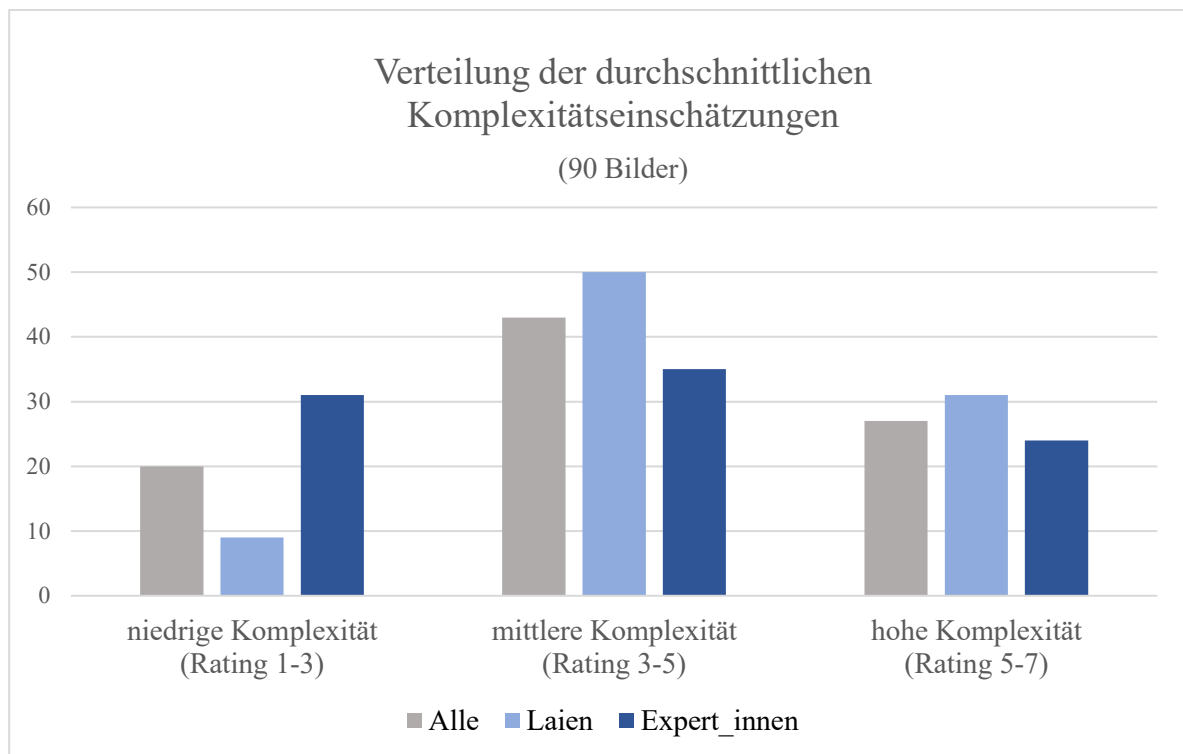


Abbildung 3 | Verteilung der durchschnittlichen Komplexitätseinschätzungen, eigene Darstellung.

Betrachten wir diese Verteilung nicht geblockt, sondern auf Bildebene (Abbildung 4), so wird deutlich, dass die Abweichungen in den Komplexitätseinschätzungen von Laien und Expert\_innen eine gewisse Struktur aufweisen. In der Visualisierung sind die Bilder von links nach rechts nach den durchschnittlich von *allen* Befragten für sie vergebenen Komplexitätswerten geordnet (vgl. grauer Balken in Abbildung 3). Jedes Ausschlagen der abgebildeten Kurve stellt jeweils eines der untersuchten Kunstwerke dar. Wie man der Grafik entnehmen kann, sind große Abweichungen in den Einschätzungen von Laien und Expert\_innen insbesondere für durchschnittlich als wenig-komplex eingeschätzte Werke zu verzeichnen. Die durchschnittlichen Komplexitätseinschätzungen beider Gruppen sind sich ab einem mittleren Komplexitätsniveau deutlich ähnlicher und bleiben dies auch für hochkomplexe Kunstwerke. Auffällig ist auch, dass die Kurven der durchschnittlichen Komplexitätseinschätzungen von Laien und Expert\_innen stark schwanken: bei manchen Werken ist der Abstand zwischen den jeweiligen Komplexitätseinschätzungen der Gruppen deutlich größer als bei anderen. Auch ist es keinesfalls so, dass die Bewertungen von Laien



im niedrigen Komplexitätsbereich *grundsätzlich* über den Bewertungen der Expert\_innen liegen würden: es gibt durchaus auch Werke die von Laien im Durchschnitt als weniger komplex eingeschätzt werden als von Expert\_innen. Dies ist insbesondere für Werke der Fall, die im Gesamtdurchschnitt als mittel bis hochkomplex wahrgenommen werden.

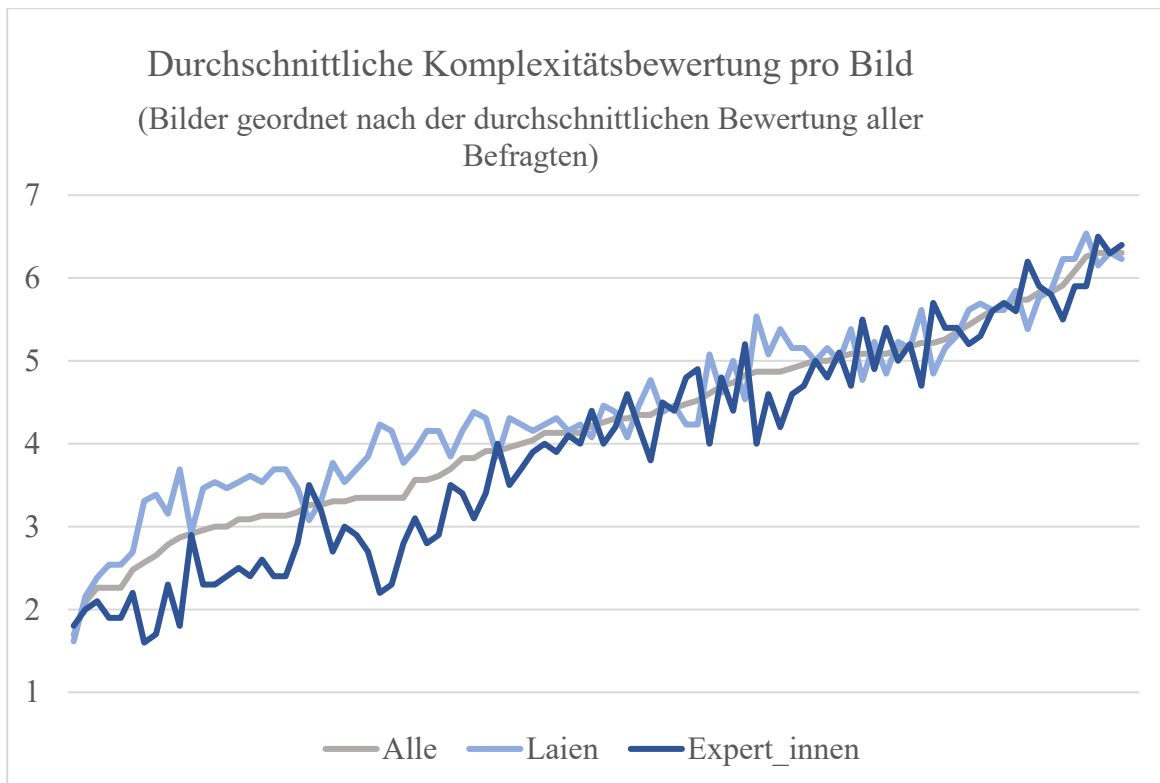


Abbildung 4 | Durchschnittliche Komplexitätsbewertung pro Bild, eigene Darstellung.

Während im Durchschnitt im höheren Komplexitätsspektrum mehr Einigkeit besteht, stellt diese Tendenz dennoch keine in sich kohärente Entwicklung dar, sondern weist – ganz im Gegenteil – immer wieder auch „Ausreißer“ an beiden Enden des Komplexitätsspektrums auf. Betrachten wir die einzelnen Werke, bei denen es zu besonders großen Abweichungen (über 1,5 Ratingpunkte auf der zugrundeliegenden 7er Bewertungsskala) zwischen Laien und Expert\_innen Urteilen kommt, im Detail, so wird deutlich, dass diese Abweichungen zwei Kategorien zuordenbar sind: a) Werke, die im Vergleich zu Laien von Expert\_innen als *weniger komplex* eingeschätzt werden (wie dies auch der generelle Durchschnitt nahelegt)

und b) Werke, die von Expert\_innen im Vergleich zu Laien als *komplexer* eingeschätzt werden.



Abbildung 5 | Werkbeispiele mit großen Abweichungen (über 1,5 Ratingpunkte) in den durchschnittlichen Komplexitäts-einschätzungen von Laien und Expert\_innen im niedrigen bis mittleren Komplexitätsspektrum (oben von links: Monet, Claude (1897). *Morning on the Seine near Giverny* (89.9 x 92.7 cm), Foto: Artstor, © Metropolitan Museum of Art, New York; van Gogh, Vincent (1889). *Portrait eines jungen Bauerns mit Hut* (65 cm × 54 cm), Foto: Artstor, © Collection E. Sforzi, Florenz; unten: Moran, Thomas (1884). *The Much Resounding Sea* (63.9 x 158.2 cm), Foto: Artstor, © National Gallery of Art, Washington DC.).

Bei den Werken, die der ersten Abweichungskategorie zugeordnet werden können, handelt es sich überwiegend um Porträt- und Landschaftsdarstellungen. Diese werden von

---

Expert\_innen in der Regel als nicht besonders komplex eingeschätzt, von Laien hingegen eher im mittleren Komplexitätsspektrum angesiedelt.

Beispiele hierfür sind im niedrigen Komplexitätsspektrum van Goghs *Portrait eines jungen Bauerns mit Hut* (1889), Thomas Morans *The Much Resounding Sea* (1884) und Claude Monets *Morning on the Seine near Giverny* (1987) (siehe Abbildung 5). Ein diesbezügliches Beispiel aus dem höheren Komplexitätsspektrum wäre etwa Frans Jansz Prosts *Brasilianische Landschaft* (1665) (Abbildung 6).



**Durchschnittliche  
Komplexitätseinschätzung**  
Laien 5.54 | Expert\_innen 4

Abbildung 6 | Werkbeispiel mit großen Abweichungen (über 1,5 Ratingpunkte) in den durchschnittlichen Komplexitätseinschätzungen von Laien und Expert\_innen im hohen Komplexitätsspektrum (Prost, Frans Jansz (1665). *Brasilianische Landschaft* (56.2 × 83.5 cm), Foto: Artstor, © Detroit Institute of Arts, Detroit.).

Die Werke, die der zweiten Kategorie von Abweichungen zwischen Expert\_innen und Laien zuzuordnen sind, welche also von Expert\_innen als komplexer eingeschätzt wurden als von Laien, sind überwiegend entweder figürlich-narrativer Natur oder aber es handelt sich um

---

abstrakte Werke. Beispiele hierfür sind etwa die in Abbildung 7 gezeigten Werke Tizians (*Himmlische und irdische Liebe*, 1514), Giuseppe Maria Crespis (*Kentaur Chiron bringt dem jungen Achilles das Bogenschiessen bei*, 1700) und Hans Memlings (*Martyrium des Hl. Sebastian*, 1475) oder auch die beiden abstrakten Arbeiten Franz Klines (*Mahoning*, 1956) und Franz Marcs (*Kämpfende Formen*, 1914), die in Abbildung 8 zu sehen sind.

Auf den ersten Blick ist nicht eindeutig, was das verbindende Element hinter all diesen beobachtbaren Abweichungen in Expert\_innen- und Laienbewertungen darstellt. Bei näherer Betrachtung ist vor dem Diskurs um semantische Komplexität jedoch auffällig, dass es sich hierbei in allen Fällen um Werke handelt, deren semantische Bedeutung sich entweder sehr leicht erschließen lässt (Landschaft; Portrait), oder aber es handelt sich um Werke die in dieser Hinsicht große Voraussetzungen an das Wissen und die Vorkenntnis der Betrachtenden stellen (figürlich-narrative Themen; abstrakte Kunst). Dabei werden diejenigen Werke, deren semantische Aspekte leichter zugänglich sind, von Laien im Vergleich zu Expert\_innen als deutlich komplexer eingeschätzt. Werke, deren semantischer Gehalt sich hingegen nicht ohne bestehendes Vorwissen erschließt, weisen ein umgekehrtes Verhältnis auf: hier liegt die Expert\_innen-Bewertung deutlich über der Einschätzung der Laien. Dieses Muster lässt sich anhand des Theorierahmens der Studie leicht erklären, da hier der Einfluss von Wahrnehmung semantischer Tiefer auf das Komplexitätsempfinden ein zentrales Moment darstellt.

Bei den Werken, die von Expert\_innen im Vergleich zum Empfinden der Laien als komplexer eingeschätzt werden, handelt es sich ausnahmslos um Arbeiten, deren motivische Dimension eines bestimmten Vorwissens bedarf, um sinnstiftend erschlossen werden zu können. Dieses Vorwissen ist dabei in allen Fällen eines, dass für Laien eher ungewöhnlich ist, womit gemeint ist, dass hier keine Themen verhandelt werden, die aufgrund des (christlichen) Kulturgutes der untersuchten westlichen Betrachtenden als grundsätzlich bekannt vorausgesetzt werden können. Das Erkennen und Dechiffrieren der hier beispielhaft genannten figürlich-narrativen Arbeiten setzt sehr spezifische Kenntnisse über bestimmte



kulturhistorisch zwar relevante, zeitgenössisch aber wenig rezipierte Erzählungen und Darstellungskonventionen voraus. Es handelt sich dabei um ein Wissen, das Expert\_innen nach einem Studium, z. B. der Kunstgeschichte, selbstverständlich haben. Laien hingegen besitzen ein solches Wissen in der Regel nicht, da die verhandelten Themen für das Zeitgeschehen von geringer Relevanz und deswegen nicht Teil ihrer Lebens- und Erfahrungswelt sind.



Abbildung 7 | Werkbeispiele mit großen Abweichungen in den durchschnittlichen Komplexitätseinschätzungen von Laien und Expert\_innen (oben: Tizian (1514). *Himmlische und irdische Liebe* (118 × 279 cm), Foto: Artstor, © Galleria Borghese; unten von links: Crespi, Giuseppe Maria (1700). *Kentaur Chiron bringt dem jungen Achilles das Bogenschiessen bei* (125 × 123cm), Foto: Artstor, © Kunsthistorisches Museum Wien; Memling, Hans (1475). *Martyrium des Hl. Sebastian* (67,4 x 67,7 cm), Foto: Artstor, © Musées Royaux des Beaux-Arts, Brüssel.).

---

Ähnliches gilt auch für die abstrakten Arbeiten, bei denen große Abweichungen in der Komplexitätsbewertung von Laien und Expert\_innen zu verzeichnen sind (vgl. Abbildung 8). Auch hier ist es für Laien aufgrund des fehlenden Vorwissens in der Regel nicht möglich eine semantische Bedeutungsstruktur zu erkennen oder zu erarbeiten. Was für den Fall von figürlich-narrativen Arbeiten gilt, können wir für den Fall abstrakter Werke in potenziierter Form annehmen: hier ist nicht nur die Problematik gegeben, dass „ikonografische“ Bedeutungsaspekte von Laien nicht erkannt und sinnstiftend bedacht werden können, sondern es fehlt überdies die Möglichkeit, sich auf einer vorkonografischen Ebene einen Überblick über die wesentlichen Elemente der Darstellung zu verschaffen. Während etwa bei der Betrachtung von Tizians *Himmlische und irdische Liebe* auch ohne Vorwissen deutlich wird, dass hier zwei Frauen und ein Putto im Freien an einem Brunnenbecken sitzen, ist bei der Betrachtung von Franz Klines *Mahoning* eine vergleichbare Orientierung hinsichtlich der grundlegenden bedeutungstragenden Bildaspekte nicht möglich. *Wo* und vor allem *wie* sich der semantische Gehalt eines solchen Werkes erschließen lässt, leitet sich nicht unmittelbar aus der sichtbaren Oberfläche des Werkes ab.



Abbildung 8 | Abstrakte Werkbeispiele mit großen Abweichungen in den durchschnittlichen Komplexitätseinschätzungen von Laien und Expert\_innen (von links: Kline, Franz (1956). *Mahoning* (204.2 × 255.3 cm), Foto: Artstor, © Whitney Museum of Art, New York; Marc, Franz (1914). *Kämpfende Formen* (91 x 131,5 cm), Foto: Artstor, © Pinakothek der Moderne, München.).

In beiden Fällen, dem abstrakten Werk wie auch dem Wissen voraussetzenden figürlich-narrativen Motiv, ist die semantische Komponente der attestierbaren Komplexität lediglich

---

Expert\_innen zugänglich, nicht aber Laien, denen die Grundvoraussetzungen fehlen hier semantische Tiefe erkennen zu können. Da die formalen und die semantischen Aspekte eines Werkes aber in Kombination bedingen, ob ein Kunstwerk subjektiv als komplex erfahren wird, ist es somit nicht weiter verwunderlich, dass gerade hinsichtlich dieser Werkbeispiele größere Differenzen in der Komplexitätsbewertung von Laien und Expert\_innen zu verzeichnen sind.

Für diejenigen Kunstwerke, bei denen die Expert\_innen Komplexitätsbewertung deutlich *unter* der Einschätzung der Laien angesiedelt ist, also etwa die zuvor beispielhaft genannten *Portrait- oder Landschaftsdarstellungen*, gelten andere Rahmenbedingungen: hier können Laien in der Regel mühelos semantischen Gehalt erkennen bzw. attestieren. Die Werke erscheinen daher subjektiv bedeutungsvoller, was sich im Umkehrschluss auf die Komplexitätserfahrung auswirken kann. Bei Expert\_innen hingegen, die für die Bewertung von semantischen Bildaspekten auf einen weitschweifigeren Referenzrahmen zugreifen können, stellt sich bei der Betrachtung von Landschafts- und Portraitdarstellungen in der Regel kein Empfinden semantischer Tiefe ein, da der Gegenstand der Darstellungen im Vergleich zum Möglichkeitsraum häufig als trivial erlebt wird.

Aus den in diesem Block dargestellten Ergebnissen lässt sich schließen, dass Expert\_innen und Laien ein unterschiedliches Komplexitätsempfinden bei der Betrachtung von Kunstwerken aufweisen – trotz der in Abschnitt (a.) dargestellten grundsätzlichen Determinanten, die den Gesamteindruck in beiden Gruppen mitbestimmen. Damit ist deutlich, dass das Urteil *komplex* keine objektiv-verbindliche Aussage darstellt. Dieses Ergebnis ist zum einen interessant, weil es eine praktische Relevanz hat, etwa für den Kontext der Museumspädagogik oder der Kunstvermittlung im Allgemeinen. Die Berücksichtigung der unterschiedlichen Bewertungen und Zugänge von Laien und Expert\_innen zum Werk können in diesem Rahmen zur Evaluierung und Verbesserung des Bildungsangebots herangezogen werden. Für den kunstwissenschaftlichen Kontext ist es allerdings viel wichtiger, dass dieses Ergebnis etwas Entscheidendes über das Komplexitätskonzept verrät,

---

nämlich, dass es sich tatsächlich um eine *Bewertung* handelt, die eine grundsätzliche Objektivierung *nicht* erlaubt. Komplexität ist also keine bildimmanente Eigenschaft, die durch Beobachtung erkannt und durch Beschreibung korrekt wiedergegeben werden kann. Ein voreiliger Schluss über die Qualitäten des Komplexitätsbegriffs als kunstwissenschaftliches Arbeitswerkzeug ist an dieser Stelle dennoch nicht angebracht. Zwar machen die bisher diskutierten Ergebnisse deutlich, dass es intersubjektive Differenzen in der Auffassung der Komplexität von Kunstwerken gibt, doch wurde bisher lediglich zwischen zwei sehr diversen Gruppen unterschieden. Ob Komplexität gruppenintern ebenso divers bewertet wird wie im Vergleich der Gruppen, lässt sich den bisherigen Ausführungen nicht entnehmen. Es ist allerdings gerade diese Information wichtig, um den Einsatz des Komplexitätsbegriffes als Arbeitswerkzeug in den Kunstwissenschaften angemessen zu bewerten: Eine kohärente gruppeninterne Komplexitätsauffassung würde für eine einheitliche Expert\_innen-Perspektive sprechen, auf deren Grundlage fachintern Verbindlichkeit angenommen werden dürfte. Ich werde daher im folgenden Diskussionsabschnitt die Ergebnisse der Studie besprechen, welche die Übereinstimmung und die Spezifika der Komplexitätserfahrung von Expert\_innen adressieren.

### *c. Übereinstimmung in der Komplexitätserfahrung von Expert\_innen*

Aus den in der Studie dargelegten und diskutierten Ergebnissen lässt sich ableiten, dass sich Expert\_innen im Durchschnitt weniger in ihren Einschätzungen und Bewertungen unterscheiden als Laien. Das gilt hinsichtlich aller Werkaspekte, also sowohl der formalen Aspekte und ihrem Einfluss auf die Komplexitätsbewertung als auch der semantischen Werkaspekte und deren Einfluss auf die finale Bewertung. Das heißt, dass die finale Einschätzung der Expert\_innen-Gruppe aufgrund ähnlicher(er) Parameter getroffen wurde. Folglich sind sich auch die einzelnen Komplexitätsbewertungen, die den zuvor diskutierten Durchschnittswerten zugrunde liegen, in der Expert\_innen-Gruppe ähnlicher als in der Laien Gruppe. Das heißt, dass die durchschnittliche Komplexitätseinschätzung die ein Werk von



---

Expert\_innen erhält, in der Regel nicht auf stark abweichenden Einzelbewertungen basiert, sondern sich die Einzelbewertungen durchaus ähnlich sind.

Obwohl es zutreffend ist, dass sich die Expert\_innen-Bewertungen hinsichtlich der Komplexität eines Werkes im Durchschnitt durch mehr Ähnlichkeit auszeichnen als dies für Laien der Fall ist, kann aus diesem Ergebnis dennoch nicht gefolgert werden, dass Expert\_innen sich hinsichtlich der Frage nach der Komplexität eines Werkes immer absolut einig wären. Weder Expert\_innen noch Laien bewerten die Komplexität eines Kunstwerkes mit absoluter Übereinstimmung – und in beiden Gruppen gibt es sowohl Werke, bei denen die Übereinstimmung zwischen den Personen hoch ist, und Werke, bei denen die Übereinstimmung zwischen den Personen vergleichsweise gering ist.

Diese „Einigkeit“ lässt sich mittels einer simplen Maßzahl einfach abbilden: Die Standardabweichung von Durchschnittswerten. Vereinfacht gesagt gibt die Standardabweichung an, wie weit die einzelnen Werte, die einem Mittelwert zugrunde liegen, im Durchschnitt voneinander abweichen. Damit ist gemeint, dass die allermeisten Bewertungen in einer gewissen Nähe bzw. einem gewissen Abstand zum Mittelwert liegen. Liegt beispielsweise für eine Verteilung eine Standardabweichung mit dem Wert 2 vor, so bedeutet dies, dass die erhobenen Bewertungen im Durchschnitt  $\pm 2$  vom Mittelwert entfernt liegen. Die Tatsächlichen Bewertungen sind also meistens entweder 2 Werte höher oder niedriger als der Mittelwert. Je größer also die Standardabweichung, desto weniger gut bildet der Mittelwert die tatsächlichen Beobachtungen ab. Das Mittel ist demnach dann besonders Repräsentativ, wenn die Standardabweichung niedrig ist. In Bezug auf Bewertungen ist die Standardabweichung daher auch als Ausdruck von Ähnlichkeit oder Einigkeit interpretierbar. Eine perfekte Übereinstimmung, die auf lauter identischen Bewertungen basiert, würde den Wert 0 aufweisen. Liegen dem Durchschnittswert hingegen voneinander abweichende Einzelbewertungen zu Grunde, so wird sich der Wert der Standardabweichung mit zunehmender polarisierter Heterogenität der Bewertungen sukzessiv erhöhen. Was dabei als hohe oder niedrige Standardabweichung zu verstehen ist, hängt von den theoretischen

---

Minima und Maxima der gemessenen Größe ab. Können Befragte beispielsweise Werte von 1 bis 100 vergeben, so ist eine Standardabweichung mit dem Wert 2 sehr klein; ist hingegen der Antwortmöglichkeitenraum auf die Werte 1 bis 7 beschränkt, wie dies in der vorliegenden Studie der Fall war, so ist eine Standardabweichung mit dem Wert 2 relativ groß.

In der nachfolgenden Abbildung 9 sind die Standardabweichungen der durchschnittlichen Komplexitätsbewertungen von Expert\_innen für alle 90 untersuchten Werke dargestellt. Wie auch im Fall der zuvor besprochenen durchschnittlichen Komplexitätsbewertungen, handelt es sich bei den in Abbildung 9 visualisierten Werten, um die im Anhang der Studie in numerischer Form gelisteten deskriptiven Ergebnisse der Untersuchung. Anders als in den vorangegangenen Abbildungen sind die Werke in dieser Darstellung nicht nach der durchschnittlichen Komplexitätsbewertung geordnet, sondern von links nach aufsteigender Höhe der ermittelten Standardabweichungen gelistet.

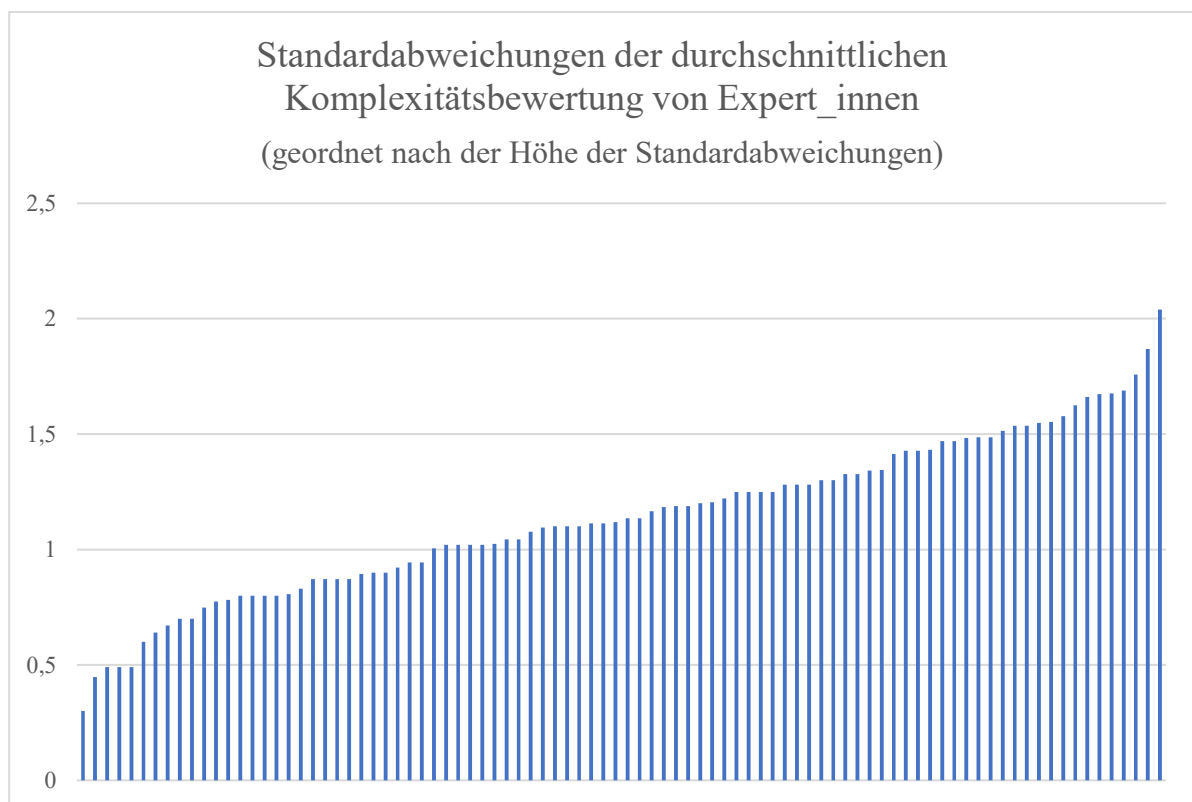


Abbildung 9 | Standardabweichungen der durchschnittlichen Komplexitätsbewertung von Expert\_innen (geordnet nach der Höhe der Standardabweichungen), eigene Darstellung.

Mit Blick auf die Grafik ist deutlich, dass ein guter Teil der Standardabweichungen unter 1 oder aber sehr nahe am Wert 1 liegt. Ein weiterer großer Anteil der Werte liegt zwischen 1,1 und 1,5. Nur sehr wenige Standardabweichungen liegen deutlich über 1,5. Tatsächlich sind hier sogar nur die drei letzten Werte deutlich höher als alle anderen.

Abbildung 10 zeigt die soeben beschriebenen Abschnitte zur besseren Veranschaulichung und Übersicht in einer eigenen Grafik. Hier ist noch deutlicher zu sehen, dass die meisten Werte tatsächlich im Bereich 0 bis 1 (36 von 90) oder im Bereich 1,1 bis 1,5 (44 von 90) liegen. Insgesamt liegen nur zehn Werte über der Marke von 1,5.

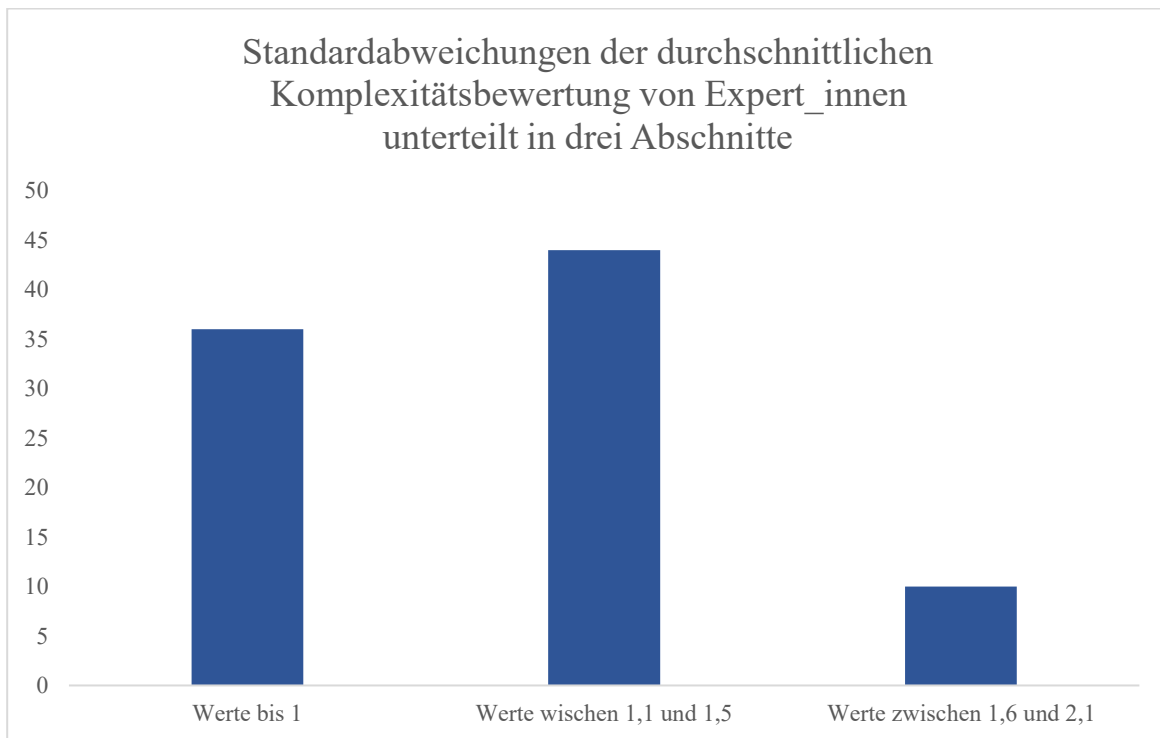


Abbildung 10 | Standardabweichungen der durchschnittlichen Komplexitätsbewertung von Expert\_innen unterteilt in drei Abschnitte, eigene Darstellung.

Die drei gewählten Wertebereiche sind dabei nicht willkürlich, sondern folgenden einer inhaltlichen Überlegung. Wie oben dargelegt besagt die Standardabweichung, dass die beobachteten Werte durchschnittlich  $\pm X$  vom Mittelwert der Verteilung entfernt liegen –

wobei X hier den jeweiligen numerischen Wert der Standardabweichungen bezeichnet. Für die besprochenen Mittelwerte der Komplexitätseinschätzungen liegen die theoretischen Minima und Maxima, wie bereits erwähnt, bei dem Wert 1 und 7. Die Auswahloptionen, zwischen denen die Befragten entscheiden konnten, reichen also, wie in Abbildung 11 zu sehen ist, vom Wert 1 (überhaupt nicht komplex) bis zum Wert 7 (sehr komplex).

1.	Wie komplex schätzen Sie das soeben gezeigte Bild ein?						
	○	○	○	○	○	○	○
	überhaupt nicht komplex	nicht komplex	eher nicht komplex	weder noch	eher komplex	komplex	sehr komplex

Abbildung 11 | Beispielfrage aus dem Fragebogen der Studie „*More Than the Sum of Its Parts. Perceiving Complexity in Painting*“, eigene Darstellung.

In der Zusammenschau der Unterteilungsabschnitte der Standardabweichungen (Abbildung 10) und der Spannweite der Antwortoptionen (Abbildung 11) ist ersichtlich, dass die gewählten Abschnitte sich an sinntragenden Referenzrahmen orientieren: Standardabweichungswerte von bis zu 1 decken maximal 3 Antwortoptionen ab, d.h. dass bei einem Mittelwert von beispielsweise 2 (nicht komplex) die Antworten in der Regel im Bereich „überhaupt nicht komplex“ bis „eher nicht komplex“ liegen. Hier ist also, trotz bestehender Abweichungen im Detail, im Wesentlichen immer etwas sehr Ähnliches gemeint – im Fall des angeführten Beispiels, dass das Werk tendenziell eine niedrige Komplexität hat. Standardabweichungswerte die über 1 liegen sprengen diese in sich kohärente Sinneinheit. Hier kann es durchaus sein, dass ein Kunstwerk das im Durchschnitt den Komplexitätswert 2 (nicht komplex) erzielt hat, von einigen als mittelmäßig komplex bewertet wurde. Da es sich bei der Standardabweichung um *durchschnittliche* Abweichungen vom Mittelwert handelt, bedeuten Werte von 1,1 bis 1,5 im vorliegenden Fall ausdrücklich, dass hier nicht alle Befragten mit ihren Bewertungen über dem Referenzrahmen einer sinntragenden Einheit liegen, sondern lediglich ein Teil davon. Da darüber hinaus liegen die meisten Werte dieses Abschnittes, wie der detaillierten Darstellung

---

in Abbildung 9 zu entnehmen ist, tendenziell eher im Bereich 1,1 bis 1,3. Was bedeutet, dass nur *sehr wenige* Bewertungen nicht innerhalb der durch den Mittelwert repräsentierten Sinneinheit liegen. Mit Blick auf die Anzahl der befragten Expert\_innen ist deutlich, dass hier tatsächlich eine *sehr* kleine Zahl gemeint ist: Bei zehn Befragten Expert\_innen liegen nur etwa ein bis zwei Personen mit ihren Bewertungen jenseits der durch den Mittelwert ausgedrückten Sinneinheit. Der Mittelwert bildet jedoch auch bei Standardabweichungswerten von bis zu 1,5 einen noch guten Richtwert, da das Überschreiten der durch den Durchschnittswert ausgedrückten Sinneinheit innerhalb eines verhältnismäßig überschaubaren Rahmens verbleibt. Zwar kann es vorkommen, dass einzelne Bewertungen in der „benachbarten“ Sinneinheit liegen, doch ist es unwahrscheinlich, dass diese deutlich weiter entfernt sind. Damit ist gemeint, dass es bei einem Mittelwert von 2 (nicht komplex) zwar möglich und denkbar ist, dass einige wenige Bewertungen die Erfahrung einer mittleren Komplexität ausdrücken, die Bewertung desselben Werkes als hoch komplex hingegen unwahrscheinlich ist. Für den größten Teil der untersuchten Kunstwerke lässt sich daher schließen, dass die Einigkeit unter den Expert\_innen als hoch zu bewerten ist – der Mittelwert ist repräsentativ und (im weiteren Sinne) Ausdruck einer kohärenten Expert\_innen Perspektive.

Dieser Schluss gilt, wie auch der grafischen Darstellung in Abbildung 9 anschaulich zu entnehmen ist, für einen Großteil der untersuchten Werke: 80 der 90 gezeigten Arbeiten weisen in ihrem Bewertungsdurchschnitt eine Standardabweichung auf, die bei 1,5 oder geringer liegt. Für Werke deren durchschnittliche Komplexitätseinschätzungen eine Standardabweichung von über 1,5 aufweisen, liegt der Fall ein wenig anders. Hier ist definitiv von einem gehäuften und zum Teil auch deutlichen Überschreiten der durch den Mittelwert repräsentierten Sinneinheit auszugehen. Das gilt natürlich insbesondere für diejenigen Werke, deren durchschnittlich vergebene Komplexitätswerte eine Standardabweichung aufweisen, die deutlich näher am Wert 2 als am Wert 1,5 liegt. Für die meisten der zehn Werte, die in den Standardabweichungsbereich von 1,5 bis 2,1 fallen, ist dies jedoch nicht der Fall – die überwiegenden Werte liegen hier deutlich näher an der

---

Wertemarke 1,5. Die Einigkeit zwischen den Expert\_innen ist hier zwar leicht niedriger als im oberen Bereich des zweiten Unterteilungsabschnittes, aber nicht in drastischem Umfang. Begreifen wir die gesetzten Abschnittsgrenzen als sinngemäß und nicht absolut, ließen sich diese Beispiele also durchaus noch dem mittleren Abschnitt zuschlagen. Das gilt allerdings nicht für die bereits erwähnten drei Werke, deren durchschnittliche Komplexitätsbewertungen von Standardabweichungen begleitet werden, die *deutlich über* der Marke von 1,5 liegen. Hier ist notwendig davon auszugehen, dass die Expert\_innen-Meinungen merklich auseinandergehen.

Ogleich es sicherlich nicht ungewöhnlich ist, dass es (immer wieder) einzelne Werke gibt, zu denen die Expert\_innen-Meinungen auseinandergehen, ist dennoch die Frage interessant, wie diese augenscheinliche Divergenz zu erklären ist. Ich werde im Folgenden daher eine kurze Reflektion zu den betroffenen Werken wagen. Bevor ich die konkreten Werke bespreche, möchte ich die berichteten Zahlen noch genauer kontextualisieren. Bisher wurden die Standardabweichungswerte im Abstrakten besprochen. Es wurde nur die Bedeutung ihrer Höhe im Vergleich zur Spannweite der theoretischen Minima und Maxima der Verteilung besprochen. Nicht zur Sprache gekommen ist bisher die Relation von Standardabweichung und Mittelwert, womit gemeint ist, dass in den bisherigen Ausführungen keine Aussagen dazu getroffen wurden, wie hohe oder niedrige Standardabweichungswerte mit hohen oder niedrigen durchschnittlichen Komplexitätsbewertungen zusammenhängen. In der nachfolgenden Abbildung 12 ist dieser Zusammenhang visualisiert. Zu sehen sind die Standardabweichungen der durchschnittlichen Komplexitätsbewertungen von Expert\_innen für alle 90 untersuchten Bilder.



Abbildung 12 | Standardabweichungen der durchschnittlichen Komplexitätsbewertung der *Expert\_innen* pro Bild (Bilder geordnet nach der durchschnittlichen Bewertung *aller* Befragten), eigene Darstellung.

Die Sortierung der einzelnen Werke entspricht der im vorherigen Diskussionsabschnitt etablierten Reihung: von links beginnend sind die Kunstwerke nach dem aufsteigenden Durchschnitt der von *allen* Versuchsteilnehmern vergebenen *Komplexitätswerte* geordnet. Die dargestellte Kurve zeigt demnach von links nach rechts die Veränderung der Standardabweichungen bei steigender durchschnittlicher Komplexitätsbewertung (nur *Expert\_innen*). Wie der Datenvisualisierung in Abbildung 12 zu eindeutig entnehmen ist, handelt es sich hier nicht um eine glatte Kurve. Vielmehr verhält es sich so, dass über das gesamte Komplexitätsspektrum hinweg kontinuierlich sowohl Ausschläge in der Höhe der jeweiligen Standardabweichungen nach oben als auch nach unten zu verzeichnen sind. Die *Expert\_innen* werden sich also nicht mehr oder weniger „einig“ je komplexer oder simpler ein Werk im Durchschnitt bewertet wird, sondern weisen bei allen Komplexitätsgraden ein gewisses Maß an „Uneinigkeit“ auf – diese liegt jedoch, wie soeben besprochen, für den überwiegenden Anteil der Werke unterhalb einer kritischen Grenze. Interessant ist auch, dass die drei „problematischen“ Fälle, also diejenigen Werke die eine größere Uneinigkeit in der

---

Expert\_innen-Gruppe hervorgerufen haben, sich alle, wenn auch nicht ganz am oberen Ende, so doch im höheren Komplexitätsspektrum befinden. Am unteren Ende der Komplexitätsskala finden sich keine vergleichbaren „Ausreißer“.

Bei den drei Werken, deren durchschnittliche Komplexitätseinschätzungen eine (verhältnismäßig) große Standardabweichung aufweisen, handelt es sich um Max Ernsts *Die Versuchung des Hl. Antonius* (1945), Jackson Pollocks *Convergence* (1952) und Gustav Klimts *Das Märchen* (nicht datiert). Um die hier zu Tage tretende „Uneinigkeit“ besser einschätzen und bewerten zu können, gehe ich auf diese drei Werke und die für sie vergebenen Einzelbewertungen im Folgenden näher ein.



Abbildung 13 | Ernst, Max (1945). *Die Versuchung des Hl. Antonius* (108 x 128 cm), Foto: Artstor, © Wilhelm-Lehmbruck Museum, Duisburg.

Max Ernsts *Die Versuchung des Hl. Antonius* (Abbildung 13) hat von den drei „problematischen“ Werken gleichzeitig die kleinste Standardabweichung und die höchste mittlere Komplexitätsbewertung – es handelt sich dabei also um den letzten größeren



Ausschlag der Kurve in Abbildung 12 auf der rechten Seite. Im Durchschnitt *aller* Bewertungen, also sowohl der Bewertungen der Expert\_innen als auch der Bewertung der Laien, hat dieses Werk einen Gesamtkomplexitätswert von 6,09 erhalten (auf einer Skala von 1 bis 7, siehe Abbildung 11). Legt man lediglich die Expert\_innenbewertungen für den Durchschnitt zugrunde, so liegt die mittlere Komplexität des Werkes ein wenig tiefer, nämlich bei 5,9. Die Standardabweichung liegt dabei in der *Expert\_innengruppe* bei 1,76. Mit Blick auf die konkreten einzelnen Bewertungen die von den jeweiligen Expert\_innen vergeben wurden (Tabelle 1) ist deutlich, dass dieser verhältnismäßig hohe Wert nicht dadurch zustande kommt, dass die Einzelbewertungen eine weite Streuung aufweisen würden; ganz im Gegenteil, bis auf *Person 6* liegen alle Bewertungen sehr nahe beieinander. Es ist also dieser eine Wert, der die *durchschnittliche* Abweichung stark beeinflusst.

Komplexitätsbewertung Max Ernst	
Person 1	7
Person 2	7
Person 3	7
Person 4	7
Person 5	6
Person 6	1
Person 7	5
Person 8	6
Person 9	6
Person 10	7
<i>Mittelwert</i>	5,9
<i>Standardabweichung</i>	1,76

Tabelle 1 | Einzelaufstellung der Komplexitätsbewertungen der zehn Expert\_innen für Max Ernsts *Die Versuchung des Hl. Antonius* (1945), eigene Darstellung.

Da hier nur eine einzige Person einen deutlich abweichenden Wert vergeben hat und dieser Wert darüber hinaus das andere Extrem bedient, ist es durchaus wahrscheinlich, dass *Person 6* bei der Beantwortung der Frage ein Fehler unterlaufen ist. Die Einschätzung von *Person 6*

---

würde nahelegen, dass das Werk als „überhaupt nicht komplex“ wahrgenommen wurde (vgl. Antwortkategorien in Abbildung 11). Da jedoch kaum von der Hand zu weisen ist, dass Max Ernsts *Hl. Antonius* sowohl in formaler als auch semantischer Hinsicht als relativ anspruchsvoll zu bewerten ist, ist eine fehlerhafte Eingabe für diesen Fall umso wahrscheinlicher. Dieser „Ausreißer“ ist ein sehr schönes Beispiel dafür, inwiefern einzelne Fehler bei der Datenerhebung zum Teil erheblichen Einfluss auf das Ergebnis der Untersuchung nehmen können. Da die diversen Fehleranfälligkeiten von Messungen nachfolgend noch ausführlich diskutiert werden, verzichte ich an dieser Stelle auf eine vertiefende Besprechung dieses Problems und verweise auf Kapitel 3.2.

Anders als bei Ernsts *Die Versuchung des Hl. Antonius* ist im Falle des zweiten und dritten „Ausreißers“ nicht von einem Fehler auszugehen – hier gehen die Expert\_innen-Meinungen tatsächlich auseinander. Jackson Pollocks *Convergence* (Abbildung 14) hat im Durchschnitt *aller* Befragten einen Komplexitätswert von 5,9 erhalten; im Durchschnitt der befragten Expert\_innen erreicht das Werk hingegen nur eine mittlere Komplexität von 4,9. Begleitet wird dieser Expert\_innendurchschnitt von einer relativ hohen Standardabweichung mit einem Wert von 1,87.

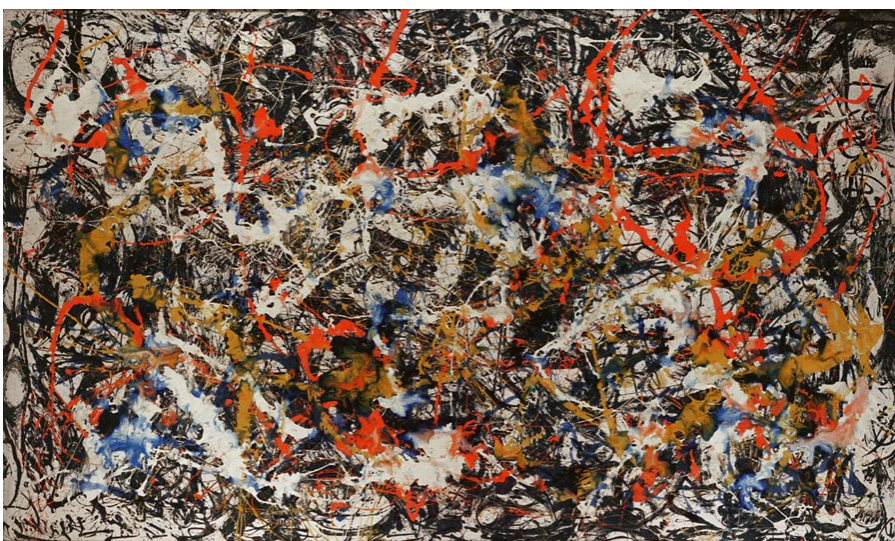


Abbildung 14 | Pollock, Jackson (1952). *Convergence* (241.9 x 399.1 cm), Foto: Artstor, © Albright-Knox Art Gallery.

Der relative hohe Wert kommt in diesem Fall aufgrund zweier unterschiedlicher Bewertungscluster zustande: In der Einzelansicht der jeweiligen Expert\_innenbewertungen ist ersichtlich, dass hier offenbar zwei „Gruppen“ existieren. Eine der beiden „Gruppen“ hat tendenziell niedrige Werte vergeben (*Person 1, Person 2* und *Person 3*); die zweite „Gruppe“ hat hingegen eher hohe Werte vergeben (vgl. Tabelle 2). Mit Blick auf das Werk könnte man annehmen, dass die beiden Gruppen jeweils verschiedene Kriterien für ihre Bewertung zugrunde legen. Eine hohe Komplexitätsbewertung würde in einem solchen Fall von Personen vergeben werden, die eine starke Gewichtung auf formale Aspekte legen, wie beispielsweise Übersichtlichkeit, Klarheit des Bildaufbaus, Symmetrie, Rhythmus, etc; eine niedrige Komplexitätsbewertung hingegen, würde man diesem Gedanken zufolge eher erwarten, wenn semantische Aspekte eine größere Rolle für die Bewertung gespielt haben. Die *Nicht-Inhaltlichkeit* von Pollocks Arbeiten sowie fehlende, im Motiv direkt verankerte, Bedeutungsschichten mögen in diesem Fall zu einer niedrigen Komplexitätsbeurteilung geführt haben.

Komplexitätsbewertung Jackson Pollock	
Person 1	2
Person 2	3
Person 3	2
Person 4	5
Person 5	5
Person 6	5
Person 7	7
Person 8	7
Person 9	6
Person 10	7
<i>Mittelwert</i>	4,9
<i>Standardabweichung</i>	1,87

Tabelle 2 | Einzelauflistung der Komplexitätsbewertungen der zehn Expert\_innen für Jackson Pollocks Convergence (1952), eigene Darstellung.

---

Eine solche Unterscheidung als gegeben anzunehmen, ist jedoch problematisch, denn die Frage, ob Pollocks Arbeiten (in diesem Fall *Convergence*) tatsächlich als in semantischer Hinsicht *simpel* zu bewerten sind, ist letztlich eine Frage der Perspektive, welche auf Werke des abstrakten Expressionismus angelegt wird. Die grundlegende Geisteshaltung, aus welcher heraus die Kunst der abstrakten Expressionisten entstanden ist, bildet ein durchaus anspruchsvolles Rahmenkonzept: die Annahme, dass sich in diesen (neuen) Ausdrucksmitteln nach dem zweiten Weltkrieg die freie Kunstform des Westens manifestiert, welche etwas später auch einen deutlichen Gegenpol zur dogmatisch überformten sozialistischen Kunst ermöglicht und gleichzeitig als Ausdruck von Freiheit die verdinglichte Inkorporation westlicher Werte darstellt, ist in Anbetracht der historischen Rahmenbedingungen des immer weiter an Fahrt aufnehmenden Kalten Krieges sowohl in gesellschaftlicher und politischer als auch in kultureller Hinsicht spannend. Angesichts einer solchen Perspektive, lässt sich kaum ausschließen, dass diejenigen Expert\_innen, die hohe Komplexitätsbewertungen vergeben haben, sowohl formale als auch semantische Aspekte in ihren Überlegungen zum Werk berücksichtigt haben. Das Ergebnis der Untersuchung zu Pollocks *Convergence* ist damit letztlich ein gutes Beispiel dafür, dass sich Kunstwerke nicht grundsätzlich „über einen Kamm scheren lassen“ in dem durch eine auf Strukturen bedachte Perspektive alle Werke bedingungslos unter gleichbleibenden und dahingehend auch verbindlichen Bedingungen betrachtet werden: Die (theoretische) Perspektive, mit welcher sich Forschende einem Werke nähern, beeinflusst ganz offensichtlich dessen Bewertung (siehe auch Methodenteil, Kapitel 3.3); bei einigen Werken – und Arbeiten des abstrakten Expressionismus sind dafür sicherlich kein schlechtes Beispiel – gehen die Perspektiven, die auf ein und dasselbe Werk angelegt werden können, durchaus beträchtlich auseinander. Damit ist dieses Teilergebnis der Untersuchung eine wichtige Erinnerung, dass bei allen Vorteilen, die eine auf Strukturen bedachte Perspektive mit sich bringen mag, das Einzelwerk doch nicht vergessen oder vernachlässigt werden darf. Die sehr speziellen Voraussetzungen und Gegebenheiten der jeweiligen Werke sind auch für den auf gröbere Strukturen abzielenden Blick von Bedeutung: nur in der Zusammenschau von systematischer

---

Perspektive und dessen konkreter Ausprägung am Beispiel des Einzelfalls ergibt sich ein vollständiges Bild.



Abbildung 15 | Klimt, Gustav (1883). *Das Märchen* (85 x 116 cm), Foto: Artstor, © Wien Museum.

Diesen Schluss legt auch eine genauere Betrachtung des dritten „Ausreißers“ der Untersuchung nahe. Für Gustav Klimts *Das Märchen* (Abbildung 15) wurde im Durchschnitt *aller* Befragten ein Komplexitätswert von 5 vergeben; im Durchschnitt der Expert\_innen liegt die Komplexitätseinschätzung bei 4,8. Die Standardabweichung ist dabei die höchste, die in der vorliegenden Untersuchung besteht: sie liegt bei 2,04. Die jeweiligen Einzelbewertungen streuen im Fall vom Klimts *Das Märchen* noch deutlich breiter als dies bei Pollocks *Convergence* gegeben war. Damit ist gemeint, dass sich nicht „nur“ zwei Gruppen in den Bewertungen ausmachen lassen, sondern dass hier sowohl niedrige, mittlere und hohe Komplexitätseinschätzungen von den Expert\_innen vergeben wurden (vgl. Tabelle 3).

Komplexitätsbewertung Gustav Klimt	
Person 1	2
Person 2	2
Person 3	6
Person 4	2
Person 5	7
Person 6	4
Person 7	7
Person 8	5
Person 9	6
Person 10	7
<i>Mittelwert</i>	4,8
<i>Standardabweichung</i>	2,04

Tabelle 3 | Einzelaufstellung der Komplexitätsbewertungen der zehn Expert\_innen für Gustav Klimts *Das Märchen* (nicht datiert), eigene Darstellung.

Diese Spannweite an unterschiedlichen Einschätzungen ist zunächst nicht mit abweichenden semantischen Überlegungen zu erklären; dies gilt insbesondere, da mit Blick in die weiteren Antworten der Befragten deutlich wird, dass alle Expert\_innen hier richtig verstanden haben, dass es sich um eine Darstellung von Fabelgeschichten handelt. Alle Expert\_innen haben hier also während der Bewertungsphase ähnliche inhaltliche Überlegungen zum Kunstwerk angestellt. Die beobachteten Abweichungen zwischen den einzelnen Bewertungen lassen sich jedoch auch nicht auf eine unterschiedliche Beurteilung der formalen Werkkriterien zurückführen: alle Expert\_innen haben für *Das Märchen* einheitlich eine mittlere formale Komplexität angenommen. Die sich unterscheidende Bewertung der *Gesamtkomplexität* des Werkes muss also an die unterschiedliche Bewertung des dargestellten Themas gekoppelt sein: während einige Expert\_innen die Darstellung von Fabeln als in semantischer Hinsicht wenig komplex begreifen, verstehen andere das Thema als durchaus vielschichtig. Ähnlich wie im Falle von Pollocks *Convergence* liegt hier also eine unterschiedliche Bewertung der semantischen Struktur vor, die sich als sehr abhängig von der speziellen Perspektive der befragten Expert\_innen zeigt. Detaillierte Daten über die Überlegungen, welche die Befragten in der Auseinandersetzung mit Klimts *Das Märchen* angestellt haben, sind in der



---

Studie nicht erhoben worden. Dass Kenntnis von diesen Überlegungen jedoch nötig wäre, um eine Aussage über die Komplexität Klimts *Das Märchen* zu treffen, ist angesichts des diesbezüglichen Ergebnisses offensichtlich. Der vorliegende Fall ist somit ein zugespitztes Beispiel für das Problem, welches bei der Bewertung von Pollocks *Convergence* bereits greifbar wurde: einerseits ist es hier so, dass – wie auch schon im Falle von *Convergence* – eine genauere Auseinandersetzung mit dem Werk und den diese Auseinandersetzung begleitenden Perspektiven notwendig wird; andererseits steht die Art der Datenerhebung einer solchen Auseinandersetzung im Weg. Hier entstehen also wegen des (spezifischen) empirischen Zugangs Probleme, die durch den Zugang selbst nicht gelöst werden können, sondern einer anderen Form der Bearbeitung bedürfen. Diese kann zwar sowohl theoretisch als auch empirisch sein, muss aber in jedem Fall über den in dieser Studie gewählten starken Formalismus hinausgehen.

Obwohl die Probleme des gewählten empirischen Ansatzes nicht unerheblich sind, sollte doch betont werden, dass die Anzahl solcher „problematischer“ Fälle sehr gering ist. Die dargelegten Probleme sind damit nicht besonders ausgeprägt. Ihr Aufscheinen stellt den grundlegenden Ansatz nicht in Frage, macht aber deutlich, dass empirische Daten immer einer ausführlichen Inspektion bedürfen. Darüber hinaus sind diese „problematischen“ Fälle auch insofern als positiv zu bewerten, als sie gewährleisten, dass sie Arbeit am und mit dem Bild trotz eines an Strukturen orientierten Forschungsparadigma seine zentrale Bedeutung aufrechterhält. Eine empirisch-arbeitende Kunstwissenschaft kann somit zwar im Sinne des Wölfflin'schen Ansatzes als „Kunstgeschichte ohne Namen“ operieren, nicht aber – um es drastisch zu formulieren – als Kunstwissenschaft die ohne Bilder auskommt.

Die Diskussion soweit zusammenfassend lassen sich die wichtigsten Ergebnisse der Studie auf drei Punkte bringen: a) Komplexität ist ein erfahrungsbasiertes Phänomen, das sowohl formale als auch semantische Kriterien umfasst; b) die Erfahrung von Komplexität variiert mit dem Expertisegrad der Betrachtenden; c) Expert\_innen weisen in ihren Komplexitätseinschätzungen eine hohe Übereinstimmung auf.

---

### *Fazit und Beitrag der empirischen Methoden*

Die im Rahmen der Studie behandelten Fragen sind aus kunstwissenschaftlicher Perspektive, wie eingangs dargestellt, aus zwei Gründen interessant: zum einen handelt es sich trotz der empirischen Herangehensweise um Problemstellungen mit kunsttheoretischem Charakter, die zur näheren Beleuchtung von rezeptionsästhetischen Überlegungen beitragen können, da hier die Wahrnehmung und Erfahrung von Kunstwerken im Zentrum des Interesses steht; zum anderen werden durch die Studie aber auch Methodenfragen adressiert, da eine Dekonstruktion des Komplexitätsbegriffs dessen Verwendung hinterfragt und sich damit ein Potenzial zur Aktualisierung kunstwissenschaftlicher Praxis realisiert.

Hinsichtlich der ersten durch die Studie abgedeckten Problemstellung lässt sich aus den Ergebnissen die klare Aussage ableiten, dass die Erfahrung von Komplexität weder ausschließlich durch das Werk, noch ausschließlich durch die Perspektive der Betrachtenden determiniert ist. Die Erfahrung ist ganz im Gegenteil wesentlich von beiden Faktoren abhängig; nur im Zusammenspiel von Werkeigenschaften und Betrachtendeneigenschaften ist der Gesamteindruck zu erklären, der bei der Betrachtung der Kunstwerke entsteht. Dieses Ergebnis ist auch für die Methodenfragen interessant, die sich durch die gesetzte Problemstellung eröffnen. Dem Ergebnis ist zu entnehmen, dass die Komplexität eines Kunstwerks keineswegs als objektives Konzept gelten kann. Aufgrund der hohen Übereinstimmung von Kunstexpert\_innen in sowohl ihren Komplexitätsbewertungen als auch in den zu diesen Bewertungen führenden Einflussfaktoren, lässt sich Komplexität im Rahmen gewisser Einschränkungen aber immerhin als quasi-objektives Konzept fassen. Legt man ausschließlich Kunstexpert\_innen zugrunde, deren Wahrnehmung und Verarbeitung von Kunstwerken qua ihrer Ausbildung ähnlich ist, so lassen sich Komplexitätserfahrungen, trotz der diskutierten „Ausreißer“, als intersubjektiv-stabile Bewertungen verstehen. Stellen wir also die Frage nach der Validität des Begriffes *Komplexität* als Arbeitswerkzeug in den Kunstwissenschaften, lässt sich darauf eine eingeschränkt positive Antwort geben: *positiv*, weil die intersubjektiv-stabile Komplexitätserfahrung der Kunstexpert\_innen, die subjektive



---

Komponente des Konzeptes „überschreibt“ und diese Subjektivität somit für die Praxis der Begriffsverwendung irrelevant ist; *eingeschränkt* positiv, weil dies nur für den Kontext ganz *bestimmter* Verwendungszusammenhänge gilt. Kommt der Begriff in der Analyse und Beschreibung von Kunstexpert\_innen zum Tragen, ist er also Ausdruck einer Expert\_innenbewertung, so ist die Verwendung des Begriffes aus den genannten Gründen unproblematisch. Wird mit dem Begriff aber auf eine allgemeine, dem Kunstwerk zugeschriebene Eigenschaft verwiesen, so ist dies nicht sinnvoll. Insbesondere Verwendungszusammenhänge, in denen Komplexität explizit in Zusammenhang mit der Erfahrung von Kunst gebracht wird, wie es etwa Usus in der psychologischen Ästhetikforschung ist, bedürfen einer reflektierteren Verwendung des Begriffes: der Verwendungszusammenhang sollte in diesem Fall die Mehrdimensionalität des bezeichneten Konzeptes berücksichtigen und die Konsequenzen dieser Mehrdimensionalität offenlegen.

Die Überlegungen zu dieser Studie abschließend möchte ich noch den Beitrag bzw. Mehrwert besprechen, der durch die empirische Herangehensweise realisiert wurde. Inwiefern also lassen sich die wichtigsten Erkenntnisse der Untersuchungen direkt auf den Einsatz von empirischen Methoden zurückführen (vgl. Kapitel 1.3)? Hierzu ist zunächst festzuhalten, dass der explizit auf intersubjektiv-stabile Strukturen gerichtete Blick, also die Suche nach Regelmäßigkeiten in der Wahrnehmung von Komplexität, notwendig war um bestehende Konstanten offenzulegen. Das Offenlegen dieser Konstanten ist im Umkehrschluss essentiell, um auch Abweichungen von bestimmten Mustern aufzuzeigen. Die Tatsache, dass Kunstexpert\_innen und Kunstlaien ein abweichendes Verständnis von Komplexität in der Kunst haben, dass durchaus praxisrelevante Unterschiede in der Wahrnehmung von Komplexität bestehen, konnte nur durch die systematische Untersuchung zweier Vergleichsgruppen und deren (Bewertungs-)Verhalten der Beobachtung zugänglich gemacht werden. Auch die Erkenntnis, dass die Rollen diverser formaler und inhaltlicher Bildaspekte in der Komplexitätserfahrung nicht äquivalent sind, sondern dass durchaus Unterschiede in Wichtigkeit und Art des Einflusses bestehen, wäre ohne den auf Strukturen bedachten Blick empirischer Ansätze nicht zu erarbeiten gewesen.

---

Tatsächlich hätten sich diese Erkenntnisse auch nur schwerlich auf der Basis kunstwissenschaftlicher Theoriebausteine antizipieren lassen. Mit einem semiotischen Ansatz, beispielsweise, hätten sich zwar die Determinanten der Komplexitätserfahrung ableiten lassen, nicht aber die Art und Weise, in welcher sie die Praxis der Erfahrung tatsächlich beeinflussen. Es ist allerdings diese konkrete Praxis der Erfahrung, die über das zugrundeliegende Konzept die beste Auskunft gibt – denn sie beeinflusst letztlich die Verwendung des Begriffes und ist somit für dessen realweltliche Bedeutung konstitutiv.

Da Komplexität als ein erfahrungs- und wahrnehmungsbasiertes Konzept zu verstehen ist, ist darüber hinaus auch die Überwindung subjektiver Forschungsperspektiven zur Fassung des Phänomens essentiell. Die einzelne Forscherpersönlichkeit würde in der subjektiv-hermeneutischen Auseinandersetzung mit dem Problem notwendig die eigene Erfahrung und Wahrnehmung beschreiben. Während dies grundsätzlich für jene Kontexte unproblematisch ist, in denen Wahrnehmungsinhalte als intersubjektiv-stabile Konzepte angenommen werden können, ist der Fall der Komplexitätserfahrung weniger trivial. Wie wir gesehen haben, bestehen signifikante Unterschiede in der Weise, wie Kunstexpert\_innen und Kunstlaien die Komplexität von Kunstwerken erfahren. Der explizite Nutzen einer empirischen Vorgehensweise ist in diesem Fall offensichtlich: Das Erkennen von Regelmäßigkeiten, Abweichungen und merkmalspezifischen Unterschieden ist für das Erfassen des Wesens von Komplexität in der Kunst unabdingbar.

Der Rückgriff auf die im Auswertungsprozess der Studie verwendeten Mehrebenenanalysen, erlaubt darüber hinaus die gleichzeitige Betrachtung verschiedener Einflussfaktoren sowie das Herausstellen des für die einzelnen Faktoren spezifischen Einflusses auf die Erfahrung.<sup>131</sup> Ohne sowohl eine breite Datenbasis wie auch ein methodisches Werkzeug, das strukturierten Einblick in diese Daten erlaubt, wäre ein Zugang zu den einzelnen Einflussfaktoren und damit auch die Bewertung ihrer jeweiligen Relevanz für das Konzept der Komplexität in

---

<sup>131</sup> Rabe-Hesketh/Skrondal 2012.

---

Kunstwerken nicht greifbar geworden. Die Frage also, welche Aspekte in einem holistisch(er) gefassten Konzept notwendig inkludiert sein müssen und welche Rolle sie jeweils für das untersuchte Phänomen spielen, lässt sich ohne eine Möglichkeit, die untersuchten Einzelteile als solche sichtbar herauszustreichen nicht klären. Die Beantwortung der im Rahmen dieser Studie behandelten Problemstellungen ist demnach nur durch die gewählte Herangehensweise zu leisten gewesen. Ein empirischer Ansatz ist für den gesetzten Fragen- und Problemraum essentiell.

---

# Studie

## Abstract Art as a Universal Language?

LITERATURANGABE

---

Hanna Brinkmann/ Laura Commare/ Raphael Rosenberg/ Helmut Leder, Abstract Art as a  
Universal Language?, in: Leonardo, 47(3), 2014, S. 256-257.

---

## 2.2 *Abstract Art as a Universal Language?*

Die zweite Studie, die ich hier besprechen möchte, ist 2014 als Kurzaufsatz in der interdisziplinären Zeitschrift *Leonardo* erschienen und trägt den Titel *Abstract Art as a Universal Language?*. Die Studie ist in Zusammenarbeit mit Hanna Brinkmann, Raphael Rosenberg und Helmut Leder entstanden.<sup>132</sup> Im Kern adressiert sie die Frage, inwiefern abstrakte Kunst einer *universellen Kommunikation* entspricht, wie es von Kunsttheoretiker\_innen und Künstler\_innen insbesondere seit den 1950er Jahren propagiert wurde. Anders als beim zuvor besprochenen Beispiel handelt es sich hier nicht um eine Studie, die im weiteren Sinne als theoriegebend zu verstehen ist. Hier werden bestehende Thesen zur menschlichen Wahrnehmung und Erfahrung bei der Betrachtung und Auseinandersetzung mit abstrakter Kunst aufgegriffen und hinsichtlich ihrer realweltlichen Praxisrelevanz überprüft. Es handelt sich bei dieser Studie also um einen *theorietestenden* Ansatz. Das hat zur Folge, dass sowohl die Diskussion der behandelten Thematik als auch der Umfang der Untersuchung deutlich weniger ausführlich ausfallen als dies im Falle eines theoriegebenden Ansatzes der Fall wäre. Theorietestende Ansätze vom Format der nachfolgenden Studie können daher gewissermaßen als Kommentar zu einem bestehenden Diskurs verstanden werden. Wie auch im vorangegangenen Abschnitt ist die Studie von einer kurzen Einführung und einer abschließenden Diskussion der Ergebnisse gerahmt. Da der Aufsatz aufgrund der Richtlinien der Zeitschrift *Leonardo* kurz gehalten ist, werde ich den folgenden Einführungsabschnitt nutzen, um die im Rahmen der Studie untersuchten Ideen ein wenig plastischer zu präsentieren. Der Diskussionsabschnitt wird die Ergebnisse der Studie kurz erläutern und besprechen, inwiefern eine empirische Herangehensweise zur Erlangung der vorgestellten Ergebnisse zwingend erforderlich war. Beide

---

<sup>132</sup> Die Mehrfachautor\_innenschaft des Textes ist durch das arbeitsteilige Vorgehen in der Vorbereitung, Durchführung und Nachbereitung der empirischeren Untersuchung bedingt. Die unterschiedlichen Arbeitsschritte sind dabei schwerpunktmäßig von jeweils bestimmten Teilen des Teams bearbeitet worden. Eine ausführliche Aufstellung der jeweiligen Aufgaben und Einzelleistungen findet sich im Anhang (Anhang *Stellungnahme zur Arbeitssteilung II*).

---

Rahmungsabschnitte sind dabei deutlich knapper gehalten als im zuvor besprochenen Beispiel.

### *Einleitung*

Die Idee von einer universalen Ausdruckskraft von Farben und Formen begleitet den Diskurs um abstrakte Kunst bereits seit ihren Anfängen. So findet sich etwa in Kandinskys *Punkt und Linie zu Fläche* (1926), das er als „organische Fortsetzung“ seines so prominenten Werkes *Über das Geistige in der Kunst* (1911) versteht, die Annahme ausgedrückt, dass Kunstwirkung auf einem universellen Verständnis der jeweiligen Wirkung seiner Grundelemente beruhe.<sup>133</sup> Die Vorstellung von abstrakter Kunst als eine Art universelle Sprache erlebt jedoch insbesondere nach dem Zweiten Weltkrieg eine Konjunktur. Hier wurde die Forderung nach einer abstrakten Bildsprache in der Kunst besonders nachdrücklich geäußert, da abstrakte Malerei als Gegenstück und Abgrenzung zur Nazikunst verstanden und politisiert wurde. Abstrakte Malerei wurde gleichsam zu einem Sinnbild für eine freie und demokratische Ausdrucksweise erhoben.

Im Rahmen dieses Nachkriegsdiskurses ist auch der bekannte, häufig Werner Haftmann zugeschriebene Slogan, „Abstraktion als Weltsprache“ entstanden. Haftmann kann als großer Fürsprecher der abstrakten Kunst und insbesondere der abstrakten Malerei gelten. Eine seiner prominentesten Thesen dazu ist, dass alle Kunstströmungen unweigerlich in der Abstraktion enden müssen. Das formuliert er bereits in seinem 1954 erschienenen und weitläufig rezipierten Werk „Malerei im 20. Jahrhundert“.<sup>134</sup> Im darauffolgenden Jahr, im Rahmen der Documenta II, an der er maßgeblich als führender Kurator beteiligt war, verfestigt sich diese Haltung. Obwohl der ihm zugeschriebene Slogan „Abstraktion als Weltsprache“ in dieser Form von Haftmann selbst nicht geäußert wurde, lassen sich Passagen im Katalog der Documenta II nachweisen, die diese Vorstellung stützen. So schreibt er beispielsweise: „Von

---

<sup>133</sup> Kandinsky 1911, Kandinsky 1926.

<sup>134</sup> Haftmann 1954.

---

Europa über die beiden Amerika und Asien bis hin zum Fernen Osten hat sie [die abstrakte Kunst] innere Übereinstimmungen wachrufen können und diese Übereinstimmungen in eine Sprachform einbetten können, die eine unmittelbare Kommunikation möglich macht. Sie kann als erster Modellfall von Menschheitskultur gelten.“<sup>135</sup> Hier wird nicht nur die Idee von abstrakter Kunst als einer Weltsprache deutlich, sondern auch die Vorstellung einer unmittelbaren Selbstvermittlung abstrakter Bilder.

Die Annahme, dass abstrakte Kunst universell verständlich und somit voraussetzungslose Kunst sei, die nicht an kulturelle, politische oder historische Kontexte gebunden ist, wurde nicht nur von Kunsthistorikern, wie Haftmann oder auch Clement Greenberg,<sup>136</sup> sondern auch von Künstlern, wie beispielweise Willi Baumeister, Barnett Newman, Jackson Pollock und Robert Motherwell, vertreten.<sup>137</sup> Willi Baumeister etwa äußert sich ganz explizit zur universalen Verständlichkeit von abstrakter Formensprache. Er schreibt: "Die sogenannte 'abstrakte' Malerei ist nicht abstrakt im Sinne von Fremdheit zum Leben und zum Menschen. Die Empfindungen des Künstlers sind ganz natürliche. Eine senkrechte, gerade Linie vermittelt einen bestimmten Empfindungswert, den alle Menschen empfangen können. Eine gekurvte Linie oder ein Fragezeichen löst dagegen andere Empfindungen aus. Ebenso ergeben die einzelnen Farben bestimmte Reaktionen."<sup>138</sup>

In den Reflektionen der Künstler\_innen wird insbesondere auch deutlich, dass jenseits des Erlebens von Ausdruck der Akt des Betrachtens für die Idee einer universellen Kunstsprache als wesentlich betrachtet und daher zwingend mitreflektiert werden muss. So schreibt beispielsweise Willi Baumeister über das *Betrachten* von abstrakter Malerei: „Kunstwerke kann man nicht naiv genug betrachten. [...] Gelingt es dem Beschauer diesen Zustand der Naivität zu erreichen, so sind alle Bedingungen gegeben, um das Kunstwerk aufnehmen zu können. Demnach ist der Zustand des Betrachters viel wichtiger, als seine mitgebrachte

---

<sup>135</sup> Haftmann 1959: 14.

<sup>136</sup> Greenberg 1986: 326.

<sup>137</sup> Hoormann 2007: 269.

<sup>138</sup> Baumeister 1971: 204.

---

Kunstkenntnis und Ableitungen davon.“<sup>139</sup> Ganz Ähnliches finden wir auch in einem Zitat von Barnett Newman: “The image we produce is the self-evident one of revelation [...] that can be understood by anyone who will look at it without the nostalgic glasses of history.”<sup>140</sup>

Wenn aus den Äußerungen zum Thema auch nicht eindeutig hervorgeht, was unter Welt- oder Universalsprache im Konkreten zu verstehen ist, so erlauben diese kunsttheoretischen und künstlerischen Positionen doch den Schluss, dass sich bei der Betrachtung von abstrakten Gemälden im Vergleich zur Betrachtung von gegenständlicher Malerei weniger intersubjektive Differenzen in der Wahrnehmung und Wirkung der Gemälde zeigen sollten. Der Kern, das Wesen, des abstrakten Kunstwerks müsste sich, sofern die Annahme einer Universalsprache zutreffend ist, eindeutig(er) und leicht(er) identifizieren lassen. Das sollte zur Folge haben, dass der Eindruck, den Betrachtende vom Werk gewinnen, im Durchschnitt deutlich homogener ist als dies bei einem gegenständlichen Vergleichsbeispiel der Fall wäre. Dies sollte jedenfalls für die Wirkung bzw. den erlebten Ausdruck der jeweiligen Werke gelten. Ob der tatsächliche Betrachtungsakt, also die Art des Sehens, der Blick auf das Bild, ebenfalls von dieser Homogenisierung betroffen sein sollte, ist den Äußerungen nicht eindeutig zu entnehmen. Nimmt man jedoch eine Übereinstimmung im Erleben des Ausdrucks an, so ist es, wenn schon nicht notwendig, dann zumindest wahrscheinlich, dass der Blick der Betrachtenden in ähnlicher Übereinstimmung auf die Trägerelemente dieses Ausdrucks fällt. Demnach müsste also auch der ganz konkrete Vorgang der Betrachtung im Fall von abstrakten Gemälden im Durchschnitt homogener ausfallen, als dies bei gegenständlicher Malerei der Fall ist.

Sowohl die Annahme eines homogenen Erlebens von Ausdruck als auch die Annahme eines homogenen Betrachtungsaktes stellen starke Thesen über die menschliche Wahrnehmung dar. Zwar sind diese Thesen nicht völlig unplausibel, trotzdem ist unklar, insbesondere eingedenk des ideologischen Hintergrunds ihres Entstehungskontextes, ob und inwiefern sie

---

<sup>139</sup> Baumeister 1952: Online ohne Seitenangabe.

<sup>140</sup> Newman 1992: 173.



---

tatsächlich zutreffen. Die Frage nach dem Praxisgehalt dieser Annahmen stellt sich umso mehr, als die sie begleitenden Aussagen zur menschlichen Wahrnehmung nur unzureichend empirisch belegt sind. Es bleibt somit offen, ob die Idee von *Abstraktion als Weltsprache* lediglich dem übersteigerten Idealtypus eines bestimmten Zeitgeistes entsprungen ist oder ob damit tatsächlich auch ein praxisrelevantes rezeptionsästhetisches Prinzip formuliert ist. Eine empirische Untersuchung, welche Antworten auf diese Fragen gestattet, indem sie die hier getroffenen Annahmen zur menschlichen Wahrnehmung der Sphäre der Spekulation enthebt, wurde bisher jedoch von der Forschung noch nicht angestrebt.

### *Bemerkungen zur Studie*

Die nachfolgende Studie adressiert die offenliegende Frage nach dem Praxisgehalt der Annahme einer sich durch Abstraktion realisierenden Universalsprache. Im Konkreten werden die Annahmen eines homogenen Erlebens von Ausdruck und die Annahme eines homogenen Betrachtungsaktes empirisch überprüft. Dazu werden zwei unterschiedliche Maße herangezogen: Zum einen ästhetische Urteile, die Rückschluss über das Erleben von Ausdruck gestatten; zum anderen mittels Eye-Tracking gewonnene Daten, welche die Analyse und den Vergleich des konkreten Blickverhaltens ermöglichen.

Die in ihrer Publikationsform beigelegte Studie hat, bedingt durch die Rahmenbedingungen des von der Zeitschrift *Leonardo* gesetzten Publikationsformates, einen sehr knappen Umfang. Das hat zur Folge, dass insbesondere Erläuterungen zur Methode hier keinen Platz gefunden haben. Eine ausführlichere Besprechung des Vorgehens bei der Messung des ästhetischen Urteils und dem Eye-Tracking findet sich im Text der *dritten* Studie, wo die Ergebnisse und Rahmenbedingungen der zweiten Studie aufgegriffen und breiter diskutiert werden.

## Studenttext

### ABSTRACT ART AS A UNIVERSAL LANGUAGE?

Hanna Brinkmann<sup>1</sup>, Laura Commare<sup>1</sup>, Helmut Leder<sup>2</sup>, Raphael Rosenberg<sup>1</sup>

<sup>1</sup>Department of Art History, University of Vienna, 1090 Vienna, Austria.

<sup>2</sup>Department of Basic Psychological Research and Research Methods, University of Vienna, 1010 Vienna, Austria.

Email: <hanna.brinkmann@univie.ac.at>.

See <www.mitpressjournals.org/toc/leon/47/3> for supplemental files associated with this issue.

Submitted: 23 October 2013

#### Abstract

The concept of abstract art as ‘world language’ became famous after documenta II (1959). Abstract art was considered as universally comprehensible and independent of cultural, political or historical contexts. However, this was never explicitly tested empirically. If these assumptions were true, there should be higher intersubjective coherence in perceiving abstract paintings compared to representational art. In order to test this hypothesis, the authors recorded the eye-movements of 38 participants and collected information on their cognitive and emotional evaluations. The results suggest that the concept of abstract art as a universal language was not confirmed and needs to be revised.

In 1959 documenta II in Kassel introduced the general public to the concept of abstract art as a “world language” [1]. This view was not only advanced by art historians such as Haftmann, the leading theorist of this documenta, or Clement Greenberg, but also by artists like Barnett Newman [2]. The ambition of artists like Jackson Pollock, Robert Motherwell and others was to promote the unprejudiced act of experiencing the pure value of lines and colors as such [3], which led to an increasing importance of the act of beholding. If this concept of abstract art as “world language” holds true, it should be seen in the act of beholding. We conducted an empirical study in which we expected that, if abstract art can be interpreted as a universal language, perception should be more homogeneous for abstract than for representational art. Assuming that the major elements of art perception are the actual seeing (eye-movements) and cognitive / emotional reactions, there should be a stronger intersubjective coherence in these elements in response to abstract, compared to representational, artworks.

#### Study Design

38 participants were shown high-quality reproductions of 5 abstract and 5 representational paintings for 2 minutes each in

random order on a high resolution computer screen (2560 x 1600 px). After viewing each painting, participants were asked to use a 7-point scale semantic differential (13 bipolar adjectives, e.g. active – passive, friendly – unfriendly, etc.) to indicate their cognitive and emotional evaluation of each artwork [4]. We matched the set of abstract artworks with a set of representational counterparts with regard to formal criteria, like dynamics, color and composition (Fig. 1 and Fig. 2). Abstract art by definition does not depict recognizable objects, but instead uses shapes and colors for their own expressive value in a nonrepresentational way. The set of abstract paintings we chose includes two works by Kandinsky and one each by Pollock, Motherwell and Götz. Works of all these painters were part of documenta II. We chose the 5 abstract paintings in order to have significant differences in color, in the number of depicted elements and dynamics without manipulating the original artworks. The set of representational counterparts was chosen according to results of a pre-study by 28 experts, who considered them as similar to the abstract paintings. Besides image characteristics, we controlled the participants’ gender and expertise. We measured expertise by taking into account the participant’s field of study, their knowledge about the selected artists and the frequency in which they conduct art-related activities in their leisure time, such as going to museums or talking about art with friends. We recorded the eye-movements of the participants while they were perceiving the artworks, using a binocular remote eye-tracker (SMI Red 120). Fixations of the eye usually correspond to visual attention. Therefore it is possible to describe participants’ attention profiles by analyzing fixation patterns. In the analyses, we compared the attention profiles evoked by the two groups of artworks. To measure the distribution of fixations we chose a bottom-up approach – not defining regions of interest in advance. Thus a cluster – or region of interest – was defined as an area with a radius of 90 px and with a min. of 5 fixations per minute. We measured the number of fixations (expressed as a percentage) located in clusters (Fig. 1 and Fig. 2). A high number of fixations located in clusters can be interpreted as a focused or directed attention, while for a low number of fixations located in clusters a broader distribution of attention can be assumed, resulting in a more diffuse attention profile. Thus the latter is a sign of lower coherence. The emotional and cognitive responses were analyzed separately from the eye-tracking data.

Fig. 1 and Fig. 2: Visualization of one participant’s eye-movements when viewing Félix Vallotton “Street Scene” (left) and Vasily Kandinsky “Accompanied Contrast” (right) analyzed with Eye-Trace. Black circles indicate single fixations, white circles indicate shared clusters (38 persons): The representational artwork shows more coherence. © Laboratory of Cognitive Research in Art History, University of Vienna)





## Results

We ran two separate mixed linear models with random intercepts and random slopes for image characteristics (intensity of color, number of depicted elements, presence of depicted humans, abstraction as dichotomous variable).

The first model (Fig. 3. Model 1) addresses the question whether there are differences in individual attention profiles of participants concerning abstract and representational artworks. Here, the dependent variable was the number of fixations located in individual clusters. The analyses revealed a highly significant difference ( $p < 0,001$ ) in participants' attention profiles when viewing abstract compared to representational art, despite controlling for image characteristics: participants show 10,17% more diffuse attention profiles when viewing abstract art. While none of the socio-demographic variables influenced the effect significantly, image characteristics played a major role in structuring attention. Bright colors tended to raise focused attention (very bright colors  $\beta = 13,63$ ,  $p < 0,001$ ; medium bright colors  $\beta = 11,02$ ,  $p < 0,001$ ) and the number of depicted elements led to more diffuse eye-movement patterns (high number of depicted elements  $\beta = -10,39$ ,  $p < 0,001$ ). As the ICC for all image characteristics was below 0,01 no random effects are reported. The second model (Fig. 3. Model 2) predicts the homogeneity of the localization of fixations.

We analyzed the collective distribution of attention by modeling the number of fixations located in shared regions of interest. We found very similar results as in the analysis of number of fixations: the effect of abstraction ( $\beta = -11,04$ ) was highly significant ( $p < 0,001$ ) and there were no significant effects for socio-demographic data, but highly significant effects for image characteristics (very bright colors  $\beta = 20,55$ ,  $p < 0,001$ ; medium bright colors  $\beta = 13,34$ ,  $p < 0,001$ ; high number of depicted elements  $\beta = -16,49$ ,  $p < 0,001$ ). In the second model the effect of depicted humans was also significant ( $p < 0,05$ ): especially faces in a painting led to slightly more consistent clustering of attention ( $\beta = 3,26$ ). As in the first model, the ICC for all image characteristics was below 0,01.

The analysis of the semantic differential concerning the two different types of paintings showed a consistent rating for some items. But the items participants rated coherent differed from painting to painting (e.g. Pollock was rated by 95% of the participants as active, excited and dynamic). Obviously different paintings have different qualities and people seem to agree strongly in items which describe the essential quality of a work, independent of its type – abstract or representational. In order to test for coherence, we compared the mean of the standard deviation to examine how homogeneous the evaluation was. Stronger coherence comes along with smaller standard deviation. A T-Test revealed no significant effect ( $p = 0,8$ ). We thus conclude that there are no differences between abstract and representational artworks concerning the coherence of cognitive and emotional response.

## Conclusion and Limitations

One of the most persistent theories in modern art claims that abstract art is a universal language. In an empirical study comparing perception and evaluation of abstract and well-matched representational counterparts we found no evidence that ab-

Fixed Effects	Model 1	Model 2
	% of Fixations in Clusters	% of Fixations in shared Clusters
abstract art	-10,17***	-11,04***
field of study is art history	2,14	-1,71
knowledge about artists	-0,07	-0,23
art-related activities	-0,13	0,01
sex	2,53	2,57
depicted humans	-2,01	3,26*
medium bright colors	11,02***	13,34***
very bright colors	13,63***	20,55***
medium number of elements	0,66	1,72
high number of elements	-10,39***	-16,49***
Intercept	51,68***	25,35***

Random Effects	Estimate	Estimate
Intercept	4,72	2,11
Residual	11,09	8,47

Legend: \*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$

Fig. 3. Mixed linear models with random intercepts and random slopes for image characteristics. Random effects are not displayed (ICC < 0,01).

stract art elicits more coherent beholder behavior. Quite the opposite: in contradiction to the idea of abstract art as a universal language, our results indicate that abstract paintings do not evoke a more homogeneous perception than representational paintings - neither concerning eye-movements nor regarding the evaluative responses. Eye-movement patterns clearly were more heterogeneous for abstract compared to representational art. Based on these findings, we provide empirical evidence that the idea of abstract art as a universal language needs to be revised.

## References

1. W. Haftmann, *Malerei nach 1945. documenta-2-Katalog* (Köln: DuMont Schauberg, 1959), p. 14; cf. G. Poensgen/ L. Zahn, *Abstrakte Kunst eine Weltsprache* (Baden-Baden: Woldemar Klein, 1958).
2. C. Greenberg, *The collected essays and criticism, Vol. 2, Arrogant Purpose, 1945-1949* ed. J. O'Brian (Chicago: Univ. of Chicago Press, 1986), p. 326; W. Haftmann, *Malerei im 20. Jahrhundert. Eine Entwicklungsgeschichte* (München: Prestel 1954); B. Newman, »The sublime is now« [1947], in: id., *Selected Writings and Interviews*, ed. J. P. O'Neill (Berkeley: Univ. of California Press, 1992), p. 173.
3. Cf. A. Hoormann, *Die Entfernung der Inhalte aus dem Bild. Zur Konstruktion des "unschuldigen Sehens" im Abstrakten Expressionismus*, in: id. *Medium und Material. Zur Kunst der Moderne und der Gegenwart* (München: Fink 2007), p. 57.
4. C. E. Osgood, G. Suci & P. Tannenbaum, *The measurement of meaning* (Urbana: University of Illinois Press, 1957).

---

## *Diskussion*

Ziel der Studie war es, zu prüfen, inwiefern Thesen zur universalen Ausdruckskraft von abstrakter Malerei in der Praxis der Auseinandersetzung mit abstrakter Kunst eine Entsprechung finden. Im Konkreten wurde zum einen untersucht, ob abstrakte Kunstwerke eine höhere Übereinstimmung im Erleben von Ausdruck hervorrufen als in formaler Hinsicht vergleichbare gegenständliche Arbeiten; zum anderen wurde untersucht, inwiefern eine solche mögliche Übereinstimmung sich auch im Vorgang der Betrachtung widerspiegelt, ob also auch der ganz konkrete Blick aufs abstrakte Bild eine intersubjektive Homogenisierung erfährt.

Das Vorgehen der Studie reflektiert den in der Fragestellung angelegten Vergleich von abstrakten und gegenständlichen Kunstwerken, indem die zwei Werkgruppen gegenübergestellt gestellt werden. Im Rahmen einer Vorstudie, die aufgrund des knappen Textumfangs in der Publikation lediglich erwähnt, nicht aber näher erläutert worden ist, wurden eine Reihe von abstrakten und gegenständlichen Werken zu Paaren gruppiert und 28 Kunsthistoriker\_innen vorgelegt, welche die Werkpaare hinsichtlich ihrer formalen Ähnlichkeit bewertet haben. Die fünf Bildpaare, die letztendlich das Bildmaterial der Studie gebildet haben, wurden von den befragten Expert\_innen als (in formaler Hinsicht) am ähnlichsten befunden (siehe Abbildung 16). Eine hinreichende formale Ähnlichkeit zwischen den verwendeten abstrakten und gegenständlichen Kunstwerken war deswegen für die Untersuchung relevant, weil durch diese Ähnlichkeit gewährleistet ist, dass gefundene Unterschiede auf den verbleibenden wesentlichen Unterschied zwischen den Arbeiten zurückzuführen ist: die Abstraktion. Gefundene Unterschiede können in einem solchen Vergleich also nicht auf abweichende formale Kriterien zurückgeführt werden, sondern stehen tatsächlich in Zusammenhang mit der interessierenden Größe.



Bildpaar 1

Bildpaar 2



Bildpaar 3



Bildpaar 4



Bildpaar 5

Abbildung 16 | Bildmaterial der Studie „Abstract Art as a Universal Language?“ (**Bildpaar 1:** Kandinsky, Wassily (1935). *Kontrast mit Begleitung* (97,1 x 162,1 cm), Foto: Artstor, © Guggenheim Museum New York & Vallotton, Félix (1895). *Straßenszene*, Foto: Artstor, Privat; **Bildpaar 2:** Pollock, Jackson (1952). *Convergence* (241,9 x 399,1 cm), Foto: Artstor, © Albright-Knox Art Gallery & Delacroix, Eugène (1837). *The Battle of Taillebourg* (489 x 454 cm), Foto: Artstor, © Musée de l'Histoire de France (Versailles); **Bildpaar 3:** Kandinsky, Wassily (1926). *Several Circles* (140,4 x 140,7 cm), Foto: Artstor, © Guggenheim New York & Bosschert, Ambrosius (1606). *Flowers in a Glass* (35,6 x 29,3 cm), Foto: Artstor, © Cleveland Museum of Art; **Bildpaar 4:** Motherwell, Robert (1946-47). *Western Air* (182,9 x 137,2 cm), Foto: Artstor, © The Museum of Modern Art New York & Cézanne, Paul (1879-82). *Stilleben mit Karaffe, Milchkanne, Kaffeetasse und Orange* (26,67 x 34,92 cm), Foto: Artstor, © Dallas Museum of Art; **Bildpaar 5:** Götz, Karl Otto (1953). *Bild vom 5.2.1953* (125 x 90 cm), Foto: Artstor, © Staatliche Museen zu Berlin & Nolde, Emil (1910). *Tänzerin in rotem Kleid* (34,8 x 28,8 cm), Foto: Artstor, © Kunsthalle in Emden.



---

Auf der Basis der gewählten Bildbeispiele wurden zwei weitestgehend getrennte Untersuchungen vorgenommen: zum einen die Untersuchung des erlebten Ausdrucks und zum anderen die Untersuchung des konkreten Betrachtungsvorgangs. Während ersteres über ein sogenanntes semantisches Differential erfasst wurde, wurde der Betrachtungsvorgang durch Eye-Tracking-Aufzeichnungen abgebildet. Als semantisches Differential bezeichnet man eine Befragungsmethode, der eine mehrstufige Skala mit zwei gegensätzlichen Adjektiven an den Polen zugrunde legt. Die Befragten würden zum Beispiel gebeten, einzuschätzen, ob sie das betrachtete Werk als eher *aktiv* oder als eher *passiv* empfinden. In dieser Studie wurden den Befragten 13 Adjektiv-Paare vorgelegt (vgl. Tabelle 4). Diese wurden mit Hilfe einer qualitativen Vorstudie ausgewählt (20 Personen), welche zum Ziel hatte Beschreibungen für die gewählten Bildbeispiele zu ermitteln, deren Passung als intersubjektiv-geteilt verstanden werden kann (siehe auch folgendes Studienbeispiel).

*Begriffspaare semantisches Differential*

---

aktiv-passiv  
aufgeregt-gelassen  
dynamisch-ruhig  
einnehmend-abstossend  
freundlich\_unfreundlich  
heiter-duester  
klar-diffus  
lastend-luftig  
angespannt-entspannt  
schwer-leicht  
warm-kalt  
bewegt-statisch  
chaotisch-harmonisch

---

Tabelle 4 | Begriffspaare semantisches Differential, eigene Darstellung.

---

Da Eye-Tracking als Methode im nächsten Abschnitt sehr ausführlich besprochen wird, verzichte ich an dieser Stelle auf eine genauere Ausführung. Zu erwähnen ist aber dennoch, dass, wie schon in der Diskussion der vorherigen Studie deutlich geworden ist, sich die menschliche Wahrnehmung kaum auf einzelne Parameter reduzieren lässt. Sowohl Werkaspekte als auch bestimmte persönliche Rahmenbedingen, wie etwa Expertise, nehmen Einfluss auf die Wahrnehmung von und Auseinandersetzung mit Kunst. Um diesem Umstand Rechnung zu tragen, wurden die Eye-Tracking-Daten um zusätzliche Kriterien ergänzt. Um beispielsweise abbilden zu können, ob die Betrachtung von abstrakten Werken homogener zwischen den Versuchsteilnehmer\_innen ausfällt, hat die Untersuchung die Expertise der Befragten berücksichtigt – denn Expertise bedingt eine Veränderung des Zugangs zu Kunst. Die zuvor diskutierte Studie ist dafür nur ein Beispiel von vielen. Ebenfalls Berücksichtigung finden formale Bildaspekte. Hierzu wurden zwei Kriterien herausgegriffen: die Anzahl der dargestellten Elemente und die Leuchtkraft der Farben. Diese beiden Aspekte sind deswegen von Interesse für die Untersuchung, weil beide Einfluss auf den Betrachtungsvorgang nehmen können. Je mehr Elemente ein Werk enthält, desto mehr gibt es zu entdecken. Der Blick aufs Bild ist daher notwendigerweise breiter gestreut als dies für Kunstwerke der Fall ist, die in dieser Hinsicht übersichtlicher gestaltet sind. Um einen Vergleich aller zugrunde gelegten Werke zu ermöglichen (Abbildung 16), ist es daher nötig diesen Faktor zu berücksichtigen. Ähnliches gilt für die Leuchtkraft der Farben. In einigen psychologischen Studien konnte bereits gezeigt werden, dass der menschliche Blick von Signalfarben angezogen wird.<sup>141</sup> Basiert die Gestaltung einiger Werke also auf dem Einsatz von Farben mit starker Leuchtkraft, während dies bei anderen Werken nicht der Fall ist, so ist der Betrachtungsvorgang beider Werkgruppen nur vergleichbar, wenn der Einfluss der Farbgebung berücksichtigt wird. Der letzte Punkt, den ich bezüglich des Studiendesigns explizit herausgreifen möchte, ist die gesonderte Betrachtung der „Darstellung von Menschen“. Bei einem Vergleich von abstrakten und gegenständlichen Werken mag diese

---

<sup>141</sup> Z.B. Underwood/ Foulsham 2006, Itti/Koch 2000.

---

zusätzliche Unterscheidung zunächst etwas merkwürdig bzw. unnötig erscheinen. Tatsächlich hat sie aber einen sehr praktischen Grund: empirische Ergebnisse zum visuellen Explorationsverhalten zeigen sehr konsistente Ergebnisse hinsichtlich des Einflusses von menschlichen Darstellungen: stärker als jedes andere gegenständliche oder ungegenständliche Element zieht die Darstellungen von Menschen den Blick an.<sup>142</sup> Dies gilt insbesondere für die Gesichter der dargestellten (oder fotografierten) Menschen. Ein Vergleich mit der Betrachtung von gegenständlichen Darstellungen, welche einmal Menschen darstellen und einmal nicht, ist also ebenfalls nur möglich, wenn diese Unterschiede explizit Berücksichtigung finden.

Die Ergebnisse der Studie weisen nicht unbedingt auf die Haltbarkeit der Thesen zur Wahrnehmung von abstrakter Kunst hin: Die tatsächliche Auseinandersetzung mit abstrakter Kunst weist keine besonders ausgeprägte Universalität der Erfahrung auf. Zum einen zeigt sich, dass *alle* Werke, ganz gleich, ob abstrakt oder gegenständlich, bestimmte Qualitäten haben, die in großer Übereinstimmung wahrgenommen und bewertet werden. Es handelt sich hierbei also ganz explizit nicht um spezifische Qualitäten, die *nur* abstrakten Arbeiten eigen sind.<sup>143</sup> Zum anderen zeigt sich, dass auch der Betrachtungsvorgang, der ganz konkrete Blick aufs Bild, welcher die Grundvoraussetzung und den Startpunkt für die Wahrnehmung und Auseinandersetzung mit dem Werk darstellt, im Fall von abstrakten Werken nicht als einheitlicher oder gerichteter zu bewerten ist, als dies bei gegenständlichen Arbeiten der Fall ist. Das Gegenteil ist der Fall: die hinsichtlich ihres formalen Aufbaus vergleichbaren gegenständlichen Kunstwerke werden im Vergleich deutlich einheitlicher und auch gerichteter betrachtet. Bei der Betrachtung abstrakter Werke richten die Betrachtenden ihre Aufmerksamkeit also ganz grundsätzlich weniger eindeutig auf bestimmte Bildaspekte als im Fall von gegenständlicher Kunst. Darüber hinaus ist die jeweilige Aufmerksamkeit bei der Betrachtung von abstrakter Malerei auch weniger häufig an Punkten konzentriert, welche

---

<sup>142</sup> Z. B. Bindemann et al. 2005, Ro/Friggel/Lavie 2007, Langton et al. 2008, Thoma/Lavie 2013.

<sup>143</sup> Eine tiefergehende Besprechung dieses Umstandes findet sich auch in der im Folgenden besprochenen dritten Studie.



---

als intersubjektiv-geteiltes gemeinsames Interesse verstanden werden können. Stattdessen wählen die Betrachtenden hier viel individuellere Betrachtungsstrategien und setzen ihr Augenmerk auf subjektiv sehr verschiedene Bildaspekte und Regionen.

Da sowohl die Expertise der Befragten als auch Aspekte des formalen Aufbaus in der Auswertung der Daten berücksichtigt wurden, kann dieses Ergebnis als unabhängig von den berücksichtigten Faktoren verstanden werden. Das heißt, dass die beschriebenen Ergebnisse in dieser Form sowohl auf Laien als auch Expert\_innen zutreffen und dass sie unabhängig von der Anzahl der dargestellten Elemente, der Leuchtkraft der verwendeten Farben und auch der Darstellung von Menschen sind. Gegenständliche Kunstwerke führen also grundsätzlich zu einer stärkeren Gerichtetheit des Blicks und zu einer größeren intersubjektiven Übereinstimmung hinsichtlich der interessantesten Bildregionen. Im Lichte der eigenen Erfahrung von der Betrachtung diverser Werke ist dieses Ergebnis kaum verwunderlich. Bei der Betrachtung eines gegenständlichen Werkes ist in der Regel offensichtlich, welches die wichtigsten (bedeutungstragenden) Elemente sind – beispielsweise menschliche Figuren. Sobald diese gesehen und identifiziert wurden, können weitere Elemente wie etwa Attribute oder die räumliche Einbettung sinnstiftend in die Betrachtung miteinbezogen werden. Das erschließen einer Szene folgt demnach einem logischen Rahmen, den wohl die meisten Betrachtenden in sehr ähnlicher Form bedienen. Ein vollständiges Abweichen von diesem Rahmen, wie beispielsweise das vollständige Ignorieren menschlicher Figuren zugunsten einer ausschließlichen Betrachtung von Hintergrundlandschaften ist für den Zugang zu einem gegenständlichen Werk schlicht unwahrscheinlich. Einem abstrakten Kunstwerk hingegen fehlen Fixpunkte, die als intersubjektiv verbindliche Bedeutungsträger markiert werden können. Es ist weniger deutlich, welche Bildelemente für den Zugang zum Werk notwendig oder bedeutsam sind. Die Betrachtung ist hier freier, weil sie nicht im Rahmen eines semantischen Korsetts stattfindet, sondern einen kontingenten Möglichkeitsraum eröffnet.

Diese freiere Betrachtung von abstrakter Malerei geht allerdings nicht, wie zuvor besprochen, mit einer weniger eindeutigen Wahrnehmung von Ausdruck einher. Die

---

Wahrnehmung von Ausdruck und der Betrachtungsvorgang scheinen demnach keiner direkten Kopplung zu unterliegen: Eine Höhere Übereinstimmung im Betrachtungsvorgang führt nicht zu einer höheren Übereinstimmung in der Wahrnehmung von Ausdruck.

### *Fazit und Beitrag der empirischen Methoden*

Für den Diskurs um die „Abstraktion als Weltsprache“ lässt sich aus den Ergebnissen der Studie schlussfolgern, dass Abstraktion keine bessere „Weltsprache“ darstellt als gegenständliche Ausdruckformen. Die Ergebnisse der Blickbewegungsuntersuchung legen darüber hinaus nahe, dass der jeweilige Zugang zu abstrakten Werken deutlich individueller gestaltet ist, als dies bei gegenständlichen Arbeiten der Fall ist. Es wäre sogar denkbar, dass bei einer detaillierteren Untersuchung des wahrgenommenen Ausdrucks zu Tage trete, dass dieser in der Auseinandersetzung mit gegenständlichen Kunstwerken homogener ausfällt als bei abstrakter Malerei. Da die Untersuchung des erlebten Ausdrucks in der Studie aber auf relativ grobe Kategorien begrenzt war, lassen sich hierzu auf der Basis der vorliegenden Ergebnisse keine weiteren Schlüsse ziehen. Wichtig ist jedenfalls festzuhalten, dass die Ergebnisse der Studie *nicht* bedeuten, dass abstrakte Bildelemente wie Linien, Farben oder Formenzusammenstellungen nicht einen in gewisser Hinsicht als universell zu verstehenden Ausdruck besitzen würden – immerhin zeigen die Ergebnisse der Untersuchung, dass abstrakte und auch gegenständliche Werke sehr übereinstimmend in ihrem jeweiligen Ausdruck wahrgenommen wurden. Dies mag durchaus daran liegen, dass bestimmte Bildelemente (die in beiden Werkgruppen gleichermaßen vorhanden sind) einen bestimmten Ausdruck transportieren oder auf Grund von bestehenden Darstellungs- und Interpretationskonventionen nahelegen. Den Ergebnissen der Studie ist deswegen lediglich zu entnehmen, dass es nicht die *Abstraktion* ist, die eine größere Einigkeit oder Uneinigkeit in erlebtem Ausdruck bedingt, sondern etwas anderes. Dieses Andere gilt es in zukünftigen theoretischen und empirischen Arbeiten zu beschreiben und zu klassifizieren.<sup>144</sup>

---

<sup>144</sup> An der Universität Wien wird derzeit zu diesen Fragen ein interdisziplinäres Forschungsprojekt mit dem Titel *Universal aesthetics of lines and colors? Effects of culture, expertise, and habituation*

---

Ähnlich wie auch schon im Falle der vorangegangenen Studie wären viele der hier erzielten Ergebnisse und Beobachtungen ohne den Rückgriff auf empirische Methoden nicht leistbar gewesen. Ohne einen auf quantifizierbare Strukturen bedachten Blick hätten intersubjektive Wahrnehmungsübereinstimmungen und –unterschiede nicht abgebildet werden können. Ohne den spezifischen Vergleich von abstrakten und gegenständlichen Werken unter weitestgehendem Ausschluss formaler Unterschiede wäre ein spezifisches Adressieren der *Abstraktion* als eigenständiger Einflussfaktor auf die Erfahrung von Ausdruck nicht möglich gewesen. Auf noch viel spezifischere Weise von einem empirischen Zugang abhängig, als es die Erfassung von Bewertungen und Urteilen ist, ist das Abbilden des tatsächlichen Vorgehens bei der Betrachtung von Kunst. Bewertungen und Urteile sind im Wesentlichen entweder als kondensiert-abstrakte Reduktionen von bewusster Erfahrung zu verstehen oder leiten sich direkt aus bewusster Erfahrung ab. Sie sind damit der subjektiven Erfahrung direkt zugänglich und können daher auch *direkt* von Versuchsteilnehmenden erfragt werden.<sup>145</sup> Das Vorgehen bei der Betrachtung von einem Werk, also die tatsächlich vollzogenen Blickbewegungen, die unserer Wahrnehmung zugrunde liegen, sind der Erfahrung hingegen *nicht* direkt zugänglich. Nicht nur sind die meisten Bewegungen unserer Augen keine willentlichen Akte – manche Augenbewegungen vollziehen sich schlicht zu schnell um unserer bewussten Erfahrung überhaupt zugänglich zu sein.<sup>146</sup> Der Nutzen von empirischen Messverfahren ist in diesem Fall also noch offensichtlicher: Der Betrachtungsvorgang bei der Auseinandersetzung mit Kunst lässt sich ohne einen empirischen Zugang zur Frage überhaupt nicht korrekt adressieren. Darüber hinaus sind Regelmäßigen in Blickbewegungsaufzeichnungen, betrachtet man sie mit dem bloßen Auge, etwa in Form einer Visualisierung im Videoformat,<sup>147</sup> nur schwerlich auszumachen. Oft wirkt der Blick

---

durchgeführt (Projektleitung Prof. Dr. Raphael Rosenberg & Prof. Dr. Helmut Leder).

Abschließende Ergebnisse liegen noch nicht vor (vgl. auch

<https://linesandcolors.univie.ac.at/publikationen-und-konferenzbeitraege/>, Zuletzt abgerufen am 08.05.18).

<sup>145</sup> Vgl. Diekmann 2007.

<sup>146</sup> Für eine detaillierte Besprechung verschiedener Blickbewegungstypen und deren zeitliche Dimensionen siehe Holmqvist et al. 2011.

<sup>147</sup> Vgl. Brinkmann/Commare 2015.

---

eines Einzelnen eher unstet, beinahe erratisch.<sup>148</sup> Erst beim Vergleich mehrerer Personen, lassen sich überhaupt Strukturen und gewisse Regelmäßigkeit antizipieren. Solche Regelmäßigkeiten sind dem bloßen Auge oft nicht einmal erkenntlich, wie beispielsweise die feinen graduellen Abstufungen im Streuungsgrad unserer visuellen Aufmerksamkeit. Um also Unterschiede in Aufmerksamkeitsverteilungsstrukturen bei der Betrachtung von Kunstwerken überhaupt beobachten zu können, ist der Rückgriff auf sowohl empirische Messungen als auch statistische Auswertungsverfahren notwendig. Dabei geht es ganz explizit nicht um die Quantifizierung von anderweitig ebenfalls erzielbaren Ergebnissen und deren breitflächige Validierung, sondern um die grundsätzliche Adressierbarkeit von für den Fragenraum relevanten Aspekten.

---

<sup>148</sup> Vgl. Betz et al. 2010.

---

# Studie

## Aesthetic Echoes in the Beholder's Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art

LITERATURANGABE

---

Laura Commare/ Hanna Brinkmann, Aesthetic Echoes in the Beholder's Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art, in: Michael Zimmermann (Hg), Vision in Motion. Streams of Sensation and Configurations of Time, Berlin 2016, S. 221-234.

---

### 2.3 *Aesthetic Echoes in the Beholder's Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art*

Die dritte und letzte Studie, die ich vorstellen möchte, ist als Aufsatz in einem von Michael Zimmermann herausgegebenen Sammelband zum Thema *Bewegtes Sehen (Vision in Motion)* erschienen. Der Text trägt den Titel *Aesthetic Echoes in the Beholder's Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art* und ist in Zusammenarbeit mit Hanna Brinkmann entstanden.<sup>149</sup> Ähnlich wie in der vorangegangenen Studie ist die zentrale Frage dieser Untersuchung eine der Literatur entlehnte These, nämlich die Frage nach der qualitativen Entsprechung von kognitiv wahrgenommenen Bildqualitäten und tatsächlichem Betrachtendenverhalten. Auch in diesem Fall handelt es sich um ein theorietestendes Vorgehen, das Annahmen zur Erfahrung von abstrakter Kunst herausgreift, welche im Verlauf eines kunstwissenschaftlichen Diskurses postuliert wurden, und diese hinsichtlich ihrer realweltlichen Praxisrelevanz untersucht. Wie auch die letzte Studie ist die nachfolgende Untersuchung als Kommentar zu bestehenden Positionen zu verstehen. Da dieser Text nicht gar so knapp ausfallen musste wie das vorhergehende Beispiel, finden sich in der Publikation etwas ausführlichere Erläuterungen zu den untersuchten Ideen und vor allem zu den zur Anwendung kommenden Methoden. Ich werde mich daher für Einführung und Diskussion der Ergebnisse im Folgenden auf wenige Worte beschränken und jeweils nur einen knappen Überblick zu den wichtigsten Punkten geben.

---

<sup>149</sup> Die Mehrfachautor\_innenschaft des Textes ist durch das arbeitsteilige Vorgehen in der Vorbereitung, Durchführung und Nachbereitung der empirischeren Untersuchung bedingt. Die unterschiedlichen Arbeitsschritte sind dabei schwerpunktmäßig von jeweils bestimmten Teilen des Teams bearbeitet worden. Eine ausführliche Aufstellung der jeweiligen Aufgaben und Einzelleistungen findet sich im Anhang (Anhang *Stellungnahme zur Arbeitssteilung III*).

---

## Einleitung

Die Beschäftigung mit abstrakter Kunst, insbesondere mit Werken des abstrakten Expressionismus geht häufig Hand in Hand mit wirkungsästhetischen Überlegungen. Besonders eindrücklich wird dies etwa am Beispiel des Diskurses um Jackson Pollocks Arbeiten deutlich. Seine Gemälde werden häufig als *action paintings* beschrieben.<sup>150</sup> Der Fertigungsprozess dieser Arbeiten ist durch eine überbordende Aktivität und Dynamik geprägt, die sich unter anderem darin äußert, dass Pollock hier – häufig unter Einsatz seines ganzen Körpers – Farbe auf riesige am Boden liegende Leinwände tropfen, fließen und spritzen lässt. Diese für den Fertigungsprozess so bezeichnende Dynamik schreibt sich in das Werk ein, wird zu dessen essentieller Qualität. Diese Qualität, so die Annahme, wird von Betrachtenden nicht nur erfahren, sondern wirkt sich überdies auch auf ihren konkreten Zugang zum Bild aus. Ulf Küster beispielsweise nimmt an, dass die Erfahrung von Aktivität und Dynamik in Pollocks Arbeiten so groß ist, dass die Betrachtenden gar nicht anders können, als physisch mit ihrem Blick darauf zu reagieren. Dazu äußert er die explizite Annahme: „There is no ‘action painting’ without ‘action viewing.’“<sup>151</sup> In dieser Annahme drückt sich die Idee einer qualitativen Entsprechung von Bildqualität und Betrachtungsmodalität aus – eine Vorstellung, die sich auch metaphorisch als ästhetischer Wiederhall, als ästhetisches Echo im Auge der Betrachtenden begreifen lässt.

Den Überlegungen Küsters nicht ganz unähnlich ist auch Gottfried Boehms Position. Boehm spricht dabei jedoch nicht, wie Küster, über die bildimmanente Dynamik von Pollocks Werken, sondern bezieht sich auf das in den Gemälden Pollocks häufig ebenfalls zur Anwendung kommende Konzept des *all-over paintings*. Die größtenteils undifferenzierte und flächendeckende Behandlung der Leinwand führt dazu, dass sich in den Werken keine dominanten Bildausschnitte herausbilden, auf die sich aus inhaltlichen oder auch formalen Gründen die Aufmerksamkeit der Betrachtenden konzentrieren könnte. Für den konkreten

---

<sup>150</sup> Rosenberg 1952: 22.

<sup>151</sup> Küster 2008: 19.

---

Anwendungsfall von Pollocks Arbeiten schreibt Boehm: „Die visuellen Vektoren verfestigen sich nirgendwo zu kompositionellen Ordnungen.“<sup>152</sup> Wie Küster schließt auch Böhm, dass diese Bildqualität einen physischen Wiederhall im Auge der Betrachtenden zur Folge hat. Wörtlich schreibt er dazu: “Pollocks Bilder bieten keine Durchblicke an, sondern fordern ein schweifendes Auge.”<sup>153</sup>

Die von Küster und Boehm geäußerten Annahmen zum Blickverhalten bei der Betrachtung von Pollocks Werken stellen starke Thesen über die Praxis menschlicher Wahrnehmung dar. Sie sind insofern also vergleichbar mit dem Kontext der letzten Studie. In beiden Fällen ist ohne Praxistest letztlich nicht zu klären, ob mit den jeweiligen Annahmen praxisrelevante, rezeptionsästhetische Konzepte formuliert sind. Für das vorliegende Beispiel ist dies sogar noch problematischer als im vorherigen Studienbeispiel, da hier konkrete, sich physisch manifestierende Handlungen Gegenstand der Argumentation sind. Die Erfahrung von Bildqualitäten wird, zumindest implizit, sowohl in Küsters Argument als auch in Boehms Argumentation, an ganz bestimmte körperliche (Blick-)Reaktionen gebunden. Ob eine solche Entsprechung von Blickverhalten und erfahrender Bildqualität aber tatsächlich vorhanden ist, ob es wirklich notwendig ist, im Bild dargestellter Aktivität oder Dynamik mit einem schweifenden Auge zu begegnen, ist alleine aufgrund der vermeintlichen Plausibilität dieser Annahmen nicht geklärt. Dies gilt umso mehr, als der Blick selbst, also die Voraussetzungen und Rahmenbedingungen des ‚*Sehens*‘, in den jeweiligen Positionen nicht thematisiert werden. Das hat zur Folge, dass letztlich unklar bleibt, was genau unter *action viewing* oder einem *schweifenden Auge* tatsächlich verstanden werden kann und soll. Eine empirisch-informierte Definition der Konzepte, sowie eine empirische Prüfung ihrer Verankerung in der Praxis menschlichen Handelns, ist somit schlicht notwendig, um zu den postulierten Thesen überhaupt Position beziehen zu können.

---

<sup>152</sup> Boehm 2008: 41.

<sup>153</sup> Ebd.



---

### *Bemerkungen zur Studie*

In der folgenden Studie wird zum einen ein *Vorschlag* für eine empirisch-informierte Definition der Konzepte *action viewing* und *schweifender Blick* entwickelt. Zum anderen erfolgt auf der Grundlage dieser Definitionen eine empirische Überprüfung der von Küster und Boehm postulierten Thesen. Der Fokus der Untersuchung liegt dabei auf dem Einfluss der Erfahrung bildimmanenter Aktivität und Dynamik auf das Blickverhalten, also auf der Frage ob – und wenn ja, wie – wahrgenommene Bildqualitäten, wie z.B. Dynamik, tatsächlich einen *schweifenden Blick* oder *action viewing* zur Folge haben. Die Frage nach dem Einfluss, welchen die Erfahrung einer *all-over Struktur* auf den Blick der Betrachtenden nehmen kann, wird in der folgenden Untersuchung nicht explizit aufgegriffen. Sie wird nur indirekt über eine Untersuchung des Einflusses von formaler Bildkomplexität auf das Blickverhalten adressiert. Die Ergebnisse der Untersuchung können somit nur bedingt Aufschluss auf die konkrete Frage nach dem Einfluss von *all-over painting* Konzepten im Kontext rezeptionsästhetischer Überlegungen geben.

Zur Prüfung dieser Thesen werden, wie auch schon in der zweiten Studie, zwei Maße herangezogen: Zum einen die ästhetischen Urteile der Teilnehmenden, die einen Rückschluss über das Erleben von Ausdruck bzw. erlebter Bildqualität erlauben; zum anderen Eye-Tracking-Aufzeichnungen, über welche die Analyse und der Vergleich des konkreten Blickverhaltens möglich ist. Methodisch gesehen stellt das folgende gegenüber dem zuletzt vorgestellten Studienbeispiel dennoch eine Erweiterung dar: Teil des hier umgesetzten Studiendesigns ist, über die genannten Aspekte hinaus, auch eine Replikationsstudie. Über die Replikation von empirischen Ergebnissen, werden grundsätzliche Fragen nach der Tragfähigkeit und Generalisierbarkeit von in Einzelstudien generierten Ergebnissen adressierbar. In der späteren gemeinsamen Diskussion aller drei vorgestellten Studienbeispiele wird daher insbesondere diese dritte Studie als ein Beispiel für Probleme und Grenzen empirischen Arbeitens dienen.

**Laura Commare and Hanna Brinkmann**

**Aesthetic Echoes in the Beholder's Eye?  
Empirical Evidence for the Divergence of Theory and Practice  
in the Perception of Abstract Art**

Artworks by Jackson Pollock are often described as *action paintings*.<sup>1</sup> This is due to the active and dynamic production process, which sometimes involved the entire body of the artist. The exuberant activity of the production is mirrored in the essential qualities of the paintings—reflecting strong intrinsic dynamics. In this context Ulf Küster stated, “There is no ‘action painting’ without ‘action viewing,’”<sup>2</sup> intending to say that the impression of action in Pollock’s paintings is so strong that one cannot simply view his artworks without being forced to react physically to them: the bustling overspray of paint is literally mirrored in the perception process. The idea of a reflection of the affective characteristics of Pollock’s artworks in the beholder’s gaze, an aesthetic echo in the eye, has also been expressed by Gottfried Boehm. Boehm refers to the concept of *all-over painting*, stating that due to the non-differential treatment of the surface, Jackson Pollock’s artworks lack a dominant point of interest. They therefore demand a rambling gaze.<sup>3</sup>

Despite their convincing theoretical reasoning, Küster’s and Boehm’s assumptions only scratch the surface of a complex problem. It is obvious that different paintings do have different characteristics, such as dynamism and activity, which in turn affect our experience. Yet, there is more to the seemingly simple process of experiencing a painting than meets the eye. Art perception evolves in the physical process of seeing (eye movements) and in the resulting cognitive and emotional reactions to the artwork. Vision, the starting point of every visual experience, is a pervasive and mostly unconscious activity. Therefore, it often appears effortless and facile. It is often seen as a habitual operation, without ever reflecting on the actual complexity of the involved processes.

1 H. Rosenberg, “The American Action Painters,” *ARTnews* 51, no. 8 (December 1952): 22.

2 U. Küster, “Action Painting—Mythos und Realität,” in *Action Painting: Jackson Pollock*, ed. Fondation Beyeler (Ostfildern: Hatje Cantz Verlag, 2008), 12–19, 19.

3 “Pollocks Bilder bieten keine Durchblicke an, sondern fordern ein schweifendes Auge. Die visuellen Vektoren verfestigen sich nirgendwo zu kompositionellen Ordnungen.” Gottfried Boehm, “Die Form des Formlosen—Abstrakter Expressionismus und Informel,” in Fondation Beyeler, *Action Painting: Jackson Pollock*, 38–46, 41.

The concepts of *action viewing* and *rambling gaze* subsume a series of aspects of human vision. An active and dynamic perception mode is not only related to the gaze's speed, but also to the directedness of visual attention and the depth of the exploration. Moreover, it is not yet clear how these aspects are related to perceived artwork characteristics. Is there such a thing as an aesthetic echo in the beholder's eye? And if yes, what does it actually look like? Which aspects of the vision process are involved in the concept of *action viewing*? Our study investigates whether there is a kind of qualitative equation of affective image characteristics and visual exploration detectable in the actual process of art perception.

In order to reliably test the assumptions stated by Küster and Boehm it is necessary to empirically measure the process of visual perception and the beholder's experiences. This can be done by the means of standardized questionnaires and eye-tracking techniques, which capture the actual motion of the eye. Eye trackers have been used in research since the late nineteenth century. The first eye trackers were mostly mechanical and very painful for the participants, who had to literally attach the hardware to the eyeball.<sup>4</sup> Nowadays, eye tracking techniques operate in a completely different manner. The dominating method to measure the gaze is the so-called video-based pupil-and-corneal-reflection method. It uses infrared cameras (for example, integrated in a computer screen) to illuminate and track the pupil's positions in the measured dihedral angle. Geometrical calculations together with a calibration procedure, which matches a certain eye position to a singular point on the perceived object, are used to map the gaze.<sup>5</sup>

### Measuring the Motion of the Eye

Visual exploration depends not only on the perceived object, but also on the subject observing. Previous studies by Rosenberg and colleagues about the perception of artworks provide empirical evidence that beholder-specific characteristics, such as art expertise, result in divergent, group-specific gaze patterns.<sup>6</sup> The authors also found that

4 K. Holmqvist et al., *Eye Tracking: A Comprehensive Guide to Methods and Measures* (New York: Oxford University Press, 2011), 9.

5 *Ibid.*, 26.

6 J. Betz et al., "Dem Auge auf der Spur. Eine historische und empirische Studie zur Blickbewegung beim Betrachten von Gemälden," *IMAGE* 11 (2010); R. Rosenberg and C. Klein, "The Moving Eye of the Beholder. Eye Tracking and the Perception of Paintings," in *Art, Aesthetics and the Brain*, ed. J. P. Huston et al. (New York: Oxford University Press, 2015); R. Rosenberg, "Dem Auge auf der Spur. Blickbewegungen beim Betrachten von Gemälden—historisch und empirisch," *Jahrbuch der Heidelberger Akademie der Wissenschaften für 2010* (2011): 76–89; R. Rosenberg, J. Betz, and C. Klein, "Augensprünge," *Bildwelten des Wissens. Kunsthistorisches Jahrbuch für Bildkritik* 6 (2008): 127–29. For other attempts that incorporate eye

image characteristics, such as compositional aspects, structure gaze patterns systematically. Knowing about this two-layered dependency of the gaze nevertheless leaves the essential question unanswered: to what degree is the actual art experience detectable in the visual exploration?

In one of our previous studies we show that the distribution of visual attention is affected by the degree of abstraction.<sup>7</sup> We tested the concept of abstract art as a universal language. In the postwar era abstract art was claimed to be a world language because it enables the beholder to experience the pure value of lines and colors as such, without being affected by any kind of prior knowledge or cultural background. If this holds true, abstract art can be interpreted as art that elicits universal understanding, or, so to speak, a universal experience. This should be reflected in the perception and valuation process itself, namely, by a more homogeneous beholding process in the perception of abstract art compared to representational art. In order to compare the perception processes, we matched selected abstract artworks with representational counterparts, which are similar with regard to formal criteria, such as dynamics, composition, and color. With this procedure we ensured that differences in the perception process appear not due to differences in artworks' characteristics, except for the degree of abstraction of the paintings. We found no evidence for the hypothesis that abstract art induces more coherent beholder behavior. Instead, our results show that abstract artworks consistently elicit a clearly more diffuse attention distribution compared to their representational counterparts. This finding contradicts the stated hypothesis of abstract art as a universal language. With regard to evaluative responses we found no significant differences in the coherence of answers. We conclude from these results that people seem to agree on the essential qualities of an artwork on a cognitive, evaluative level, whether the painting is abstract or representational. However, there are large differences in the attention distributions across the studied art categories (abstract vs. representational).<sup>8</sup> We can infer from these results that coherent cognitive experience is not necessarily dependent on coherent perception processes: different perceptual procedures can still lead to the same impression.

tracking methods in art-history research, see also C. Wagner, "Der Mythos vom 'unschuldigen Auge'. Blickbewegungen, Wahrnehmungsprozesse in der Bildenden Kunst," in *Aisthesis: Wahrnehmungsprozesse und Visualisierungsformen in Kunst und Technik*, ed. M. Greenlee, C. Wagner, and C. Wolff (Regensburg: Schnell & Steiner, 2013), 31–50; Christoph Wagner, "Mapping the Eye. Blickbewegungen und Verstehensprozesse in der Kunst," in *Bilder sehen: Perspektiven der Bildwissenschaft*, ed. M. Greenlee, C. Wagner, and C. Wolff (Regensburg: Schnell & Steiner, 2013), 69–86.

7 H. Brinkmann et al., "Abstract Art as a Universal Language?," *Leonardo* 42 (2014): doi:10.1162/LEON-a-00767.

8 See also D. Massaro et al., "When Art Moves the Eye: A Behavioural and Eye Tracking Study," *PLOS ONE* 7 (2012): e37285. doi:10.1371/journal.pone.0037285.



In the previous study we did not take artwork-specific affective characteristics into account for the analysis of eye movements. Instead we focused on general differences in visual exploration between abstract and representational art. Abstract art by definition does not depict recognizable objects, but instead uses shapes and colors for their own expressive value in a nonrepresentational way.<sup>9</sup> Yet the term abstract covers a variety of very different kinds of works of art. A painting by Vasily Kandinsky, for example, evokes a different effect than an artwork by Jackson Pollock. As shown in our previous study, there seems to be high agreement on the affective characteristics of an artwork. This is a basic prerequisite for the assumption that the characteristics of an artwork may be able to systematically elicit certain experience types and corresponding perception modes (eye movements)—action viewing will only arise when we experience the observed object as active.

### The Process of Visual Perception

Visual perception can be defined as the process of interpreting surrounding environments by using information derived from visible light. Light enters the eye through the pupil; the lens turns the image upside-down and then projects it onto the retina. The light-sensitive cells on the retina transform the incoming light waves into electric signals and send the received information through the optic nerve to a region in the back of the brain, the visual cortex. Here the information is processed. Human vision is not able to capture an entire scene or object at once in full acuity. Instead, we only see a small region in our field of sight sharply. This is related to the fact that solely the fovea, a very small part of the retina, contains enough light-sensitive cells to reach a full acuity of vision. The region from which the fovea produces a sharp image covers only approximately 2° of the entire field of vision. To give a more comprehensible scale: 2° of a human's field of vision is roughly the size of a thumb nail as it appears at arm's length. Most objects we look at are far bigger than a thumbnail. In order to produce a sharp image of the entire scene or object the eye therefore has to jump rapidly from point to point. At each focused point the eye stands still for a short while to gather the relevant information. These rest points are called fixation points or fixations (first proved by Erdmann and Dodge in 1898).<sup>10</sup> A fixation lasts for an estimated minimum of 100–120

<sup>9</sup> M. Rosenthal, *Abstraction in the Twentieth Century: Total Risk, Freedom, Discipline* (New York: Guggenheim Museum Publications, 1996), 3.

<sup>10</sup> B. Erdmann and R. Dodge, *Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage* (Halle: Niemeyer, 1898).

milliseconds. The rapid jumps of the eye that connect the single fixations are much shorter. They only last for about 30–80 milliseconds, which is the fastest movement the human body can perform. This movement, the so-called *saccade*, is so fast that the light entering the eye during this time span cannot be processed properly. Therefore, we are mostly blind during a *saccade*.<sup>11</sup>

The most reported event in eye-tracking data is related to the moments when the eye remains still—the fixation points. Needless to say, even if a fixation measures only 2° of the field of vision, we are not blind in the blurred areas around the fixation point. This brings on the question: how should we interpret a fixation, if it is obviously not a measure of what we see, but only of what we see sharply? Fixations can be interpreted as a measure of attention.<sup>12</sup> We draw the greatest attention to the point that we can detect best. This becomes clear if we take the reading process as an example. In order to read a given text, our eyes need to scan line by line and sometimes even word by word. The word you are currently reading is also the word you are paying most attention to. Hence, with eye-tracking methods we are able to measure the attention distribution during the perception process.

The term *attention distribution* refers to a pattern of fixations, which is established over a certain perception time. The localization of attention is important, because it tells us something about where the beholder was looking. Yet, only taking into account the distribution of attention falls short of the mark. Attention might be diffuse or directed. A diffuse attention profile is defined as a distribution of fixations with no or only a few regions of major interest that are frequently looked at by the beholder. A focused or directed attention, on the other hand, describes a pattern with clear regions of interest. Besides the degree of directedness of attention, the modalities of the ongoing attention distribution may also differ. The beholder's gaze can be fast-paced and restlessly flying over the painting, or even-tempered and consciously wandering. A fast-paced and flying gaze with short fixations would correspond to the concept of *action viewing* described above. According to the idea of a quality equation in artwork and gaze, a consciously wandering gaze with long fixations occurs when the beholder is having a calm and ruminant art experience, as might be the case for paintings by Mark Rothko or Barnett Newman. Of course the gaze's speed is not an "either-or" choice between bustling and calm; it can take any gradation between these extremes.

We analyzed the connection between affective image characteristics and visual perception on two levels. First, we tested whether *attention distributions* occur in a more

11 Holmqvist et al., *Eye Tracking*, 23.

12 *Ibid.*, 22.

diffuse manner in dynamic paintings as opposed to rather passive artworks, as suggested by Boehm. Second, we investigated whether the *modality of attention distributions* and the speed of the gaze vary systematically with the perceived activity of a painting.

### Study Design

We studied the five abstract paintings that were employed in our previous study,<sup>13</sup> and extended the given set with five additional abstract paintings. We ensured that the additional paintings can be interpreted as companion pieces to the set of our previous study with respect to formal qualities, such as color, dynamics, and the number of depicted elements. The number of depicted elements can be interpreted as a measure of formal complexity.<sup>14</sup> The matching was done to ensure that our findings are replicable for all artworks with similar qualities and have not emerged by chance. In total, the selected artworks form a sample of ten paintings, dividable in five pairs with corresponding qualities. The set includes works by Karl Otto Götz, Vasily Kandinsky, Robert Motherwell, Sam Francis, and Jackson Pollock. None of the paintings was altered.

To ensure the reproducibility of our results, we conducted two separate experiments of identical type with a total of eighty participants, forty in each setting. For both experiments, the recruitment of participants was limited to students from the University of Vienna, only varying in age, gender, and expertise. While age and gender have not been systematically controlled, we balanced the number of experts and the number of laypeople. We measured expertise by taking into account the field of study of the participants and their answers to a detailed questionnaire about their art expertise. The questionnaire covers a broad variety of aspects, including questions that reveal the participants' knowledge about selected artists as well as questions about the frequency with which they conduct art-related activities in their leisure time, such as going to museums or talking about art with their friends.

In each experiment the forty participants were asked to look at five of the selected abstract paintings, which were presented in random order on a high-resolution computer screen (2560 x 1600 pixels) in the Laboratory for Cognitive Research in Art History in Vienna. Each artwork was presented for two minutes. While the participants were

13 Brinkmann et al., "Abstract Art as a Universal Language?"

14 D. E. Berlyne, *Aesthetics and Psychobiology* (New York: Appleton-Century-Crofts, 1971); D. E. Berlyne, J.C. Ogilvie, and L. C. C. Parham, "The Dimensionality of Visual Complexity, Interestingness, and Pleasingness," *Canadian Journal of Psychology* 22 (1968): 376–87; M. Nadal et al., "Visual Complexity and Beauty Appreciation: Explaining the Divergence of Results," *Empirical Studies of the Arts* 28 (2010): 173–91.

viewing the presented paintings, their eye movements were recorded with a binocular remote eye tracker (made by SMI). After the presentation of each artwork, participants were asked to answer a couple of questions concerning their art experience, starting with liking and familiarity ratings and continuing with a detailed thirteen-paired semantic differential. A semantic differential is a seven-point bipolar rating scale, which has two oppositional adjectives at the extreme values. It is used to interpret the participant's emotional and cognitive response to a given stimulus, in our case, artworks.<sup>15</sup> In order to obtain reasonable adjective pairs for the description of the essential affective characteristics of the presented artworks, we ran a small-scale pre-study. We asked twenty participants to describe how they felt about the artwork they were currently looking at and what kind of qualities they would attribute to it. The thirteen adjective pairs used in the main study were derived from the qualitative answers in this pre-study.

Most of the descriptive adjective pairs refer to qualities expressing dynamics, movement, and activity (e.g., active—passive, dynamic—calm, strained—unstressed, chaotic—harmonic). We combined all of these categories of the semantic differential in an activity measure. The higher the overall score of a painting on this scale, the more dynamic and active it is perceived to be. The scale ranges from 6 to 42 points, with a score of 6 referring to almost no perceived activity and a score of 42 referring to maximal perceived activity. In line with our previous findings, we found in both of our experiments that the participants highly agree in their ratings.<sup>16</sup> As expected, the paintings by Jackson Pollock (*Convergence*, 1952, Albright-Knox Art Gallery, and *Number 3*, 1950, Saint Louis Art Museum) were rated as the most active artworks, with 35.75 and 35.17 points on the activity scale. While most of the other artworks received average activity ratings below 30, there is one other painting in the set that was evaluated to be almost as active and dynamic as Jackson Pollock's artworks: Karl Otto Götz's *Bild vom 5.2.1953* (1953, Berlin State Museums). It reaches an average activity score of 33.94. The counterpart to *Bild vom 5.2.1953* was a work by Sam Francis from 1958 (*Untitled*, Norton Simon Museum of Art). Its average activity score is 31.81, which is very similar to the score for *Bild vom 5.2.1953*, showing that the matching was adequate. Also, the other matches show comparable activity-score ranges. We find the lowest scores for paintings by Robert Motherwell (*Western Air*, 1946–47, Museum of Modern Art; *The Little Spanish Prison*, 1941, Museum of Modern Art; and *The Garden Window*, 1969, Modern Art Museum of Fort Worth). They range from 19.29 to 20.97. The differences in the activity-rating scores across all

15 C.E. Osgood, G. Suci, and P. Tannenbaum, *The Measurement of Meaning* (Urbana: University of Illinois Press, 1957).

16 Brinkmann et al., "Abstract Art as a Universal Language?"



paintings allow us to test for activity-dependent differences in the visual exploration. For our analysis it is of clear virtue that at least one of the less formally complex artworks (with fewer depicted elements) in our set, *Bild vom 5.2.1953*, reaches activity scores comparable to Jackson Pollock's action paintings, which are fairly complex with regard to their formal characteristics. In order to test whether certain results ensue from perceived activity or if they depend on other image characteristics like complexity, we must have differences in both categories.

### Results

To measure the directedness of the gaze, we followed the procedure developed in our previous paper.<sup>17</sup> In contrast to classical approaches, we measured the distribution of attention with a bottom-up approach, focusing on the number of fixations per image section. Image sections are circular areas on the painting with a maximum radius of 90 pixels. They are not determined by a pre-defined location on the canvas. Their position rather depends on where the observed attention clusters. As a consequence, cluster locations may vary across beholders, because the focus of their interest possibly lies on different parts of the perceived image. Our measure allowed us to address the question of whether the visual exploration is less directed for artworks with high average activity scores compared to artworks with low average activity scores. Moreover, this measure enabled us to model the general directedness of the gaze, as we did not need to restrict the analysis to pre-defined areas identical for all participants. A directed attention can be measured by observing fixations that strongly cluster in certain image sections. A cluster of fixations is defined as a minimum of five fixations occurring in an image section per minute. The more clusters a person generates and the more distinct they are, the more directed is the visual exploration. To ease the interpretation of the results, we focused on the percentage of the total number of fixations that are located in a cluster: the higher the percentage of fixations in clusters, the more directed is the total attention distribution.

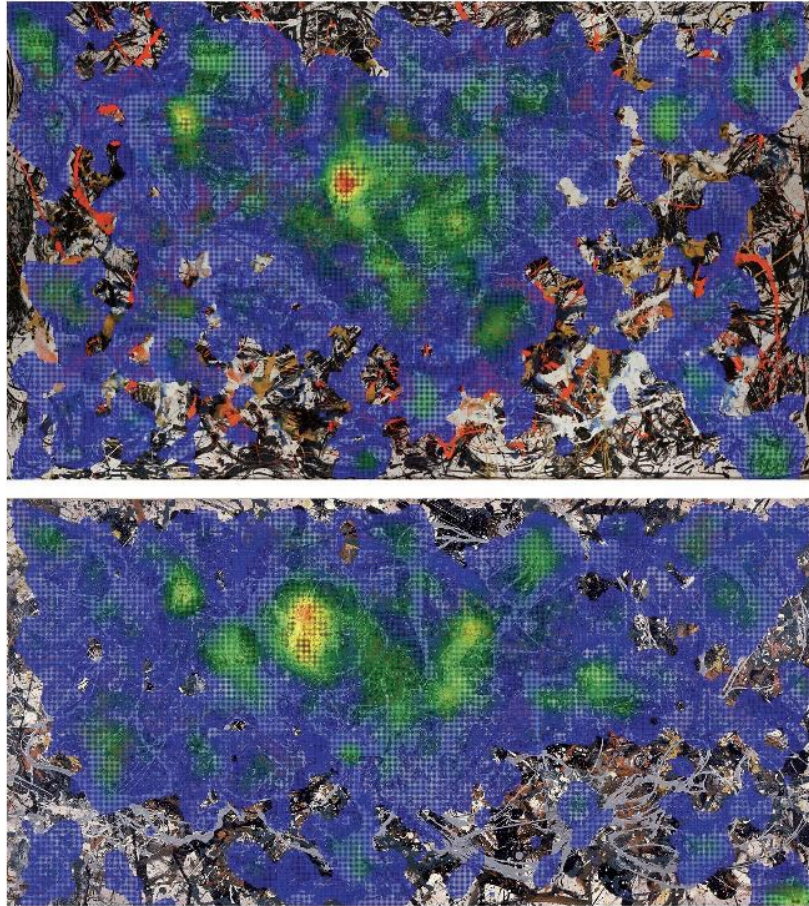
In a first step we analyzed the individual attention patterns, not taking into account whether the individual clusters resemble each other or not. For each experiment we ran a single mixed-effects linear model with random slopes for the average activity ratings and the formal complexity of the artwork. We found no significant effects in either of

<sup>17</sup> Ibid.

the groups for perceived activity, independent of the expert or lay status of the participant. In fact, the only aspect that significantly influenced visual attention patterns in both experiments is formal complexity, which is operationalized by the number of depicted elements in the paintings. For both experiments, highly complex paintings lead to significantly less attention clustering of over 25 percent ( $\beta = -34.89$ ,  $p > 0.001$  and  $\beta = -25.07$ ,  $p > 0.005$ ) as compared to simple paintings. We infer from these results that the perceived activity level of an artwork does not change individual exploration strategies with regard to attention-clustering rates.

This changes when it comes to shared attention clustering. Shared attention clusters are defined as image sections that, on average, receive high attention rates from the entire sample. They represent the collective distribution of attention. While the number and location of these image sections is constant for all participants, the percentage of the individual fixations located in these areas varies across the sample. High numbers of individual fixations located in shared clusters indicate homogeneous visual-exploration strategies. On the other hand, low numbers of individual fixations in shared clusters are a sign of heterogeneous visual-exploration strategies. Again employing mixed-effects linear models with random slopes for the activity ratings and the formal complexity of the artwork, we found that perceived activity has indeed a significant effect on the percentage of individual fixations in areas of shared interest in both groups ( $p > 0.001$  and  $p > 0.001$ ). The more active a painting is rated, the more heterogeneous is the perception process for all participants ( $\beta = -2.67$  and  $\beta = -0.56$ ), regardless of the formal complexity of the artwork and the participants' expertise level. Figures 1 and 2 visualize these results with so-called heat maps, a graphical representation of data that colors highly fixated areas in red, intermediately fixated areas in yellow and green, and rarely fixated areas in blue.

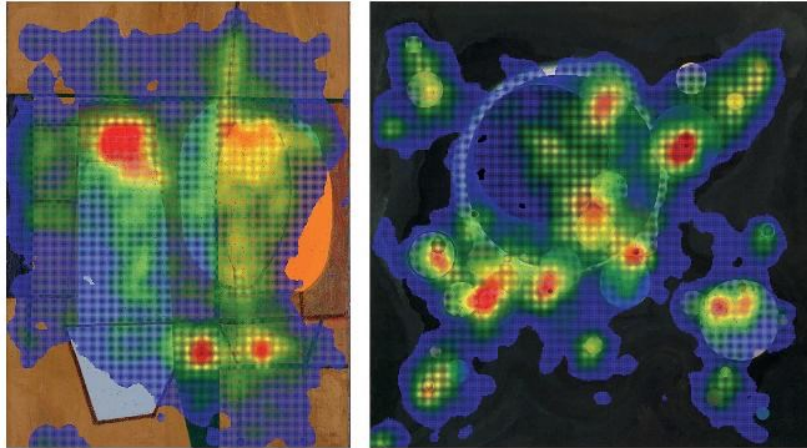
Although attention clustering takes place in the perception of abstract art, it is, as we showed in our previous paper, not very elaborated. The visualization threshold for highly fixated areas is three fixations per image section. Figure 1 shows two works by Jackson Pollock (*Convergence*, 1952, Albright-Knox Art Gallery, and *Number 3*, 1950, Saint Louis Art Museum), each covered by the visualization of the mean attention distribution of the forty participants. In both images of figure 1 only one area (colored in red) shows remarkable clustering. Figure 2 shows two examples of paintings that received rather low activity ratings on average: a painting by Robert Motherwell (*Western Air*, 1946–47, Museum of Modern Art) and a painting by Vasily Kandinsky (*Several Circles*, 1926, Guggenheim New York).



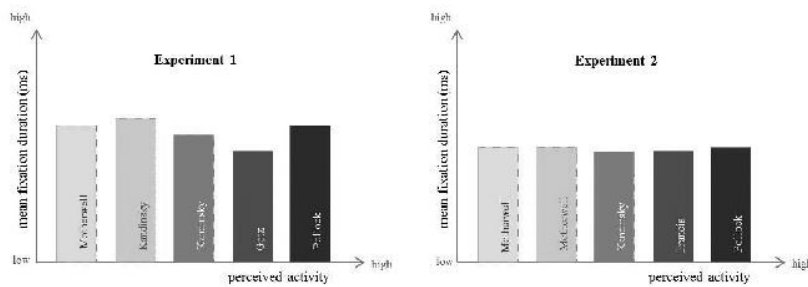
1 Collective Distribution of Attention—forty participants viewing Jackson Pollock's *Convergence*, painted in 1952 (top), and Jackson Pollock's *Number 3*, painted in 1950, for 2 minutes, 2014.

There is a considerable difference between the paintings in figure 1 and figure 2 concerning the number and intensity of red areas, although the visualization threshold is the same for all four artworks. Formal complexity, like individual attention clustering, also has a significant effect on collective attention distribution in both experiments ( $\beta = -20.46$ ,  $p > 0.002$  and  $\beta = -25.89$ ,  $p > 0.001$ ). The more formally complex a painting is, the more heterogeneous is the visual exploration among the participants. Perceived activity rates do not seem to alter the directedness of individual visual exploration, but have an effect on how homogeneous the exploration strategies are with regard to the localization of attention.





2 Collective Distribution of Attention—forty participants viewing Robert Motherwell's *Western Air*, painted in 1946–47 (right) and Vasily Kandinsky's *Several Circles*, painted in 1926 (left), for 2 minutes, 2014.



3 Mean fixation duration in ms for ten abstract artworks.

To examine whether the modality of attention distribution and the speed of the gaze varies with perceived activity, we compared the average fixation durations. Long fixations indicate a rather calm and steady exploration. Short fixations denote a fast-paced, restless, and jumping gaze. Figure 3 shows the average fixation duration for each artwork for both experiments.

The artworks are ordered from left to right according to their average activity ratings, beginning with the lowest score (from left to right, experiment 1: Robert Motherwell, *Western Air*; Vasily Kandinsky, *Accompanied Contrast*; Vasily Kandinsky, *Several Circles*; Karl Otto Götz, *Bild vom 5.2.1953*; Jackson Pollock, *Convergence*. From left to right,

experiment 2: Robert Motherwell, *The Garden Window*; Robert Motherwell, *The Little Spanish Prison*; Vasily Kandinsky, *Pink Compensation*; Sam Francis, *Untitled*; Jackson Pollock, *Number 3*). The vertical position of the upper bound of the bars shows the average fixation time for the paintings in milliseconds (ms). As we can infer from the diagrams, there is no systematic decrease or increase of fixation duration with increasing perceived activity. This, as expected, holds true when statistically tested: according to a mixed linear model with individual slopes for perceived activity on a subjective level, perceived activity does not significantly predict fixation durations.

### Discussion and Limitations

In contrast to Küster's assumption, we find that it is quite possible to perceive action paintings without action viewing. Moreover, there is no evidence for a qualitative equation of gaze speed or gaze dynamic and perceived activity in images at all. The gaze speed is not altered by the perceived activity of an artwork, whether it is a dynamic work by Jackson Pollock or a calm and easy painting by Vasily Kandinsky.

Boehm expected that the non-differential treatment of the surface in Jackson Pollock's artworks would lead to a rambling gaze. In opposition to his theoretical argumentation, our results show that perceived activity has no significant influence on the directedness of the individual exploration. This finding contradicts the empirical results reported by Massaro and colleagues, who showed that dynamic properties lead to less directed attention in representational artworks.<sup>18</sup> As we showed in our paper "Abstract Art as Universal Language?"<sup>19</sup> there are major differences in the perceptual strategies used to look at abstract and representational art. Thus, the divergent results might be related to the fact that the perception of abstract and representational art depends on different parameters.

Perceived activity, however, has an impact on the homogeneity of the beholder's attention distributions. The individual exploration patterns differ more during the perception of active and dynamic paintings compared to static and calm ones. Dynamics and activity can also be understood as formal image characteristics. Often they go hand in hand with a rather complex formation of elements. This holds especially true for our sample of artworks. Thus, the effect we found might not or not solely be related to the dynamic quality of the artwork, but rather to the fact that perceivable activity in

18 Massaro et al., "When Art Moves the Eye."

19 Brinkmann et al., "Abstract Art as a Universal Language?"

abstract artworks is a form of complexity.<sup>20</sup> We found complexity to be the most dominant predictor for both the directedness of attention distribution and the homogeneity of attention patterns. This is due to the fact that in formally complex paintings the number of potential attractors is higher. In order to grasp the depicted information, the beholder needs to spread his attention widely. Furthermore, the increasing number of potential points of interest leads to a decreasing probability of all participants choosing attractors that resemble each other.

We can conclude from our results that with regard to perceived image characteristics there are no aesthetic echoes in the beholder's eye. Despite participants' being consistent in their cognitive reasoning about affective image qualities, we do not find these qualities to influence the motion of the eye. The modalities of the physical process of seeing are thus not altered by cognitive responses to perceived image qualities. It is, rather, the formal structure of the artwork that moves the eye. Formal complexity and dominant points of interest structure how and where the gaze dwells.

Finally, we note that even if the results are robust in both experiments, using different participants and different artworks, the total number of ten observed artworks is still fairly low. To give a definitive answer to the question of the existence of a qualitative mirroring of activity in gaze and artwork, further research with higher sample and image numbers is needed. Nevertheless, our results provide the first empirical evidence for the divergence of theory and practice in the perception of abstract art.

### Acknowledgments

We would like to take the opportunity to thank Georgina Bábinszki, Solomiya Husak, Chiara Pompermaier, and Luisa Senkowsky for their assistance in data acquisition and data beneficiation. Furthermore, we thank the Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF) (Project CS 11-023, Helmut Leder and Raphael Rosenberg: *Time Makes the Difference! Uncovering the Nature of Aesthetic Experience*) and the Austrian Academy of Sciences (Österreichische Akademie der Wissenschaften, ÖAW) (Projekt: DOC-team Bildpraktiken: *Transdisziplinäre Studien zu Materialität und Habitualität visueller Konventionen*) for financial support.

20 See R. P. Taylor et al., "Perceptual and Physiological Responses to the Visual Complexity of Pollock's Dripped Fractal Patterns," *Nonlinear Dynamics, Psychology, and Life Sciences* 9 (2005): 89-114.

## Bibliography

- Berlyne, D. E. *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts, 1971.
- Berlyne, D. E., J. C. Ogilvie, and L. C. C. “Parham, the Dimensionality of Visual Complexity, Interestingness, and Pleasingness.” *Canadian Journal of Psychology* 22 (1968): 376–87.
- Betz, J., M. Engelbrecht, C. Klein, and R. Rosenberg. “Dem Auge auf der Spur: Eine historische und empirische Studie zur Blickbewegung beim Betrachten von Gemälden.” *IMAGE* 11 (2010).
- Boehm, Gottfried. “Die Form des Formlosen—Abstrakter Expressionismus und Informel.” In *Action Painting: Jackson Pollock*, edited by Fondation Beyeler, 38–46. Ostfildern: Hatje Cantz Verlag, 2008.
- Brinkmann, Hanna, Laura Commare, H. Leder, and R. Rosenberg. “Abstract Art as a Universal Language?” *Leonardo* 42 (2014): doi:10.1162/LEON—a—00767.
- Erdmann, B., and R. Dodge. *Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage*. Halle: Niemeyer, 1898.
- Holmqvist, K., M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Winter. *Eye-Tracking: A Comprehensive Guide to Methods and Measures*. New York: Oxford University Press, 2011.
- Küster, Ulf. “Action Painting—Mythos und Realität.” In *Action Painting: Jackson Pollock*, edited by Fondation Beyeler, 12–19. Ostfildern: Hatje Cantz Verlag, 2008.
- Massaro, D., F. Savazzi, C. Di Dio, D. Freedberg, V. Gallese, G. Gilli, and A. Marchetti. “When Art Moves the Eye: A Behavioural and Eye-Tracking Study.” *PLOS ONE* 7 (2012): e37285. doi:10.1371/journal.pone.0037285.
- Nadal, M., E. Munar, G. Marty, and C. J. Cela-Conde. “Visual Complexity and Beauty Appreciation: Explaining the Divergence of Results.” *Empirical Studies of the Arts* 28 (2010): 173–91.
- Osgood, C. E., G. Suci, and P. Tannenbaum. *The Measurement of Meaning*. Urbana: University of Illinois Press, 1957.
- Rosenberg, H. “The American Action Painters.” *ARTnews* 51, no. 8 (December 1952), 22.
- Rosenberg, R. “Dem Auge auf der Spur. Blickbewegungen beim Betrachten von Gemälden—historisch und empirisch.” *Jahrbuch der Heidelberger Akademie der Wissenschaften für 2010* (2011): 76–89.
- Rosenberg, R., J. Betz, and C. Klein. “Augensprünge.” *Bildwelten des Wissens. Kunsthistorisches Jahrbuch für Bildkritik* 6 (2008): 127–29.
- Rosenberg, R., and C. Klein. “The Moving Eye of the Beholder: Eye-Tracking and the Perception of Paintings.” In *Art, Aesthetics and the Brain*, edited by J. P. Huston, M. Nadal, F. Mora, L. Agnati, and C. J. Cela-Conde, 79–108. New York: Oxford University Press, 2015.
- Rosenthal, M. *Abstraction in the Twentieth Century: Total Risk, Freedom, Discipline*. New York: Guggenheim Museum Publications, 1996.
- Taylor, R. P., B. Spehar, J. A. Wise, C. W. G. Clifford, B. R. Newell, and T. P. Martin. “Perceptual and Physiological Responses to the Visual Complexity of Pollock’s Dripped Fractal Patterns.” *Nonlinear Dynamics, Psychology, and Life Sciences* 9 (2005): 89–114.
- Wagner, C. “Der Mythos vom ‘unschuldigen Auge’. Blickbewegungen, Wahrnehmungsprozesse in der Bildenden Kunst.” In *Aisthesis: Wahrnehmungsprozesse und Visualisierungsformen in Kunst und Technik*, edited by M. Greenlee, C. Wagner, and C. Wolff, 31–50. Regensburg: Schnell & Steiner, 2013.
- Wagner, C. “Mapping the Eye. Blickbewegungen und Verstehensprozesse in der Kunst” in *Bilder sehen: Perspektiven der Bildwissenschaft*, edited by M. Greenlee, R. Hammwöhner, B. Körber, C. Wagner, and C. Wolff, 69–86. Regensburg: Schnell & Steiner, 2013.



---

## *Diskussion*

Ziel der Studie war es, einerseits eine empirisch informierte Definition der Konzepte *action viewing* und *schweifender Blick* zu entwickeln und andererseits die von Küster und Boehm postulierten Thesen zur Wahrnehmung und Betrachtung von Jackson Pollocks *actions paintings* in einer auf diesen Definitionen aufbauenden empirischen Untersuchung zu prüfen. Der Schwerpunkt der Studie liegt dabei auf dem abstrakten Grundgedanken, der diesen Thesen zugrunde liegt: der Möglichkeit einer Entsprechung von wahrgenommenen Bildqualitäten, wie in diesem Fall Dynamik, und konkretem Betrachtungsvorgang.

Dazu wurde zunächst auf der Basis von bestehenden Erkenntnissen zur Funktionsweise und Modalität des menschlichen Sehapparates ermittelt, welche Merkmale für die Konzepte *action viewing* und *schweifender Blick* in der tatsächlichen Praxis des Betrachtungsvorgangs als charakteristisch verstanden werden können. Wie der Argumentation des Textes zu entnehmen ist, wurde als Arbeitsdefinition hier angenommen, dass die beiden Konzepte durch eine breite Streuung von Aufmerksamkeit und einer geringen Verweildauer des Blicks an spezifischen Interessenpunkten beschrieben werden können; dass es sich also um einen schnellen und sprunghaften Betrachtungsvorgang mit breiter Flächenstreuung handelt.

Um zu ermitteln ob ein solches Betrachtungsverhalten tatsächlich in besonderem Maße durch die Wahrnehmung von Dynamik angeregt wird, wurde für die Untersuchung zum einen – unter Zuhilfenahme eines semantischen Differentials (vgl. Tabelle 4) – die Erfahrung von bildimmanenter Dynamik und Aktivität erfasst und zum anderen eine Bildauswahl zugrunde gelegt, die hinsichtlich dieser Merkmale Unterschiede aufweist, sodass ein Vergleich des Betrachtungsvorgangs bei mehr oder weniger dynamisch-wahrgenommen Kunstwerken möglich war (siehe Abbildung 17).



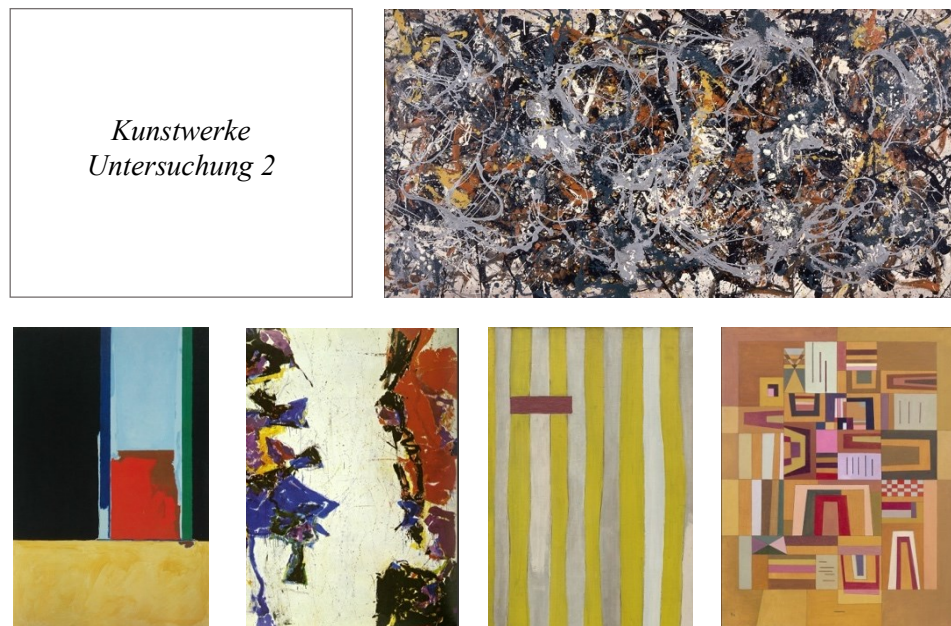
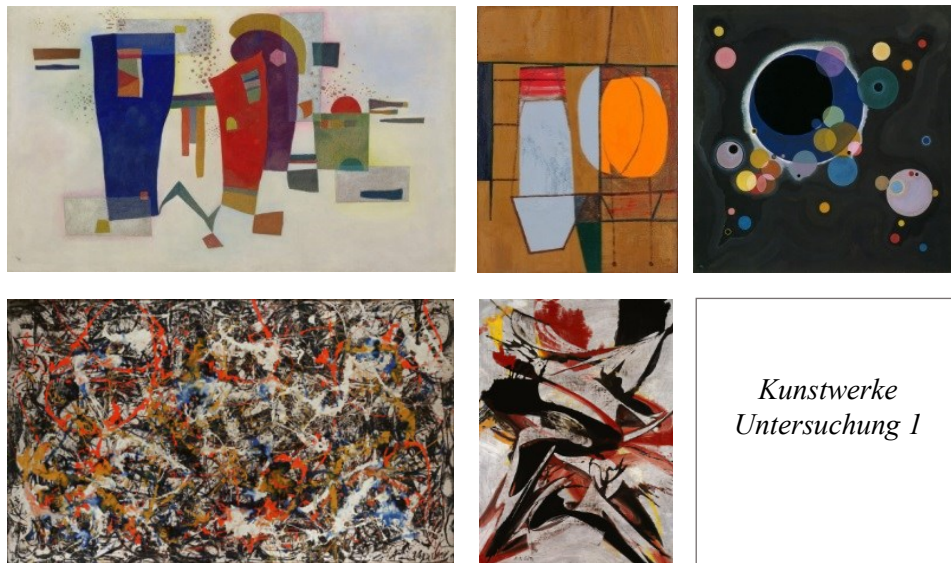


Abbildung 17 | Bildmaterial der Studie „*Aesthetic Echoes in the Beholder’s Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art*” (**Kunstwerke Untersuchung 1**, von oben und links: Kandinsky, Wassily (1935). *Kontrast mit Begleitung*, Foto: Artstor, © Guggenheim Museum New York; Motherwell, Robert (1946-47). *Western Air* (182.9 x 137.2 cm), Foto: Artstor, © The Museum of Modern Art New York; Kandinsky, Wassily (1926). *Several Circles* (140.4 x 140.7 cm), Foto: Artstor, © Guggenheim New York; Pollock, Jackson (1952). *Convergence* (241.9 x 399.1 cm), Foto: Artstor, © Albright-Knox Art Gallery; Götz, Karl Otto (1953). *Bild vom 5.2.1953* (125 x 90 cm), Foto: Artstor, © Staatliche Museen zu Berlin. **Kunstwerke Untersuchung 2**, von oben und links: Pollock, Jackson (1950). *Number 3* (122 x 145 cm), Foto: Artstor, © Saint-Louis Art Museum; Motherwell, Robert (1969). *The Garden Window* (152.4 x 124.5), Foto: Artstor, © Modern Art Museum of Fort Worth; Francis, Sam (1958). *Untitled* (275.6 x 192.4 cm), Foto: Artstor, © Norton Simon Museum of Art; Motherwell, Robert (1941). *The Little Spanish Prison* (69.2 × 43.5 cm), Foto: Artstor, © Museum of Modern Art New York; Kandinsky, Wassily (1933). *Ausgleichrosa* (92 x 73 cm), Foto: Artstor, © Musée national d'Art moderne, Centre Georges Pompidou.).

---

Gegenüber den beiden vorangegangenen Studien zeichnet sich diese Studie durch eine spezifische Besonderheit aus: es handelt sich um eine Replikationsstudie. Die zehn Werke, welche für die Untersuchung zugrunde gelegt wurden, sind also das Material für zwei identische Studien mit jeweils anderen Befragten gewesen. Hierbei wurden jedoch nur je fünf der ausgewählten Bildbeispiele pro Untersuchung verwendet.

Der Vorteil einer Replikationsstudie besteht darin, dass die Wahrscheinlichkeit eines fehlerhaften Befundes minimiert wird. Für sich stehende empirische Untersuchungen bergen immer die Gefahr eines irreführenden Zufallsbefundes, der durch die spezifischen Merkmale der Befragten oder des zugrunde gelegten Untersuchungsmaterials zustande kommt. Auf andere Befragte ist ein solcher Befund nur schwerlich übertragbar. Die Replikation der Ergebnisse mit anderen Befragten und anderem Untersuchungsmaterial reduziert die Wahrscheinlichkeit eines solchen zufälligen und für eine Gruppe von Befragten spezifischen Befundes erheblich. Der Rückgriff auf Replikationsstudien kann somit auch als Versuch oder Bestreben gewertet werden, das Argument, welches durch die Studie zu einem bestehenden Diskurs geleistet wird, stärker zu unterfüttern.

Die Auswertung der erhobenen Daten ist dem vorangegangenen Studienbeispiel sehr ähnlich, da hier weitestgehend auf dieselben Maße zurückgegriffen wird. Wie breit die Aufmerksamkeit bei der Betrachtung der Werke jeweils streute, wurde anhand der individuellen Gerichtetheit des Blicks überprüft: je weniger der Blick der Betrachtenden an bestimmten Interessenpunkten verweilt, desto *schweifender* das Auge. Darüber hinaus wurde auch die Übereinstimmung zwischen den Betrachtenden hinsichtlich der interessierenden Bildregionen erfasst. Neben diesen beiden die Streuung der Aufmerksamkeit beschreibenden Maßen wurde außerdem auch die Verweildauer der einzelnen Fixationspunkte der Betrachtung betrachtet: je kürzer die durchschnittliche Verweildauer des Auges, desto *sprunghafter* und *aktiver* der Blick.

---

Wie schon im Fall der zuletzt besprochenen Studie haben die Befragten hier ebenfalls in beiden Studien die wesentlichen Qualitäten der untersuchten Werke in großer Übereinstimmung wahrgenommen und bewertet. Ein Einfluss der jeweils wahrgenommenen großen oder kleinen Aktivität bzw. Dynamik auf den Blick hat sich hingegen nicht nachweisen lassen. Weder sinken die durchschnittlichen Verweildauern an Fixationspunkten, wenn die wahrgenommene Aktivität oder Dynamik steigt, noch streut der individuelle Blick breiter über die Bildfläche.<sup>154</sup> Beide Untersuchungen kommen trotz der abweichenden Befragtengruppe und dem unterschiedlichen zugrunde gelegten Bildmaterial zum selben Ergebnis. Ein spezifischer Einfluss auf die Gerichtetheit der individuellen Betrachtungsvorgänge konnte im Rahmen der Studie nur für die Anzahl der dargestellten Elemente nachgewiesen werden: je mehr Elemente die jeweiligen Darstellungen enthalten, desto breiter war die Streuung der individuellen Blickmuster. Wie zuvor bereits erwähnt, ist dies intuitiv kaum überraschend, denn je mehr auf einem Werk zu sehen ist, desto breitflächiger wird der Blick streuen, um sich einen Überblick zu den unterschiedlichen Elementen und ihren Eigenheiten zu verschaffen.

Obwohl die wahrgenommene Aktivität bzw. Dynamik der Werke keinen nachweisbaren Einfluss auf die individuelle Streubreite der Aufmerksamkeit während der Betrachtung gezeigt hat, haben sich Effekte hinsichtlich der *Übereinstimmung* in den interessierenden Bildregionen gezeigt: je aktiver und dynamischer die Kunstwerke von den Betrachtenden bewertet wurden, desto heterogener oder individueller waren die jeweiligen Betrachtungsvorgänge. Es haben also bei besonders dynamischen Werken deutlich weniger Bildregionen ein intersubjektiv-geteiltes Interesse hervorgerufen. Ob dieses Ergebnis aber tatsächlich allein auf die wahrgenommenen Bildqualitäten zurückzuführen ist, ist fraglich. Es ist durchaus möglich, dass sich hier ein Effekt ausdrückt, der eigentlich einer formalen Charakteristik zuzuschreiben ist. Das Erzielen eines dynamischen Ein- bzw. Ausdrucks ist in abstrakten Arbeiten naturgemäß durch eine bestimmte Linienführung und

---

<sup>154</sup> Diese Ergebnisse sind aus einigen (u.a. methodischen) Gründen nicht unproblematisch. Eine Auseinandersetzung mit diesen Problemen findet sich in Kapitel 3.3.

---

Formenzusammenstellung bedingt – etwa durch eine langschweifige und offene Gestaltung. Eine dynamische Bildgestaltung kann somit genuin eine breitere Streuung von Interessenfoki implizieren; bei tendenziell eher kompakt-konzentrisch orientierten Formen- und Gestaltungsmaximen ist das Gegenteil der Fall. Betrachten unterschiedliche Personen also eine dynamische Komposition ist der Schwerpunkt der Formen nicht eindeutig definiert, was die Wahrscheinlichkeit reduziert, dass alle Betrachtenden übereinstimmende Aspekte der jeweiligen Formen schwerpunktmäßig in den Blick nehmen. Bei ruhigeren Arbeiten mit kompakt-konzentrischen Elementen sind hingegen eindeutiger Formenschwerpunkte und spezifische Begrenzungslinien auszumachen, die als Orientierungspunkte hervorstechen und somit das Potenzial besitzen als intersubjektiv-geteilte Aufmerksamkeitsattraktoren zu fungieren. Die beobachtete stärkere Übereinstimmung der Betrachtungsvorgänge bei weniger dynamischen Werken mag daher durchaus eher auf die formalen „Angebote“ der Werke zurückzuführen sein, denn auf den erlebten Ausdruck. Dies gilt umso mehr, als eine stärkere Gerichtetheit des Betrachtungsvorgangs in Abhängigkeit von erlebtem dynamischen Ausdruck nicht nachweisbar war.

#### *Fazit und Beitrag der empirischen Methoden*

Gegeben der gewählten Arbeitsdefinition lässt sich für den gesetzten Diskurs aus den Ergebnissen der Studie schlussfolgern das Ulf Küsters Annahmen zu einer qualitativen Entsprechung von wahrgenommener Bildqualität und Blick nicht zutreffen. Die Erfahrung von bildimmanenter Dynamik führt nicht zu einem aktiverem oder dynamischerem Betrachtungsverhalten; das heißt, dass bei der Betrachtung von dynamischen Kunstwerken weder die einzelnen Verweilperioden des Blicks, die Fixationspunkte, kürzer werden, noch sind diese breiter gestreut, was einen sprunghaften und unruhigeren Betrachtungsvorgang andeuten würde. Damit ist deutlich, dass in der Praxis der Kunstbetrachtung ein Konzept wie Küsters *action viewing* nicht nachweisbar ist – weder bei der Betrachtung von Jackson Pollocks Arbeiten, noch bei anderen ebenfalls als dynamisch begreifbaren Werken. Da in der

---

Studie sowohl Laien als auch Expert\_innen untersucht wurden, gilt dieses Ergebnis für beide Gruppen.

In Falle von Gottfried Boehms Konzept des *schweifenden Auges* liegt der Fall ein wenig anders: interpretiert man das *schweifendes Auge* als analog zu Küsters *action viewing*, so ist das Konzept für die tatsächliche Praxis der Kunstbetrachtung ebenso wenig tragfähig; interpretiert man das Konzept hingegen *nicht* als ein dynamisches Betrachtungsverhalten, im Sinne von Schnelligkeit und Sprunghaftigkeit, sondern eher im Sinne eines konstant wandernden Blicks, findet sich eine Entsprechung dieser These in den Ergebnissen der Untersuchung. Obwohl wahrgenommene Dynamik keinen Einfluss auf den Blick nimmt, beeinflusst der formale Aufbau der Werke die Betrachtung sehr wohl: je mehr Elemente in einem Werk dargestellt sind, desto breiter streut die Aufmerksamkeit der Betrachtenden. Die fehlenden „kompositionellen Ordnungen“, wie Boehm es beschreibt, führen zu einem schweifenden Betrachtungsvorgang ohne explizite Schwerpunkte. Das ist allerdings explizit *nicht* auf die wahrgenommene Dynamik zurückzuführen, sondern durch die formalen „Angebote“ zu begründen, die sich den Betrachtenden bei der Auseinandersetzung mit dem Werk bieten. Diese Schlussfolgerung wird von einem weiteren Ergebnis der Studie gestützt: trotz des nicht vorhandenen Einflusses wahrgenommener Dynamik auf die individuelle Streubreite der Betrachtung zeigt sich bei Kunstwerken die als dynamisch begriffen werden, weniger Übereinstimmung in den individuellen Betrachtungsmustern. Je dynamischer also die Werke, desto unterschiedlicher werden sie betrachtet. Da der einzelne Betrachtungsvorgang nicht in seiner Modalität von der hier wahrgenommenen Bildqualität beeinflusst wird, liegt die Vermutung nahe, dass der Einfluss auf kollektiv-geteilte Muster ebenfalls nicht durch diesen Faktor bedingt ist. Vielmehr scheint es wahrscheinlich, dass die sich zeigende Korrelation nicht aufgrund eines direkten Kausalzusammenhangs der untersuchten Merkmale (wahrgenommene Dynamik und geteilte Aufmerksamkeitsstrukturen) zustande kommt, sondern aufgrund eines dritten, nicht direkt beobachteten Merkmals – nämlich, wie oben auch dargestellt, der formalen Bedingtheit wahrgenommener Dynamik.

---

Ein *schweifendes Auge* wäre demnach nicht das Resultat von ideellen Ausdrucksqualitäten, sondern ganz im Gegenteil das Resultat *formaler* Bildqualitäten. Im Falle von Jackson Pollocks Arbeiten, welche den Referenzrahmen des zugrundeliegenden Diskurses bilden, wäre es also das hier zur Anwendung kommende Prinzip des kleinteiligen *all-over paintings*, welches Einfluss nimmt auf die Praxis des Betrachtungsvorgangs. Dass Pollocks Arbeiten dabei gleichzeitig durch eine starke Dynamik im Ausdruck geprägt sind, ist für die Modalitäten der Betrachtung nicht direkt relevant.

Was für die beiden vorangegangenen Studien gilt, kommt auch in diesem Fall zum Tragen: die berichteten Ergebnisse und Beobachtung wären ohne den Rückgriff auf eine empirische Herangehensweise nicht leistbar gewesen. Der Rückgriff auf einen auf quantifizierbare Strukturen bedachten Blick war notwendig, um intersubjektive Wahrnehmungsübereinstimmungen hinsichtlich der Bildqualitäten abzubilden, und erst durch den spezifischen Vergleich von stark- und wenig-dynamisch begriffenen Werken wurde ein direktes Adressieren der *wahrgenommenen Dynamik* als eigenständigen Einflussfaktor auf die Betrachtung von Kunst ermöglicht.

Wie auch im Falle der zuletzt diskutierten Studie, sind es hier insbesondere die Erfassung und Diskussion der tatsächlichen Blickbewegungen bei der Betrachtung der Werke, die eines empirischen Zugang zwingend bedurften. In diesem Falle ist die Notwendigkeit einer (empirischen) Untersuchung des konkreten Betrachtungsvorgangs noch deutlicher gegeben, als dies für das vorherige Studienbeispiel der Fall war: während für die letzte Studie die Frage durchaus berechtigt war, ob eine Untersuchung des Blicks für die gesetzte Fragestellung zwingend notwendig ist, wird der Blick und das Betrachtungsverhalten im vorliegenden Diskurs eindeutig und direkt adressiert. Die Tatsache, dass die tatsächlich vollzogenen Blickbewegungen unserer Wahrnehmung in der Regel nur sehr begrenzt und teilweise überhaupt nicht direkt zugänglich sind, ist hier also noch problematischer als im zuletzt diskutierten Beispiel. Der Rückgriff auf technikgestützte, empirische Messungen und Auswertungen der im konkreten vollzogenen Bewegungen des Auges ist die einzige

---

Möglichkeit, die gesetzten Begrenzungen der menschlichen Beobachtungsfähigkeit zu umgehen. Im vorliegenden Fall bedarf die grundsätzliche Adressierbarkeit von für den Fragenraum relevanten Aspekten also ebenfalls zwingend eines empirischen Zugangs; denn ein Kommentar zu den postulierten Thesen Küsters und Boehms wäre ohne eine empirische Studie wenig mehr als Spekulation.

---

## 2.4 Zwischenfazit

Ziel dieses Abschnittes war es, die Möglichkeiten empirischer Forschungsansätze an konkreten Beispielen nachzuvollziehen, um so einen plastischeren Ansatzpunkt für den Nachvollzug der Argumente der *positiven Kritik* zu schaffen. Die von der positiven Kritik hervorgehobenen inhaltlichen Chancen beziehen sich hauptsächlich auf die Möglichkeit, empirischer Methoden neue Fragen- und Themenbereiche für die Kunstwissenschaft zu erschließen. Das umfasst sowohl das Aufzeigen noch unbeantworteter (klassischer) Fragen wie auch das Zugänglichwerden völlig neuer Kontexte oder Perspektiven. Dabei ist insbesondere die in quantitativen Ansätzen gegebene Schwerpunktsetzung auf Strukturen und Gesetzmäßigkeiten von Interesse, da der Kunstwissenschaft bisher kein Instrumentarium mit vergleichbaren Möglichkeiten zur Verfügung stand. Diesen Möglichkeiten und Chancen lässt sich zunächst schwerlich etwas entgegensetzen; ob sich durch den Einsatz empirischer Methoden allerdings *tatsächlich* ein für die Kunstwissenschaften relevanter und nutzenbringender Erkenntniszugewinn realisieren lässt, ist diesen allgemeinen Ausführungen nicht zu entnehmen. Erst der Nachvollzug an konkreten Beispielen erlaubt eine Aussage zum spezifischen Nutzen empirischer Ansätze und dem durch die Anwendung von empirischen Zugängen realisierten Erkenntnisgewinn.

Die drei in diesem Kapitel diskutierten Studien zeigen deutlich, dass es durchaus sehr unterschiedliche, für bestehende Diskurse relevante Fragen- und Problemstellungen gibt, deren Beantwortung zwingend eines empirischen Zugangs bedarf. In den vorgestellten Beispielen waren die jeweiligen empirischen Herangehensweisen zum einen in der Lage sowohl theoriegebende als auch theorietestende Ansätze zu unterstützen. Sie waren zentral in der näheren Beleuchtung von Prämissen unterschiedlicher Argumentationen und haben wesentlich dazu beigetragen theoretische Annahmen über Regelmäßigkeiten auf ihren Praxisgehalt zu prüfen. Das empirische Methodeninstrumentarium hat dabei die grundsätzliche Adressierbarkeit der jeweiligen für die gesetzten Fragenräume relevanten Aspekte erst ermöglicht. Der Einsatz der gewählten Methoden hat zum einen der subjektiven



---

Erfahrung *nicht* oder nur schwer zugängliche Strukturen und Regelmäßigkeiten aufgezeigt und zum anderen durch den Einsatz statistischer Auswertungsverfahren den Zusammenhang verschiedener, teils überlappender Einflussfaktoren offengelegt. Keine der so realisierten Chancen der empirischen Zugänge hätte mit anderen als empirischen Mitteln ebenfalls erreicht werden können.

---

*Kapitel* **3**

**Methodenreflexion**

---

Während das vorangegangene Kapitel schwerpunktmäßig die *inhaltlichen Möglichkeiten* behandelt hat, welche von der positiven Kritik an empirischen Ansätzen besonders hervorgehoben werden, widmet sich dieses Kapitel den *methodischen Möglichkeiten* empirischer Arbeiten. Insbesondere diese werden, wie eingangs diskutiert, von der negativen Kritik als problematisch begriffen. Vor allem besteht Zweifel an der Übersetzbarkeit geisteswissenschaftlicher Konzepte in messbare Größen. Im Folgenden soll daher der Schwerpunkt auf einer Reflektion und Diskussion der in den diskutierten Studien verwendeten Methoden liegen, um dadurch die Argumente der negativen Kritik und deren Prämissen besser zu beleuchten.

Da die einzelnen Kritikpunkte zugleich mehrere Merkmale von empirischen Methoden betreffen, werde ich die Diskussion nicht entlang der Kritikpunkte, sondern entlang von grundsätzlichen Überlegungen und Problemen empirischer Methoden strukturieren. Im Folgenden werde ich zunächst kurz das in den vorgestellten Studien verwendete Methodenrepertoire rekapitulieren und die jeweiligen Besonderheiten der einzelnen Messinstrumente herausstreichen. Dann werde ich besprechen, welche Probleme im Allgemeinen bei Datenerhebungen auftreten können, und diskutieren, inwiefern diese in den vorgestellten Studien präsent sind. Anschließend werde ich auf Methoden zur Analyse- und Interpretation eingehen, allgemeine Probleme und Fallstricke benennen und hier ebenfalls aufzeigen, in welcher Form diese in den diskutierten Studien vorliegen.

Wie im Falle der inhaltlichen Diskussion wird es auch in der Methodendiskussion nicht möglich sein, trennscharf zwischen inhaltlichen und methodischen Aspekten zu unterscheiden. „Grenzüberschreitungen“ wird es auch hier wiederholt geben, da bestimmte methodische Vorgehensweisen nur mit Blick auf die umgesetzten Konzepte sinnvoll fassbar sind.

---

### 3.1 Methodenübersicht

Wie eingangs angedeutet bilden die gewählten Studien immerhin eine kleine Methodenbreite ab. Alle drei Studien arbeiten jedoch hauptsächlich mit der Erfassung von Antworten durch *standardisierte Fragebögen*. Damit ist gemeint, dass Teilnehmende zur Beantwortung der gestellten Fragen aus einer Reihe von Antwortoptionen die zutreffendste Antwort auswählen. Die Antwortoptionen stellen dabei in der Regel ein Kontinuum dar und gestatten die Positionierung der eigenen Antwort zwischen zwei Polen. Die jeweiligen Extrema sind abhängig von der Fragestellung. Zum Beispiel würden die Pole bei einer Frage nach der Häufigkeit einer bestimmten Handlung (z. B. Wie häufig besuchen Sie in ihrer Freizeit Ausstellungen oder Museen?) am einen Ende durch die Auswahloption „nie“ und am anderen Ende durch einen plausiblen Höchstwert ausgedrückt werden, etwa „mehrmals die Woche“. Die weiteren Antwortoptionen stellen graduelle Annäherungen zwischen den Extrema dar, also etwa „einmal im Jahr“, „einmal im halben Jahr“, „alle drei Monate“, usw. Im Fall von Wahrnehmungsinhalten sind die Pole und auch die Antwortoptionen dazwischen anders auszuzeichnen, beispielsweise durch den Grad der Zustimmung zu einer bestimmten Aussage („Ich empfinde das gezeigte Kunstwerk als übersichtlich.“). Hier würde sich das Kontinuum von „stimme überhaupt nicht zu“ bis „stimme vollkommen zu“ erstrecken. Wie viele Antwortoptionen (inklusive der Pole) den Teilnehmenden jeweils zur Verfügung stehen, ist nicht genormt; in der Regel werden zwischen fünf und sieben Optionen angeboten. Das liegt daran, dass insbesondere bei Fragen, die einen Zustimmungsgrad abfragen, immer feiner werdende Zwischenschritte ab einem gewissen Punkt nicht mehr sinnvoll von den Teilnehmenden interpretiert werden können. Studien haben gezeigt, dass sich 5er oder 7er Skalen am besten eignen.<sup>155</sup> In den diskutierten Studien wurden jeweils 7er Skalen verwendet, da diese eine genauere Differenzierung erlauben als 5er Skalen (für ein Beispiel siehe Abbildung 11).

---

<sup>155</sup> Matell/Jacoby 1971.

---

Die zweite Studie greift neben standardisierten Fragebatterien auch auf *offen gestellte Fragen* zurück, die es den Teilnehmenden erlauben, ihre Wahrnehmung in eigenen Worten zu beschreiben. Der Vorteil einer solchen Datenerhebungsstrategie liegt darin, dass Wahrnehmungsinhalte hier nicht vorweggenommen werden, indem die Antwortmöglichkeiten auf vordefinierte Kategorien eingeschränkt sind. Die Möglichkeit, frei zu antworten, erlaubt es vielmehr, individuelle und gruppenspezifische Wahrnehmungsinhalte zu erfassen, die praktische Relevanz besitzen. Die zweite Studie hat sich dieser Strategie in Form einer Vorstudie bedient. Hier wurde in einem ersten Schritt ermittelt, welche Wahrnehmungsinhalte durch die Auseinandersetzung mit bestimmten Bildern evoziert werden. Diejenigen Wahrnehmungsinhalte, die als intersubjektiv bedeutsam verstanden werden können, weil sie von allen oder zumindest den meisten Teilnehmenden berichtet wurden, wurden anschließend für die eigentliche Studie in standardisierte Fragen überführt. Die dritte Studie greift auf dieselben, in der Vorstudie der zweiten Studie entwickelten Kategorien zurück. Vorgehen und Strategie der zweiten und dritten Studie unterscheiden sich damit fundamental von denen der ersten Studie. Während die abgefragten Aspekte in der ersten Studie im Wesentlichen aus der Literatur abgeleitet wurden, handelt es sich bei der zweiten und dritten Studie um nachweislich in der Praxis der Erfahrung evozierte Wahrnehmungsinhalte. Die erste Studie verfolgt also einen theoriebasierten *top-down* Ansatz, die Studien zwei und drei hingegen zeichnen sich durch eine praxis-orientierte *bottom-up* Perspektive aus.

Neben der Erfassung und Beschreibung von Wahrnehmungsinhalten ist für alle drei Studien auch die Frage zentral, inwiefern diese Wahrnehmungsinhalte mit personenbezogenen Merkmalen variieren. Insbesondere die Frage nach dem Einfluss von Expertise ist in den gewählten Studien schwerpunktmäßig adressiert worden. Die Erfassung von Expertise und die Klassifikation der Teilnehmenden gemäß ihres Expert\_innen-Status erfolgte jedoch mit unterschiedlichen Mitteln. Während die erste Studie Expertise schlicht als *dichotome Kategorie* führt, wobei zwischen dem Vorhandensein eines Studienabschlusses in Kunstgeschichte und dem Fehlen eines solchen die Trennlinie gezogen wird, verwendet die

---

zweite Studie eine Auswahl verschiedener Indikatoren. Das Studienfach spielt ebenfalls eine große Rolle, zusätzlich werden jedoch auch Wissen über Kunstwerke und Künstler sowie die freizeithliche Kunstaffinität betrachtet. Die dritte Studie wiederum arbeitet mit einem mehrdimensionalen *Index*, der Expertise in Form eines Punktesystems abbildet. Dabei werden Punkte vergeben – auf Wissensfragen, Häufigkeit von kunstbezogenen Handlungen, etc. In der Summe bilden diese den Expertise-Index – viele Punkte bedeuten eine hohe Expertise, wenige Punkte hingegen eine geringe Expertise. Expertise wird in diesem Fall also in Form eines Kontinuums konzeptualisiert und gemessen.

Jenseits der Fragebögen fanden in Studie zwei und drei auch *physiologische Messungen* des Blickverhaltens (*Eye-Tracking*) statt. Physiologische Maße erlauben es körperliche Reaktionen auf Gesehenes zu erfassen. Dabei spielt es keine Rolle, ob die Reaktion der Betrachtenden bewusst oder unbewusst erfolgt. Dies hat den Vorteil, dass auch unbewusste Vorgänge, welche bei der Wahrnehmung von und Auseinandersetzung mit Kunst stattfinden, erfasst werden können. Diese können zum besseren Verständnis der bewusst erlebten Aspekte herangezogen werden und tragen somit zur Erklärung von subjektiven Erfahrungen bei. Hierbei gilt jedoch die Annahme, dass körperliche Reaktionen tatsächlich in direkter Relation zu kognitiven Prozessen stehen. Für den Fall von Blickbewegungsmessungen wird dazu die sogenannte *Eye-Mind-Hypothese* zugrunde gelegt.<sup>156</sup> Diese geht im Wesentlichen davon aus, dass Fixationen und die kognitive Verarbeitung von visueller Information eng verknüpft sind. Damit ist im Konkreten gemeint, dass diejenigen Bereiche des Sichtfeldes, welche vom Auge zu einem bestimmten Zeitpunkt scharf gestellt sind, gleichzeitig auch aktiv verarbeitet werden. Ein gutes Beispiel ist das Lesen: diejenigen Worte oder Silben, die jeweils für einen Moment fixiert werden, sind auch diejenigen, die im selben Moment mit der größten Aufmerksamkeit verarbeitet werden. Dabei ist jedoch das ständige Neuausrichten des Fixationspunktes selbst, also die sich fortziehende Bewegung des Auges, nicht notwendig ebenfalls bewusster Wahrnehmungsinhalt. Die unbewusste, fortlaufende

---

<sup>156</sup> Wright/Ward 2008.

---

Verschiebung des Punktes der größten Aufmerksamkeit erlaubt aber dennoch Rückschlüsse auf die am Leseprozess beteiligten Verarbeitungsschritte und deren zeitliche Dimensionen.<sup>157</sup>

Die Messung von Blickbewegungen bei der Betrachtung von Kunstwerken wurde in den Studien zwei und drei im *Laborkontext* vorgenommen. Die Versuchsteilnehmenden haben die ausgewählten Bildbeispiele auf einem Bildschirm in hochauflösender digitaler Reproduktion betrachtet. Das zur Aufzeichnung des Blicks verwendete Gerät ist ein *binocular remote Eye-Tracker* der Marke SensoMotoric Instruments (SMI), welcher in einer *Frequenz von 120 Herz (Hz)* Datenpunkte generiert. Diese technischen Informationen sind insofern relevant, als sie einerseits die Bedingungen für die Datenerhebung vorgeben und andererseits die Auswahl bestimmter Auswertungsverfahren bedingen. Wie der Name bereits suggeriert, erfordern *remote* Geräte keinen physischen Kontakt mit der Versuchsperson. Diese kann also ohne die Einschränkung einer Brille, Kinnstütze oder Ähnlichem, vor einem Bildschirm sitzend die ausgewählten Bildbeispiele betrachten (siehe Abbildung 18). *Remote* Geräte sind somit deutlich weniger invasiv und ermöglichen einen natürlicheren Betrachtungsvorgang als Geräte, die einen physischen Kontakt mit dem betrachtenden Subjekt erfordern (vgl. Abbildung 19).

---

<sup>157</sup> Ich verzichte an dieser Stelle auf eine genaue Beschreibung von Eye-Tracking als Methode bzw. Messtechnik, da die dritte Studie ein Unterkapitel enthält, das bereits detaillierte Auskünfte bietet. Für eine noch umfassendere Beschreibung siehe auch Holmqvist et al. 2011.



Abbildung 18 | Nachstellung Laborsituation im *Labor für empirische Bildwissenschaft (CReA)* am Institut für Kunstgeschichte der Universität Wien mit SMI SensoMotoric Instruments Eye-Tracker, Foto © Karl Pani und René Steyer.



Abbildung 19 | Beispiel Labornachstellung für ein Eye-Tracking-Experiment mit Kinnstütze, Foto © <http://translationalneuromodeling.org/forschungsambulanz/eye-tracking/> (20.08.18).



---

Da es sich bei dem in Studie zwei und drei verwendeten Eye-Tracker um ein binokulares Gerät handelt, können bei der Aufzeichnung der Blickbewegungen beide Augen berücksichtigt werden. Ist es für Fragestellung oder Auswertungsmethode von Relevanz, können so auch Abweichungen zwischen den Blickpunkten der beiden Augen ermittelt werden. Für die hier diskutierten Studien wurden solcherart Abweichungen nicht gezielt untersucht. Die binokulare Messung war hier hauptsächlich deswegen von Vorteil, weil Unterschiede zwischen Personen in der jeweiligen *Augendominanz*<sup>158</sup> keine Verzögerung der Messung aufgrund von Umbauten oder Einstellungsänderungen erfordert haben. Die binokulare Messung hat demnach einen flüssigen Versuchsablauf garantiert. Der letzte technische Aspekt der Eye-Tracking-Messung, auf den ich an dieser Stelle verweisen möchte, bezieht sich auf die erwähnte Messfrequenz von 120 Hz. Bei einer Messfrequenz von 120 Hz entstehen 120 Datenpunkte pro Sekunde, also ein Datenpunkt alle 8 Millisekunden (ms). Damit rangiert das eingesetzte Gerät am Niedrig-Frequenz-Ende der technischen Möglichkeiten; *high-speed* Geräte messen mit einer Frequenz von etwa 2000 Hz. Der Umstand, dass es sich gewissermaßen um eine „langsame“ Messung handelt, hat für die Aufzeichnung der Daten keine Konsequenz. Auswertung und Interpretation der gemessenen Blickpunkte stehen jedoch in Abhängigkeit der jeweiligen Aufzeichnungsfrequenz, weswegen diese im Folgenden noch Gegenstand der Diskussion sein wird.

Alle verwendeten Datenerhebungsmethoden, seien sie nun Fragebogen-basiert oder physiologischer Natur, sind in den vorgestellten Studien in ein sehr bestimmtes Erhebungssetting eingebunden gewesen. Zentral für alle drei Studien war, wie im vorangegangenen Kapitel bereits diskutiert, die randomisierte Darbietung der untersuchten

---

<sup>158</sup> Unter *Augendominanz* wird der Umstand gefasst, dass, obgleich im binokularen Sehen jeweils zwei retinale Bilder vom Gehirn verarbeitet werden, eine Präferenz für eines dieser beiden Bilder zu verzeichnen ist. Welches Auge dabei jeweils das „dominante“ darstellt, ist von Mensch zu Mensch unterschiedlich. Um Fehler in der räumlichen Verortung von Blickbewegungen zu vermeiden, muss das „dominante“ Auge in der Aufzeichnung und Auswertung von Blickbewegungsdaten berücksichtigt werden (vgl. auch Holmqvist et al. 2011; Reiss/Reiss 1997).

---

Bildbeispiele. Die Versuchsteilnehmenden haben die einzelnen Kunstwerke also jeweils in unterschiedlicher Reihenfolge gesehen und bedacht. Die *Randomisierung* stellt dabei sicher, dass keine systematischen Reihenfolgeeffekte die Ergebnisse verzerren. Ein weiterer Punkt, der zum spezifischen Erhebungssetting gezählt werden kann, ist die *Reihenfolge oder Kopplung* der gestellten Fragen. Hier wurde für die erste Studie ein anderes Vorgehen zugrunde gelegt als für die Studien zwei und drei. In der ersten Studie erfolgte die Befragung der Teilnehmenden in zwei gesonderten Blöcken. Im ersten Block wurde zunächst ausschließlich nach dem Komplexitätseindruck gefragt, der bei der Betrachtung der ausgewählten Werke spontan entsteht. Bewertungshinweise, wie etwa die Bitte sowohl auf inhaltliche wie auch auf formale Aspekte zu achten, wurden bewusst unterlassen. Erst in einem zweiten Block wurden Fragen zur Wahrnehmung von inhaltlichen sowie formalen Bildaspekten gestellt. Im Versuchsaufbau der Studien zwei und drei wurde eine solche Trennung nicht vorgenommen. Hier haben die Versuchsteilnehmenden immer jeweils eines der gewählten Bildbeispiele für zwei Minuten betrachtet und dann im direkten Anschluss Fragen zu ihren Wahrnehmungsinhalten beantwortet. Während eine *Entkopplung*, wie sie in der ersten Studie zur Anwendung kam, den Vorteil hat, dass die Antworten des ersten Blocks kaum durch Inhalte des zweiten beeinflusst werden, hat die in der zweiten und dritten Studie verwendete *Kopplung* den Effekt, dass Wahrnehmungsinhalte, die während der Auseinandersetzung mit dem Werk von Bedeutung waren, noch deutlich präsent sind.

Diese knappe Übersicht der angewandten Methoden abschließend ist noch zu bemerken, dass die dritte Studie sich gegenüber den Studien eins und zwei durch eine erwähnenswerte Besonderheit auszeichnet: anders als in den Studien eins und zwei wird hier die *Reproduzierbarkeit der Ergebnisse* thematisiert. Zu diesem Zweck wurde eine inhaltlich-identische Studie mit jeweils unterschiedlichen Personen und Bildern durchgeführt, um anschließend zu vergleichen, ob sich die Ergebnisse unterscheiden.

---

### 3.2 Datenerhebung

Alle methodischen Vorgehensweisen haben spezifische Vorteile. Diese bedingen in der Regel die Entscheidung auf eine bestimmte Methode zurückzugreifen. Der vorangegangene Abschnitt hat einige diesbezügliche Überlegungen benannt. Dennoch haben alle methodischen Vorgehensweisen auch Nachteile. Viele dieser Nachteile realisieren sich im Kontext der Datenerhebung, da empirische Methoden naturgemäß Daten mit jeweils ganz bestimmten Eigenschaften und Einschränkungen produzieren. Sich dieser Eigenschaften und Einschränkungen bewusst zu sein, ist für die Arbeit mit empirischen Daten wesentlich, da diese Daten die Grundlage aller später gezogenen Schlüsse und Ergebnisse darstellen. Im Verlauf dieses Kapitels trage ich daher Überlegungen zu allgemeinen Problemen der Datenerhebung sowie zu spezifischen Problemen der benannten methodischen Aspekte zusammen, um so einen differenzierteren Blick auf die Fundamente empirischer Prinzipien werfen zu können. Ziel ist es hierbei, die diesbezüglich geäußerten Bedenken der negativen Kritik an empirischen Arbeiten im Kontext der Kunstwissenschaft zu adressieren; und zwar in der Zusammenschau mit konkreten, sich tatsächlich genuin aufgrund der Methodenwahl ergebenden Einschränkungen.

Hierzu ist zunächst anzumerken, dass alle empirischen Methoden grundsätzlich drei Haupt-Gütekriterien Genüge tun sollten: Sie sollten valide, objektiv und reliabel sein.<sup>159</sup> Damit sind im Kontext der *Testtheorie* jeweils sehr bestimmte Konzepte beschrieben. Die *Validität* eines Testinstrumentes bestimmt den Grad der Genauigkeit, mit welchem das Testinstrument misst, was gemessen werden soll. Validität bezeichnet im Kontext empirischer Methoden also die Übereinstimmung von theoretischem Konzept und empirischem Maß. Dabei ist es gleichgültig, welcher Art ein Testinstrument ist; die Forderung nach Validität gilt für Fragebögen genauso wie für Geräte, die physiologische Messungen erlauben. Selbiges gilt auch für die Forderung nach Objektivität und Reliabilität. Der Forderung nach *Objektivität*

---

<sup>159</sup> Für eine ausführlichere Diskussion als die folgende siehe Diekmann 2007.

---

ist genüge getan, wenn ein Testinstrument Ergebnisse generiert, die unabhängig von den Bedingungen ihrer Messung sind. Darunter fallen beispielsweise die räumlichen Bedingungen der Messung oder auch die anwesenden Test- und Untersuchungsleiter\_innen. Weder ein Ortswechsel noch ein Wechsel der Test- bzw. Untersuchungsleiter\_innen sollte zu anderen Ergebnissen führen. Schließlich sollte sich ein Testinstrument noch durch eine hohe *Reliabilität* auszeichnen. Reliabel sind Testinstrumente, wenn sie bei wiederholter Messung unter gleichen Rahmenbedingungen das jeweils gleich Ergebnis erzielen. Reliable Ergebnisse sind demnach nicht zufällig und deswegen auch reproduzierbar.

Es ist offensichtlich, warum Validität, Objektivität und Reliabilität die wesentlichen Stützpfeiler und Gütekriterien von empirischen Methoden darstellen. Trotzdem zeichnet sich kaum ein Testinstrument tatsächlich dadurch aus, *absolut* valide, objektiv und reliabel zu sein. Wenn die negative Kritik an empirischen Arbeiten also betont, die reine Tatsache, dass empirische Messungen stattgefunden haben, führe nicht notwendigerweise zu validen und objektiven Ergebnissen, so ist dies ein berechtigter Einwand der Beachtung verdient.

Die Gründe, die sich für dieses Problem anführen lassen, können drei groben Überkategorien zugeordnet werden – sie beziehen sich auf a) Merkmale der Befragten, auf b) Merkmale der Fragen/Messung oder auf c) Merkmale der Interview- bzw. Testsituation.

### *Merkmale der Befragten*

Zu den *Befragtenmerkmalen*, welche Einschränkungen in der Datenqualität und damit auch bei den Ergebnissen implizieren, gehören eine Bandbreite unterschiedlicher subjektiver Befindlichkeiten. Im Kontext empirischer Untersuchungen können Befindlichkeiten auf verschiedene Arten Bedeutung erlangen. Zum einen könnten sich die Befragten von einer bestimmten Frage oder Aufgabe angegriffen fühlen. Ist dies der Fall, kann es sein, dass die Beantwortung der Frage bzw. die Bearbeitung der Aufgabe als unangenehm empfunden wird. Darüber hinaus kann die Antizipation der sozial „erwünschten“ und somit moralisch

---

„richtigen“ Antwort dazu führen, dass die Befragten keine wahrheitsgemäßen Aussagen über ihre Einstellungen, Erfahrungen und Wahrnehmungen erteilen, sondern sich stattdessen am antizipierten sozialen Desiderat orientieren.<sup>160</sup> Deutlich wird dieses Problem, wenn wir uns eine Situation vorstellen, in der die Blickbewegungen männlicher Probanden bei der Betrachtung weiblicher Aktdarstellungen aufgezeichnet werden. Soziale Erwünschtheit kann hierbei dazu führen, dass subjektive Meinungen oder natürlich-gewohnheitsmäßige Verhaltensweisen in Richtung der antizipierten sozialen bzw. moralischen Erwartungen verzerrt werden.

Eine solche Verzerrung kann *bewusst* geschehen, soziale Erwünschtheit kann allerdings auch *unbewusst* Einfluss auf das Antwortverhalten von Versuchsteilnehmenden nehmen. So ist es etwa möglich, dass dem Testleiter oder der Testleiterin von den Befragten implizit bestimmte Erwartungshaltungen unterstellt werden – beispielsweise was den Umfang der Antwort auf eine bestimmte (offene) Frage anbelangt. Das Bedürfnis nicht negativ aufzufallen, kann in diesem Fall dazu führen, dass eine Antwort umfangreicher ausfällt als es der betreffenden Erfahrung angemessen ist. Auch ist es in so einem Fall denkbar, dass, selbst wenn es für die Befragten eigentlich überhaupt nichts nennenswertes zu berichten gibt, dennoch der Versuch unternommen wird eine elaborierte Antwort zu geben.<sup>161</sup>

Ein weiteres Problem besteht darin, dass die meisten Personen nur ungern eingestehen, über bestimmte Sachverhalte keine Kenntnisse zu besitzen – selbst dann nicht, wenn es die Möglichkeit gibt eine „kann ich nicht sagen“-Antwort auszuwählen.<sup>162</sup> In Kombination mit der Tatsache, dass viele Fragen darauf abzielen, das Wissen über oder die Bewertungen von diversen Untersuchungsgegenständen zu erfassen, kann dies zu deutlich verzerrten Ergebnissen führen. Ein sehr spezifischer Fall dieser Antworttendenz ist die *Akquieszenz* von bestimmten Befragten. Dabei handelt es sich um eine Ja-Sage-Tendenz.<sup>163</sup> Damit ist gemeint,

---

<sup>160</sup> Vgl. Diekmann 2007: 448.

<sup>161</sup> Für eine größere Anzahl an Beispielen vgl. Diekmann 2007: 464-471.

<sup>162</sup> Ebd.: 479-488.

<sup>163</sup> Atteslander/Kneubühler 1975.

---

dass bestimmte Versuchsteilnehmende die Tendenz haben Fragen wie etwa „Haben Sie dieses Bild schon einmal gesehen?“ oder auch „Beschäftigen Sie sich in Ihrer Freizeit mit Kunst?“ grundsätzlich mit Ja zu beantworten.

Selbstverständlich gibt es jenseits der sozialen Erwünschtheit noch weitere Befragtenmerkmale, die sich negativ auf die erhobene Datenqualität auswirken können. Ohne weiter ins Detail gehen zu wollen, lassen sich an dieser Stelle Beispiele nennen, wie etwa die persönliche Motivation der Versuchsperson, an der Studie teilzunehmen, die Tagesform der Versuchsperson, ihre Müdigkeit, ihre derzeitige Stresssituation, ihre Konzentrationsfähigkeit, ihre Erfahrungen mit und Gewohnheiten im Ausfüllen von Fragebögen, ihre möglicherweise vorhandene Tendenz *keine* Extremwerte anzugeben, ihre Tendenz *immer* Extremwerte anzugeben, usw. Diese (und weitere) Aspekte können alle Einfluss nehmen auf das Antwortverhalten.<sup>164</sup>

Alle diese Probleme sind grundsätzlicher Natur. Jede empirische Studie, die mit der Befragung von Versuchspersonen arbeitet, sieht sich mit ihnen konfrontiert – die drei zur schwerpunktmäßigen Diskussion gewählten Studien stellen dabei keine Ausnahme dar. Auch hier muss davon ausgegangen werden, dass die Antizipation sozialer Erwünschtheit das Antwortverhalten der Versuchsteilnehmenden beeinflusst haben kann. So wurden etwa in allen drei Studien Fragen gestellt, die darauf abzielen zu ermitteln, ob die gezeigten Bildbeispiele den Teilnehmenden bereits vor der Studienteilnahme bekannt waren. Im Fall, dass bestimmten Versuchsteilnehmenden tatsächlich alle gezeigten Bildbeispiele unbekannt waren, ist es durchaus denkbar, dass dies wahrheitsgemäß zu berichten dem Bedürfnis zuwiderläuft, nicht ungebildet zu erscheinen.

In den Studien zwei und drei wurden darüber hinaus Fragen zu Kenntnissen und Wissen über Kunstwerke und Künstler sowie zur Rolle von kunstbezogener Freizeitgestaltung gestellt.

---

<sup>164</sup> Systematische Antwortmuster dieser Art, die unabhängig vom Inhalt der Fragen sind, werden auch als *Response-Set* bezeichnet (vgl. Diekmann 2007: 451-455).

---

Auch hier ist es denkbar, dass Erwartungshaltungen zur positiven sozialen Bewertung von Kunstaffinität bestehen, die einer wahrheitsgemäßen Beantwortung der Fragen in manchen Fällen hinderlich waren. Dies gilt umso mehr als die Durchführung der diskutierten Studien in allen drei Fällen an einer kunstwissenschaftlichen Einrichtung stattfand: Die Annahme, dass Kunstaffinität tatsächlich im Sinne sozialer Erwünschtheit ein Desiderat darstellt, lag für die Versuchsteilnehmenden nahe.<sup>165</sup>

Auch unbewusste Aspekte von sozialer Erwünschtheit mögen während der Datenerhebung eine Rolle gespielt haben. Die zuvor skizzierten impliziten Erwartungshaltungen, die in Bezug auf den Umfang von geforderten oder gewünschten Antworten bestehen, könnten insbesondere in der zweiten Studie zu einer Verzerrung der Ergebnisse geführt haben. Die im Rahmen dieser Studie durchgeführte qualitative Vorerhebung von Wahrnehmungsinhalten, basierte ja auf der freien Beantwortung von offen gestellten Fragen, wie zum Beispiel „Welche Gedanken hatten Sie bei der Betrachtung des Kunstwerkes?“. Keine der Versuchspersonen hat hier mit „keine spezifischen Gedanken“ oder schlicht „keine“ geantwortet. Ganz im Gegenteil, die Antworten waren in den meisten Fällen erstaunlich detailliert.<sup>166</sup>

Für die gewählten Studien sind die aufgezeigten Einflüsse sozialer Erwünschtheit auf die erhobenen Daten jedoch nicht einfach nur anzunehmen, sondern müssen, da alle drei Studien auch mit Kunstexpert\_innen gearbeitet haben, in speziellem Maße berücksichtigt werden. Kunstexpert\_innen sehen sich einem besonderen Druck ausgesetzt, wenn sie gebeten werden Fragen zu ihrem kunstbezogenem Wissens- und Kenntnisstand zu beantworten: sie finden

---

<sup>165</sup> Solche systematischen „Antwortfehler“, die aufgrund der Kenntnis des „Auftraggebers“ der Studie entstehen, werden auch als *Sponsorship-Effekt* bezeichnet (vgl. Diekmann 2007: 470).

<sup>166</sup> Dieses Problem lässt sich lösen, indem ‚meinungslose Personen‘ durch eine Filterfrage ausgeschlossen werden. Damit ist gemeint, dass vor der eigentlich interessierenden Frage zunächst gefragt wird: „Haben Sie eine Meinung zum Thema/Sachverhalt?“. Nur wenn diese Filterfrage mit ‚ja‘ beantwortet wird, werden die Teilnehmenden aufgefordert die eigentlich interessierende Frage zu beantworten (vgl. Diekmann 2007: 454). In den diskutierten Studien wurde eine solche Filtertechnik jedoch nicht verwendet.

---

sich gewissermaßen in einer Prüfungssituation wieder. Darüber hinaus liegt es nahe, dass Kunstexpert\_innen annehmen, dass an sie gestellte Erwartungshaltungen eine kunstbezogene Freizeitgestaltung inkludieren. Diese sozialen Erwartungshaltungen ließen sich, salopp formuliert, zusammenfassen mit: „Gute“ Kunstexpert\_innen müssen alles über Kunst wissen und mehrmals die Woche Ausstellungen besuchen. Die Tatsache, dass Kunstexpert\_innen sich dieser sozialen Erwartungshaltungen bewusst sind, mag durchaus dazu geführt haben, dass sie ihre Antworten im besonderen Maße „geschönt“ haben.<sup>167</sup>

Auch in Bezug auf die Blickbewegungsaufzeichnungen der zweiten und dritten Studie kann soziale Erwünschtheit eine Rolle gespielt haben; allerdings vermutlich nicht in einer so drastischen Form, wie im obigen Beispiel. Die gezeigten Bildbeispiele beider Studien stellen keine „sensiblen“ Inhalte dar. Eine systematische Verzerrung der Blickbewegungsdaten, die sich auf die bewusste Vermeidung bestimmter Bildregionen zurückführen lässt, können wir daher ausschließen. Es gibt jedoch andere Aspekte sozialer Erwünschtheit, die bei der Betrachtung der Bilder durchaus eine Rolle gespielt haben mögen. So ist es denkbar, dass die Tatsache, dass den Versuchsteilnehmenden bewusst war, dass sie sich in einer Testsituation befinden, dazu geführt hat, dass sie die Bilder aufmerksamer betrachtet haben als es ihren Gewohnheiten entsprechen würde. Dies hätte zur Konsequenz, dass die Daten nicht ein natürliches, sondern ein test-spezifisches Blickverhalten abbilden.<sup>168</sup>

Jenseits der sozialen Erwünschtheit können andere Merkmale der Befragten ebenfalls negative Auswirkungen gehabt haben. Vermutlich hatten auch Motivation, Müdigkeit und momentane Konzentrationsfähigkeit einen Einfluss auf das Antwortverhalten der Versuchspersonen. Genauso ist es wahrscheinlich, dass die Teilnehmenden jeweils subjektive Gewohnheiten im Umgang mit Fragebögen besitzen, welche sich ebenfalls auf das Antwortverhalten ausgewirkt haben.

---

<sup>167</sup> Vgl. auch Diekmann 2007: 466.

<sup>168</sup> Ebd.: 468.



---

Keines der in diesem Abschnitt skizzierten Probleme lässt sich vollständig eliminieren. Das heißt jedoch nicht, dass Ergebnisse empirischer Studien deswegen im Grunde genommen keine Aussagekraft besitzen. Viele dieser Probleme lassen sich durch bestimmte methodische Kniffe weitestgehend auf ein Minimum reduzieren, etwa indem Fragen mit sensiblem Inhalt vermieden und Erwartungshaltungen deutlich kommuniziert werden. Ein solche Kommunikation kann vor dem Beginn der eigentlichen Testung erfolgen. In der Regel werden Versuchsteilnehmende in diesem Zuge daran erinnert, dass es keine richtigen und falschen Antworten gibt, dass eine wahrheitsgemäße Beantwortung der Fragen für das Studienziel wichtig ist und, vor allem, dass die Studie die Anonymität der Versuchsteilnehmenden wahrt, die jeweiligen Antworten also nicht dem Namen einer Person zuordenbar sind. Neben der Kommunikation von Erwartungshaltungen können einige der beschriebenen Probleme auch bei der Auswertung und Interpretation der Daten berücksichtigt werden, so dass sie immerhin keine grundlegenden Fehlschlüsse zur Folge haben. Es lassen sich beispielsweise Gewohnheiten beim Ausfüllen von Fragebögen, wie die Tendenz entweder keine oder immer Extremwerte anzugeben, durch eine *Zentrierung* der Daten berücksichtigen.<sup>169</sup> Einflüsse wie Müdigkeit können ebenfalls bei der Auswertung berücksichtigt werden, sofern im Verlauf der Testung hierzu Antworten von den Versuchsteilnehmenden erfragt wurden. Für die diskutierten Studien war dies nicht der Fall. Hier wurden neben der Kommunikation der Erwartungshaltungen keine weiteren Schritte unternommen, um die Einflüsse von Befragtenmerkmalen zu reduzieren.

### *Merkmale der Fragen/Messung*

Spezifische Merkmale von einzelnen Fragen bzw. der durchgeführten Messung können ebenfalls Einfluss auf die Validität, Objektivität und Reliabilität von empirischen Methoden nehmen. Das Antwortverhalten von Befragten kann dadurch beeinflusst sein, *wie* bestimmte

---

<sup>169</sup> Laursen/Hoff 2006.

---

Fragen formuliert sind, in welcher *Reihenfolge* sie aufscheinen und welche *Antwortkategorien* den Befragten zur Verfügung stehen.

Die Art und Weise, wie Fragen formuliert sind, beeinflusst das Antwortverhalten zum Teil erheblich. Beispielsweise verleiten Fragen, die einen *suggestiven Charakter* haben, Befragte häufig dazu der durch die Fragenform kommunizierten sozialen Erwünschtheit nachzugeben und somit nicht ihrem eigenen Empfinden gemäß zu antworten.<sup>170</sup> Das Antwortverhalten kann auch durch die *Verwendung „starker“ Begriffe* beeinflusst werden. So konnten sozialwissenschaftliche Studien beispielsweise zeigen, dass es eine erhebliche Differenz im Antwortverhalten von Befragten gibt, wenn zwischen den Worten „verbieten“ und „nicht erlauben“ unterschieden wird.<sup>171</sup> Darüber hinaus ist es möglich, dass ein bestimmter Begriff nicht von allen Versuchsteilnehmenden im gleichen Sinne verstanden wird. Es ist beispielsweise nicht abwegig, dass es Unterschiede zwischen den Befragten hinsichtlich der *Interpretation von bestimmten Worten* gibt. Gerade Begrifflichkeiten, die keine Beschreibung darstellen, sondern sich auf abstrakte Konzepte beziehen, können in dieser Hinsicht zu Problemen führen. Normativ überformte Begriffe wie etwa Liebe, Glück oder Anstand basieren nicht zwingend auf einem kollektiv geteilten Verständnis, welches für alle sozialen Milieus gleichermaßen gilt.<sup>172</sup> Aber auch weniger stark soziokulturell geprägte Begriffe können von Befragten unterschiedlich verstanden werden. Die erste Studie hat beispielsweise gezeigt, dass *Komplexität* ein Begriff ist, der durchaus auf unterschiedliche Weise begriffen werden kann: Kunstlaien und Kunstexpert\_innen haben ein unterschiedliches Verständnis davon, welche formalen und semantischen Aspekte Komplexität in einem Kunstwerk bedingen. Ihre Antworten auf die Frage nach der an einem Kunstwerk wahrgenommenen Komplexität implizieren voneinander leicht abweichende Konzepte. Während dies im Fall der ersten Studie nicht weiter problematisch ist, da hier explizit die Frage nach ebendiesen Unterschieden gestellt wurde, verhält es sich für Studien,

---

<sup>170</sup> Diekmann 2007: 471-479.

<sup>171</sup> Ebd.

<sup>172</sup> Vgl. z. B. Schulze 1990.

---

die einen solchen Unterschied nicht untersuchen anders. Dort würde dieselbe Frage personenabhängig etwas Unterschiedliches messen. Dieses Problem lässt sich durchaus in Bezug zur Problematik der Nicht-Übersetzbarkeit von geisteswissenschaftlichen Konzepten setzen.

Für standardisierte Fragebögen kann sich ein weiteres Problem durch die *Auswahl von Antwortkategorien* ergeben. Die zuvor bereits angesprochene Anzahl der Antwortoptionen spielt nicht nur eine Rolle, weil bei steigender Anzahl die immer feineren Zwischenschritte ab einem gewissen Punkt nicht mehr sinnvoll gedeutet werden können;<sup>173</sup> sondern auch weil Befragte sich zur Beantwortung mancher Fragen an *Schätzungen* orientieren, für welche die vorgegebenen Antwortkategorien als Referenzrahmen dienen. Das ist besonders dann der Fall, wenn post-hoc Fragen zur Häufigkeit bestimmter Handlungen gestellt werden. Die meisten Menschen haben keine genaue Kenntnis über den tatsächlichen Umfang von alltäglichen Routinehandlungen, da diese als wiederkehrende Fixpunkte von geringer subjektiver Bedeutung sind und somit im episodischen Gedächtnis nur unscharf gespeichert werden.<sup>174</sup> Werden sie dennoch gebeten hierzu genaue Angaben zu machen (z. B. „Wie häufig sprechen Sie in ihrer Freizeit mit Freunden und Familie über Kunst?“), werden sie sich zur Beantwortung der Frage mit einer Schätzung behelfen müssen. Eine solche Schätzung kann sich, wie in vielen Studien gezeigt werden konnte, an den vorgegebenen Antwortkategorien orientieren.<sup>175</sup> Haben Befragte also den subjektiven Eindruck, dass sie eine bestimmte Handlung eher häufig vollziehen, werden sie sich tendenziell in derjenigen Kategorie verorten, die der gesetzten Skala gemäß ein „eher häufig“ ausdrückt – und zwar unabhängig davon, ob diese Antwortkategorie betitelt ist mit „monatlich“ oder „wöchentlich“.<sup>176</sup>

---

<sup>173</sup> Matell/Jacoby 1971.

<sup>174</sup> Diekmann 2007: 446-471.

<sup>175</sup> Schwarz et al. 1985.

<sup>176</sup> Vgl. ebd.

---

Über die Einflüsse von Formulierungen und Antwortkategorien hinaus kann, wie bereits erwähnt, auch die Reihenfolge der Fragen eine entscheidende Rolle spielen, da einzelne Fragen einen sogenannten *Halo-Effekt* auslösen (können).<sup>177</sup> Dieser beschreibt das Phänomen, dass durch manche Fragen bestimmte Gefühle, Gedanken oder Assoziationen angeregt werden, welche auf die Beantwortung der nachfolgenden Fragen „ausstrahlen“. Das kann zur Folge haben, dass Antworten die Befragte auf Folgefragen formulieren von bestimmten Aspekten der vorangegangenen Fragen beeinflusst werden. Ein im Zuge dieses Arguments oft bemühtes Beispiel ist die Kopplung der Fragen nach dem derzeitigen Beziehungsstatus einer Person und ihrer derzeitigen Lebenszufriedenheit.<sup>178</sup> Folgen diese beiden Fragen direkt aufeinander, so sind für die Befragten bei der Beantwortung der Frage nach der derzeitigen Lebenszufriedenheit noch Gedanken zum aktuellen Beziehungsstatus präsent. Je nachdem wie die befragte Person ihren derzeitigen Beziehungsstatus bewertet, können diesbezügliche Gedanken die Beantwortung der Folgefrage positiv oder negativ beeinflussen. Halo-Effekte dieser Art sind jedoch keineswegs an emotional bedeutsame Aspekte gebunden. Sie können auch bei völlig neutralen Themen auftreten. So würde etwa die Kopplung von Fragen nach semantischen Bildaspekten und wahrgenommener Komplexität die Wahrscheinlichkeit deutlich erhöhen, dass semantische Aspekte für die Komplexitätsbewertung stärker berücksichtigt werden als es vielleicht der Fall wäre, wenn die vorangegangenen Fragen keine solchen Gedanken angeregt hätten. Die Problematik des Halo-Effektes zeigt sich überdies nicht nur in Bezug auf die Reihenfolge von Fragen, sondern kann sich auch in Bezug auf die Reihenfolge gezeigter Bildbeispiele zeigen. Auch ein zuvor gesehenes Kunstwerk kann zu Gedanken und Assoziationen anregen, welche die Auseinandersetzung mit dem nächsten betrachteten Werk beeinflussen. Dies kann nicht nur die Beantwortung von Fragen beeinflussen, sondern selbstverständlich auch den Vorgang der Betrachtung. Somit kann der Halo-Effekt auch bei der Messung von Blickbewegungen ein Problem darstellen.

---

<sup>177</sup> Thorndike 1920.

<sup>178</sup> Murphy/Jako/Anhalt 1993.

---

Formulierungs- und Antwortkategorieeffekte lassen sich weniger eindeutig auf den Kontext von Blickbewegungsaufzeichnungen übertragen als der Halo-Effekt. Dennoch finden sich vergleichbare Effekte auch hier: die Art und Weise, wie der Betrachtungsvorgang in eine Aufgabe eingebettet ist, beeinflusst das Blickverhalten. So konnte beispielsweise Alfred L. Yarbus in einer empirischen Untersuchung von 1967 zeigen, dass ein und dasselbe Gemälde je nach Aufgabenstellung unterschiedliche Blickbewegungsmuster evoziert.<sup>179</sup> Eine freie, nicht durch eine bestimmte Aufgabe gerahmte Betrachtung führte beispielsweise zu einem viel weniger deutlich gerichtetem Blickverhalten als die Aufgabe die sozialen Beziehungen der dargestellten Personen zu erfassen. Neben diesen sehr konkreten auf das Explorationsverhalten gerichteten Aufgaben kann das Blickverhalten auch durch Aufgaben oder Fragen beeinflusst werden, die den Betrachtungsvorgang rahmen. So konnte beispielsweise gezeigt werden, dass Versuchspersonen ein jeweils anderes Blickverhalten an den Tag legen, wenn sie nach der Betrachtung der einzelnen Werke entweder um eine ästhetische oder um eine semantische Bewertung gebeten werden.<sup>180</sup> Durch die Kopplung von Bildbetrachtung und nachfolgender Frage wird demnach implizit an den Betrachtungsvorgang eine Aufgabe gestellt, auf welche Versuchsteilnehmende mit einem bestimmten (häufig unbewussten) Verhalten reagieren.

Einige der in diesem Abschnitt formulierten Probleme lassen sich mit Hilfe einer umsichtigen Versuchsplanung umgehen. Formulierungseffekte können durch die Vermeidung von Suggestivfragen und „starken“ Begrifflichkeiten sowie durch das Bereitstellen von Begriffserläuterungen auf ein Minimum reduziert werden.<sup>181</sup> Ein ausführlicher *Pretest* kann zudem Sicherheit bieten, dass die Maßnahmen zur Minimierung von Fehlern erfolgreich waren. Ein solcher *Pretest* ist im Wesentlichen eine Testung des Messinstrumentes. Hierbei werden, beispielsweise anhand von *Think-Aloud-Techniken*, Vorgehensweisen bei der Beantwortung der Fragen und Verständnisprobleme aufdeckt.<sup>182</sup> Auch das Verständnis von

---

<sup>179</sup> Yarbus 1967.

<sup>180</sup> Molnar/ Ratsikas 1987.

<sup>181</sup> Diekmann 2007: 442 und 449.

<sup>182</sup> Ebd.: 415f. und 483-488.

---

Antwortkategorien lässt sich mit einem Pretest evaluieren. In der Regel wird ein Messinstrument so lange getestet und gegebenenfalls angepasst, bis keine fragenbezogenen Fehler mehr auftreten. Die drei zur Diskussion stehenden Studien stellen hierbei keine Ausnahme dar.

Weniger einfach zu beheben, sind Probleme die in Bezug auf Reihenfolgeeffekte bestehen. Hier ist es zwar manchmal möglich das Problem zu umgehen, indem – wie im Falle der ersten Studie – bestimmte Fragen in entkoppelter Form gestellt werden, doch ist eine solche Entkopplung nicht in jedem Fall eine sinnvolle Option. Die Studien zwei und drei etwa untersuchen neben dem Blickverhalten auch die bei der Betrachtung evozierten Wahrnehmungsinhalte. Eine (zeitliche) Entkopplung von Bildbetrachtung und diesbezüglichen Fragen würde bedeuten, dass sich die Befragten bei der späteren Beantwortung der Fragen an den jeweiligen Betrachtungsvorgang zurückerinnern müssten. Dies ist umso schwieriger, je mehr Bildbeispiele die Befragten im Zuge der Untersuchung betrachtet haben. Außerdem steigt mit wachsender zeitlicher Distanz auch die Wahrscheinlichkeit, dass wichtige Aspekte nicht oder sogar falsch „erinnert“ werden. Zeigt man zur Erleichterung der „Erinnerungsaufgabe“, wie im Fall der ersten Studie dasselbe Bild zweimal, so lässt sich nicht mehr unterscheiden, ob die berichteten Wahrnehmungsinhalte sich aufgrund der vorangegangenen Betrachtung oder aufgrund der zweiten Betrachtung ergeben haben. Im Fall der zweiten und dritten Studien sind Fragenreihenfolgeeffekte daher das „kleinere Übel“; insbesondere da auch hier Pretests helfen können einige Interaktionen zu identifizieren, weswegen sie bei der Auswertung und Interpretation der erhobenen Daten berücksichtigt werden können.

#### *Merkmale der Interview- bzw. Testsituation*

Die Interview- bzw. Testsituation stellt den Kontext der Datenerhebung dar. Zu diesem Zählen sowohl räumliche als auch soziale Aspekte. Der räumliche Kontext umfasst alle mit dem Ort der Befragung in Verbindung stehenden Aspekte. Der soziale Kontext hingegen

---

umfasst die Beziehung und Interaktion zwischen den während der Befragung anwesenden Personen. Sowohl räumlicher wie auch sozialer Kontext wirken auf die Befragten und können deswegen Einfluss auf deren Antwortverhalten nehmen.

Der räumliche Kontext der Befragung, also der Ort an welchem die Datenerhebung stattfindet, kann in besonderem Maße bedeutsam für die Erfahrung von Untersuchungsgegenständen sein.<sup>183</sup> Obwohl dies im Allgemeinen für alle Untersuchungsgegenstände gilt, kann dieser Umstand für die Erfahrung von Kunstwerken nicht genug betont werden. Kunsterfahrung findet außerhalb von Test- und Datenerhebungssettings in der Regel in ganz *bestimmten* räumlichen Settings statt: wir betrachten Kunst in Museen, Galerien oder öffentlichen Bauwerken, wie beispielsweise Kirchen. Dies sind die ökologischen Settings, in welchen Kunst betrachtet wird und in welchen eine Auseinandersetzung mit den Werken normalerweise stattfindet. Die meisten empirischen Datenerhebungen finden jedoch nicht in einem solchen „natürlichen“ Umfeld statt, sondern in speziellen Testlaboren. In der Regel sind Testlabore verhältnismäßig sterile Umgebungen, die wenig Ablenkung durch Geräuschkulisse, andere Personen oder auch Einrichtung bieten. Kunst wird hier nicht in Originalgröße an Wänden hängend betrachtet, sondern in digitaler Reproduktion auf einem Computerbildschirm. Die Befragten sitzen dabei meist an einem Schreibtisch und klicken sich durch einen vorbereiteten Testablauf, bei dem sich Fragen und Bilder in unterschiedlichen Rhythmen abwechseln (vgl. Abbildung 18). Der Unterschied zwischen „natürlicher“ Situation und Testsituation ist demnach gravierend. Nicht nur sind die Räumlichkeiten selbst völlig andere, auch die Untersuchungsobjekte präsentieren sich in „unnatürlicher“ Form. Größenverhältnisse, Textur und Materialität beispielsweise können in klassischen Testsituationen kaum erfahren werden.

Die „Aura des Originals“, um es in Anlehnung an Walter Benjamin auszudrücken, fehlt in den meisten Testsituationen.<sup>184</sup> Was Benjamin bereits 1935/36 formulierte, wurde jüngst

---

<sup>183</sup> Diekmann 2007: 468.

<sup>184</sup> Vgl. Benjamin 2012.

---

auch empirisch belegt: die Erfahrung von Originalen unterscheidet sich deutlich von der Erfahrung (digitaler) Reproduktionen. In einem Vergleich von Museums- und Laborsetting konnte gezeigt werden, dass die jeweilige Erfahrung von Kunstwerken im natürlichen Kontext positiver und darüber hinaus in allen untersuchten Aspekten stärker ausgeprägt war.<sup>185</sup> Dieser Unterschied lag dabei nicht ausschließlich daran, dass die Werke im Testlabor in der Regel in Reproduktion gezeigt werden. Der Vergleich der Erfahrung von Originalen in ihrem natürlichen Kontext mit der Erfahrung von Originalen im Laborkontext erbrachte dasselbe Resultat: die im natürlichen Kontext gemachte Erfahrung war positiver und stärker ausgeprägt jene im Labor. Im Vergleich aller drei Optionen zeigte sich das „schlechteste“ Ergebnis im Fall der Labortestung mit digitalen Reproduktionen.<sup>186</sup> Die schlichte Tatsache, dass Testungen im Labor stattfinden, ob nun anhand von Reproduktionen oder Originalen, verändert also die Erfahrung von den untersuchten Objekten. Für den Fall von Kunstwerken haben die empirischen Untersuchungen aber zumindest zeigen können, dass das Laborsetting die Erfahrung nicht in etwas völlig anderes verkehrt. Die Erfahrung wird lediglich deutlich abgeschwächt. Dieser Umstand lässt sich kaum beheben, da er durch die Testsituation selbst bedingt ist – er lässt sich aber in der Interpretation der Ergebnisse berücksichtigen. Es sollte stets bedacht werden: je „unnatürlicher“ die Testsituation, desto weniger repräsentativ werden die Daten für die „reale“ Situation sein.<sup>187</sup>

Die Testung oder Messung im Laborkontext hat jedoch nicht nur Nachteile. So können im Labor beispielsweise die Testbedingungen für alle Befragten konstant gehalten werden. Aspekte wie zum Beispiel Beleuchtung, Temperatur, Setting und Betrachtungsdauer können im Labor genormt werden, sodass die Datenaufzeichnungen in dieser Hinsicht für alle Befragten unter denselben Bedingungen stattfinden. Das hat den Vorteil, dass ausgeschlossen werden kann, dass hinsichtlich dieser konstant-gehaltenen Aspekte subjektive verschiedene Einflüsse auf die Daten entstehen. Auch weitere negative Einflüsse

---

<sup>185</sup> Brieber/Nadal/Leder 2015, Brieber et al. 2014.

<sup>186</sup> Ebd.

<sup>187</sup> Vgl. auch Diekmann 2007: 468.



---

auf die erhobenen Daten können durch den Laborkontext reduziert werden. So kommt es beispielsweise vor, dass qua Ablenkung (die Anwesenheit Dritter, Geräusche, etc.), zusammenhangslose Inhalte assoziiert werden, die den Lauf eines Gedankengangs unterbrechen. Solche negativen Einflüsse lassen sich in der kontrollierten Umgebung eines Labors ausschließen. In der Konsequenz sind die im Laborsetting erhobenen Daten besser vergleichbar als die Daten des „natürlichen“ Settings.

Wie oben erwähnt beeinflusst auch der soziale Kontext der Testsituation die Datenerhebung. Dieser beschreibt die Beziehung und Interaktion zwischen den während der Befragung anwesenden Personen, also in der Regel zwischen der Testleiter\_in und den Befragten. Dabei spielen die spezifischen Merkmale der Testleiter\_in eine große Rolle. Sprache, Kleidung, Alter, Mimik, Gestik, usw. können Ausdruck von sozialer Distanz oder Nähe zwischen Testleiter\_in und Befragten sein. Wahrgenommene soziale Nähe oder Distanz kann sich auf das soziale Wohlbefinden in der Testsituation auswirken. Darüber hinaus wird das (soziale) Wohlbefinden auch dadurch beeinflusst, ob die Anwesenden einander sympathisch oder unsympathisch sind.<sup>188</sup> Diese sozialen Aspekte können insofern Einfluss auf das Antwortverhalten nehmen, als sie die Motivation, an der Studie teilzunehmen, erhöhen oder reduzieren können. Sie können weiter dazu beitragen, Aspekte sozialer Erwünschtheit stärker zu betonen oder auch abzuschwächen. Darüber hinaus kann das (soziale) Wohlbefinden auch im ganz Allgemeinen das Antwortverhalten beeinflussen, weil ein angenehmes soziales Umfeld die Konzentration auf die Studie und die gestellten Fragen erleichtert. Sozialer Stress führt, wie auch in jedem anderen Kontext, in Testsituationen zu einer Verzerrung des „natürlichen“ Verhaltens.<sup>189</sup> Dabei ist es gleichgültig, ob es sich um positiven sozialen Stress handelt, der in der sozialen Situation das Bedürfnis nach sozialer Anerkennung steigert, oder um sozialen Stress negativer Art, der dazu führt, dass die (soziale) Testsituation als unangenehm bewertet wird, was Konzentration und Motivation ebenfalls stören kann. Da soziale Interaktionen immer und automatisch einer Bewertung unterworfen werden, ist es –

---

<sup>188</sup> Vgl. Diekmann 2007: 466-470.

<sup>189</sup> Diekmann 2007: 466ff.

---

außer im Fall einer postalischen Befragung – unmöglich sämtliche sozialen Einflüsse auf die Testung zu eliminieren. Wird die Interaktion zwischen Testleitung und Befragten allerdings so gering wie möglich gehalten, reduziert sich auch der Einfluss des sozialen Kontextes. Praktisch lässt sich dies etwa durch eine größtenteils schriftliche Einführung sowie durch eine räumliche Trennung von Testaufsicht und Testort umsetzen. Insbesondere die Trennung von Aufsichtsort und Testort führt dazu, dass Befragte während der Teilnahme ungestört sind und sich keiner direkten Beobachtung ausgesetzt fühlen.<sup>190</sup>

Die in diesem Abschnitt skizzierten Einflüsse auf die Datenerhebung stellen insbesondere für die vorgestellten Studien ein Problem dar. Die erhobenen Daten haben den Nachteil, in einem wenig „natürlichen“ Kontext entstanden zu sein, gleichzeitig aber den Vorteil, sich durch gute Vergleichbarkeit auszuzeichnen. Der soziale Kontext wurde in allen Fällen so neutral wie möglich gestaltet. Einflüsse der Testsituation, vor allem räumlicher Art, sind dennoch insbesondere für die beiden Eye-Tracking Studien anzunehmen. Sie müssen daher bei der Auswertung und – vor allem – der Interpretation der Daten berücksichtigt werden.

Abschließend lässt sich festhalten, dass die meisten empirischen Testungen ausgesprochen *reaktive* Erhebungsmethoden darstellen.<sup>191</sup> Womit gesagt sein soll, dass die Modalitäten der Testung Einfluss auf das Antwortverhalten von Befragten und damit auch auf die Daten nehmen. Die erhobenen Daten können damit niemals als objektive Repräsentationen des untersuchten Problems verstanden werden. Vielmehr stellen sie vielfach eingeschränkte Näherungswerte dar, deren Einschränkung bei der Interpretation und weiteren Verwendung der Ergebnisse Berücksichtigung finden sollte.

---

<sup>190</sup> Vgl. Diekmann: 514.

<sup>191</sup> Ebd.: 434 und 627.

---

### *3.3 Auswertung und Interpretation*

Trotz der vielen Probleme, die bei der Datenerhebung auftreten können, bestehen die größten Schwierigkeiten von empirischen Ansätzen in der Auswertung und Interpretation der erhobenen Daten. Das liegt nur zu einem gewissen Teil in den Auswertungsmethoden selbst begründet. Viele der vordergründigen Probleme sind eindeutig eher den gängigen Analysepraxen sowie dem weiteren Umgang mit den Ergebnissen geschuldet. Die Tragweite dieser Auswertungs- und Interpretationsproblematik ist zum Teil gravierend und die von der negativen Kritik an empirischen Ansätzen mit Nachdruck betonten Einwände, dass empirische Methoden nicht an-sich die Herstellung objektiver, valider und vor allem generalisierbarer Ergebnisse garantieren, finden hier ihren wichtigsten Ansatzpunkt. Im Folgenden werde ich diesen wesentlichen Kritikpunkt näher beleuchten, indem ich die in den gewählten Studien zur Anwendung kommenden Auswertungs- und Interpretationsmethoden kritisch hinterfrage. Zu diesem Zweck werde ich zunächst zwischen den unterschiedlichen Ansätzen unterscheiden und jeweils spezifische Probleme diskutieren, die bei der Auswertung und Interpretation jeweils auftreten (können). Ich unterscheide zwischen qualitativen Zugängen mit offenen Fragen, welche zur Messung eines bestimmten Konzeptes angelegt werden, physiologischen Eye-Tracking Messungen und Problemen, die sich in der Auseinandersetzung mit standardisierten Fragen ergeben. Letztere werde ich am Beispiel der Messung von Expertise besprechen, da die gewählten Studien dank ihrer drei unterschiedlichen Ansätze diesbezüglich eine illustrative Bandbreite bieten. Im Anschluss an diese Diskussion werde ich allgemeine Probleme statistischer Auswertungs- und Interpretationsmodi besprechen, welche über die spezifischen Probleme der jeweiligen Methoden hinaus bestehen. Abschließend werde ich auf die Publikationsform von empirischen Ergebnissen zu sprechen kommen, um aufzuzeigen, wie sich der Umgang mit den zuvor beschriebenen Problemen in der Praxis niederschlägt.

---

## *Offene Fragen*

Offene Fragen, wie sie zum Beispiel in der Vorstudie der zweiten Studie zur Anwendung gekommen sind, haben den Vorteil, dass sie die Befragten nicht durch vorgegebene Kategorien in ihren Antwortmöglichkeiten einschränken. Die Befragten können in ihren eigenen Worten, in einem von ihnen völlig frei wählbaren Umfang auf die Fragen antworten. Offene Fragen können die bewusst erlebten Wahrnehmungsinhalte der Befragten somit besser abbilden als standardisierte Fragen mit vorgegeben Antwortkategorien. Trotz dieses entscheidenden Vorteils bestehen wegen des qualitativen Zugangs zu offen gestellten Fragen einige Probleme. Da die Befragten in ihren Antwortmöglichkeiten völlig frei sind, lässt sich im Vorfeld kaum ein konkretes Auswertungsschema erarbeiten, welches garantieren würde, dass die Interpretation der Ergebnisse gemäß diesen Regeln als objektiv verstanden werden kann. Statt einer solchen „objektiven“ Interpretationsschablone, folgt die Auswertung solcher Daten in der Praxis einem eher „weichen“ Zugang. Damit ist im Wesentlichen gemeint, dass die Daten durch die Forschenden subjektiv interpretiert werden. Ob Gemeinsamkeiten oder Unterschiede in den Äußerungen und Beschreibungen der Befragten vorliegen, wird von den Forschenden auf der Basis ihrer eigenen Empfindungen interpretiert. Im Konkreten werden dazu die Äußerungen der Befragten in Sinnabschnitte (Codes) unterteilt. Diese Sinnabschnitte oder Codes können in weiteren Interpretationsschritten zu größeren Gruppen oder Kategorien zusammengefasst werden, sofern sie ähnliche Aspekte beschreiben.<sup>192</sup> Es obliegt also den Forschenden zu entscheiden, ob, und wenn ja, wie stark zwei Aussagen inhaltlich übereinstimmen. Im Falle der zweiten Studie wurden etwa unter der Kategorie „Dynamik“ eine ganze Reihe von Äußerungen gefasst, die den Begriff selbst jedoch nicht verwenden. Dazu gehören beispielsweise die Worte „aktiv“ und „bewegt“. Während dies auf den ersten Blick nicht unbedingt eine problematische Zuordnung darstellt, können die Befragten im Detail mit ihrer jeweiligen Aussage dennoch etwas anderes gemeint haben. Im Umgang mit qualitativen Daten stellt die subjektive Interpretation der Forschenden

---

<sup>192</sup> Für eine ausführliche Übersicht zu qualitativen Auswertungsmethoden und den jeweiligen Unterschieden spezifischer Zugänge siehe Bohnsack 2008.

---

also eine mögliche Fehlerquelle dar. Um die Fehler so klein wie möglich zu halten, werden die Interpretationsschritte daher meistens von einer größeren Anzahl an Forschenden durchgeführt, die unabhängig voneinander die Aussagen der Befragten kodieren und erst in einem zweiten Schritt ihre Ergebnisse konsolidieren.<sup>193</sup> Damit ist sichergestellt, dass die Interpretation zumindest in einem gewissen Maß auf einer intersubjektiven Perspektive beruht.

Jedoch sind auch diese intersubjektiven Perspektiven nicht völlig unproblematisch. Forschende, die gemeinsam an einem Projekt arbeiten, teilen in der Regel bestimmte Auffassungen in Bezug auf Fragestellung und relevante Theoriekonzepte. Die jeweiligen Interpretationen werden daher von einer gemeinsamen Perspektive auf das erhobene Material beeinflusst sein. Das kann dazu führen, dass bestimmte Aspekte in den Daten eine besondere Betonung erfahren, während andere weitestgehend außen vor bleiben, obwohl sie für das Verständnis des untersuchten Phänomens von Relevanz wären.<sup>194</sup> Es ist daher für die Kommunikation qualitativer Forschungsergebnisse von großer Wichtigkeit sowohl die vorgenommenen Interpretationsschritte als auch die jeweiligen Perspektiven offenzulegen.

### *Standardisierte Fragen*

Die Interpretationsproblematik ist bei standardisierten Fragen weniger stark ausgeprägt als bei offenen Fragen, da die Antworten der Befragten hier in der Regel in klare disjunkte Kategorien unterteilbar sind. Dies ist sowohl ein Vorteil als auch ein Nachteil von standardisierten Fragen. Es ist von Vorteil, weil die Interpretation der Daten weniger fehleranfällig ist; und es ist von Nachteil, weil die erhobenen Daten ein viel reduktionistischeres Bild zeichnen. Dieser Reduktionismus stellt jedoch nur dann einen echten Nachteil dar, wenn er in der Auswertung und Interpretation der Daten nicht gebührend

---

<sup>193</sup> Vgl. Diekmann: 545.

<sup>194</sup> Vgl. Bohnsack 2008: 194f.

---

berücksichtigt wird. In der Regel bilden die meisten statistischen Auswertungsverfahren die Unvollständigkeit des gemessenen Konzeptes hinreichend ab (siehe Abschnitt Statistik weiter unten). Viel problematischer als die sich durch standardisierte Fragen ergebende reduktionistische Perspektive ist jedoch die inhaltliche Interpretation der gesetzten Kategorien.

Im vorangegangenen Abschnitt zur Datenerhebung wurde bereits darauf hingewiesen, dass vorgegebene Kategorien Einfluss auf das Antwortverhalten von Befragten nehmen können. Die inhaltliche Interpretation der Daten sollte diese Problematik berücksichtigen. Das ist nur möglich, wenn die gesetzten Antwortkategorien nicht als absolut verstanden werden. Damit ist beispielsweise gemeint, dass eine Einschätzung, die besagt, dass die Komplexität eines bestimmten Kunstwerkes auf einer Skala von 1 bis 7 bei etwa 5 liege, nicht bedeutet, dass die Komplexität des Kunstwerkes tatsächlich den Wert 5 besitzt. Eine solche Einschätzung besagt viel mehr nur, dass gegeben der gesetzten 7er-Skala und den Vergleichsbeispielen die für die Bewertung zugrunde gelegt werden der Komplexitätswert bei etwa 5 liegt. Die Komplexität dieses bestimmten Werkes wurde damit als höher eingeschätzt als für Werke, denen der Wert 3 zugeordnet wurde. Ändern sich Bewertungsskala oder Vergleichsbeispiele, ist es durchaus auch möglich, dass die Bewertung eines bestimmten Kunstwerkes ebenfalls anders ausfällt. Die einzige Aussage, die sich den vergebenen Werten also tatsächlich entnehmen lässt, ist, dass das Werk mit dem Komplexitätswert 5 als komplexer bewertet wird als das Werk mit dem Komplexitätswert 3; nicht aus den Daten ableiten lässt sich hingegen der tatsächliche Komplexitätswert. Dieser wird je nach angelegter Skala variieren und ist damit nicht objektiv fassbar. Die Interpretation quantitativer Fragebogendaten ist demnach immer als Beschreibung relationaler Beziehungen zu verstehen, nicht aber als Erfassung objektiver Wahrheiten. Das gilt für alle Fragen, die im weiteren Sinne Antwortkategorien in Form von Abstufungsskalen vorgeben.<sup>195</sup>

---

<sup>195</sup> Das hier angesprochene ‚Problem‘ lässt sich auf die spezifischen Merkmale der Daten zurückzuführen. Ein Datenkorpus unterteilt sich in der Regel in verschiedene Variablen. Dabei bezeichnet das Wort ‚Variable‘ ein bestimmtes Merkmal oder eine bestimmte Eigenschaft, die in den Daten ausgedrückt ist. Komplexität wäre eine solche Variable. Variablen sind, wie der Begriff

---

Bei standardisierten Fragen, die dichotome Antwortoptionen anbieten, verhält es sich ein wenig anders. Beispielsweise ließe sich die offene Frage nach der wahrgenommenen Komplexität eines Kunstwerkes auch als Ja/Nein-Frage formulieren. In diesem Fall würden die Befragten eine subjektive Trennlinie für sich definieren. Alle Bildbeispiele, die gefühlsmäßig *unterhalb* dieser Trennlinie liegen, würden als nicht komplex eingestuft werden; alle Bildbeispiele, die gefühlsmäßig *oberhalb* der Trennlinie liegen, würden hingegen als komplex gelten. Der Nachteil einer solchen Messung besteht darin, dass weder exakt ermittelbar ist, wo diese subjektive Trennlinie genau liegt, noch ob alle Befragten dieselbe oder zumindest eine ähnliche Trennlinie zugrunde legen. Eine dichotome Messung der wahrgenommenen Komplexität würde somit zwar einerseits eindeutig die Unterscheidung zwischen Werken erlauben, die komplex wahrgenommen werden und solchen die nicht als komplex wahrgenommen werden; gleichzeitig würde die Messung aber die Interpretierbarkeit und Vergleichbarkeit der Daten reduzieren. Außerdem würden durch eine dichotome Erfassung der Erfahrung auch deutlich weniger Informationen abgebildet werden. So ließen sich etwa die Werke der Kategorie „komplex“ nicht weiter unterscheiden, in beispielsweise „komplex“ und „sehr komplex“.

Die Entscheidung, in welcher Form bestimmte Konzepte gemessen werden, welche Kategorien angelegt werden, ist – wie das Komplexitätsbeispiel illustriert – wesentlich für

---

bereits nahelegt, *variabel* – d.h. sie sind nicht immer gleich *ausgeprägt*. Eine Variable kann viele unterschiedliche Ausprägungen haben (ein Kunstwerk kann beispielsweise hoch-, mittel- oder auch niedrig-komplex sein). Variablen sind darüber hinaus unterscheidbar in unterschiedlichen „Typen“. Die Variablentypen (meist *Skalenniveau*) sind dabei abhängig von den möglichen Ausprägungen der Variable. Je nachdem wie bedeutsam die Abstände zwischen zwei Ausprägungen sind, lassen sich die Antworten einer Person entweder im Sinne des tatsächlich vergebenen Wertes interpretieren oder, wie in dem hier vorliegenden Fall, nur im Sinne ihrer relationalen Beziehung. Zur Verdeutlichung: der Abstand zwischen 200 und 300 ist exakt bestimmbar. Er beträgt 100. Der Abstand zwischen den Kategorien ‚nicht-komplex‘ und ‚*eh*r nicht-komplex‘ ist hingegen nicht exakt bestimmbar. Zwar ist eindeutig, dass die beiden Ausprägungen in einer Hierarchie stehen und dass der eine Ausdruck „mehr“ ist als der andere, doch können wir nicht *eindeutig* feststellen *wie viel* „mehr“. Die Interpretation der Unterschiede ist daher auf die Feststellung der Hierarchie beschränkt. Für eine detaillierte Diskussion von Skalenniveaus und deren exakter Interpretation siehe Diekmann 2007: 285-302.

---

die Interpretation der Daten. Dabei zeigt schon die Tatsache, dass es hier überhaupt Wahlmöglichkeit *gibt*, dass die Entscheidung über die exakte Form der Messung Ausdruck einer bestimmten Perspektive der Forschenden auf das Problem ist. Insofern weisen die gesetzten Kategorien, die ja auch die Grundlage der späteren Interpretation bilden, durchaus Parallelen zur Interpretationsproblematik offener Fragen auf. Welche Option der Kategoriensetzung letztlich gewählt wird, hängt häufig von den Details der Fragestellung ab. Da in den drei gewählten Studien die Konzeptualisierung und Messung von *Expertise* anhand jeweils unterschiedlicher Mittel erfolgte, werde ich im Folgenden die sich durch die Kategoriensetzung ergebenden Einschränkungen für Auswertung und Interpretation der Daten an diesem Beispiel verdeutlichen.

Soll *Expertise* untersucht werden, steht zu Anfang die Frage nach dessen Definition. Es muss also geklärt sein, was genau unter Kunstexpertise zu verstehen ist bzw. welche Merkmale hinreichend für die Zuschreibung eines *Expert\_innenstatus* sind. Hier wird durch die Forschenden bereits die erste wesentliche, durch eine bestimmte Perspektive auf den Sachverhalt geprägte Entscheidung getroffen. In den drei diskutierten Studien wurden unterschiedliche Strategien zugrunde gelegt. In der ersten Studie wurde zwischen Kunstexpert\_innen und Laien unterschieden, indem das Vorhandensein einer formalen Ausbildung als Unterscheidungskriterium zugrunde gelegt wurde. Die Problematik der genauen Merkmalsbestimmung von Kunstexpertise wurde hier also umgangen, indem *Expertise* schlicht über soziale Marker (z.B. Bildungsabschlüsse und berufliche Position) erfasst wurde. Solche sozialen Marker stellen in der Regel verlässliche Indikatoren für Kenntnisse und Kompetenzen dar, obgleich sie diese nicht notwendigerweise im Detail transparent machen. Eine dichotome Erfassung anhand von sozialen Markern ist daher oft unproblematisch – trotz der in diesem Erhebungsmodell fehlenden Rückversicherung über die tatsächlichen Kompetenzen und Kenntnisse der Befragten.

In vielen Fällen ist der Rückgriff auf „echte“ *Expert\_innen* jedoch nicht möglich. Tatsächlich arbeiten die meisten empirischen Studien hauptsächlich mit Studierenden unterschiedlicher



---

Fachrichtungen – so auch die hier vorgestellten Studien zwei und drei. Das hat zur Folge, dass eine genauere Auseinandersetzung mit den Expertise-bedingenden Merkmalen notwendig wird. Sollen Kunstgeschichte-Studierende im Kontext einer Studie als Kunstexpert\_innen geführt werden, muss vorab geklärt werden, ob sie bereits hinreichende Kenntnisse und Kompetenzen erworben haben. Eine schlichte Unterscheidung zwischen Kunstgeschichte-Studierenden und Studierenden anderer Fachrichtungen reicht in diesem Fall nicht mehr aus – es werden weitere Informationen benötigt. Eine naheliegende Lösung wäre es natürlich, auf die Anzahl der studierten Semester (Semesterzahl) zurückzugreifen. Bei näherer Betrachtung zeigt sich allerdings, dass diese allein ebenfalls wenig Aussagekraft besitzt. Zwar ist davon auszugehen, dass Kunstgeschichte-Studierende in niedriger Semesterzahl noch weit weniger Expertise erworben haben als Kunstgeschichte-Studierende höherer Semesterzahlen; doch ist es auch denkbar, dass Studierende mit unterschiedlicher Motivation und auch abweichendem Erfolg studieren. Ebenfalls möglich ist, dass Studierende, die nicht in einem Studienprogramm für Kunstgeschichte eingeschrieben sind, ein großes Interesse an Kunst und Kultur entwickeln und sich autodidaktisch Kenntnisse über Kunst aneignen. Für den Fall, dass keine „echten“ Expert\_innen für die Untersuchung gewonnen werden können, ist es also nötig, Kunstexpertise differenzierter zu messen als über die Zuordnung zu einer von zwei Kategorien.

Eine differenzierte Messung von sozial bedeutsamen Konzepten, wie Expertise, erfolgt in der Regel durch die Erhebung von drei Komponenten: der Einstellung zum interessierenden Konzept, dem Wissen darüber und typischen konzeptbezogenen Aktivitäten.<sup>196</sup> Im Fall der Studien und drei wurde Expertise anhand eines Fragebogens gemessen, der diese drei Aspekte abdeckt.<sup>197</sup> Es handelt sich dabei um eine Vorstufe des *VAlAK Fragebogens*, eines Fragebogens zur Erfassung von Kunstwissen und Kunstinteresse.<sup>198</sup> Obwohl das Konzept „Expertise“ in beiden Studien auf die gleiche Art gemessen wurde, erfolgte die Auswertung

---

<sup>196</sup> Diekmann 2007: 230ff.

<sup>197</sup> Siehe Anhang *Expertisefragebogen*.

<sup>198</sup> Specker et al. *in press*.

---

und Einbindung in die Untersuchung auf unterschiedliche Weise. In der zweiten Studie wurden einzelne Aspekte als jeweils eigene Einflussfaktoren aufgenommen; in der dritten Studie wurden die erhobenen Expertisemerkmale, wie oben bereits beschrieben, in Form eines kontinuierlichen Punkte-Index in die Untersuchung aufgenommen. An den Kategorisierungs- und Auswertungsentscheidungen der drei Studien lässt sich gut nachvollziehen, welchen eklatanten Unterschied beide Entscheidungen für die Ergebnisse und deren Interpretation nach sich ziehen. Obwohl in allen drei Fällen Expertise gemessen wurde, wurde sie immer anders gemessen oder zumindest auf abweichende Art in die Auswertung der Daten eingebunden. Die sich unterscheidenden Auswertungsmethoden haben zur Folge, dass die Ergebnisse in allen drei Fällen andere Schlüsse über Expert\_innen erlauben.

Die Ergebnisse der ersten Studie erlauben es, zwischen Expert\_innen und Laien als groben Kategorien zu unterscheiden. Die Gruppen bleiben in-sich jedoch weitestgehend undifferenziert und ermöglichen keine weiteren Schlüsse über die Expertise der Versuchsteilnehmer\_innen. Die zweite Studie erlaubt es, zwischen den einzelnen Expertisemerkmale und ihren jeweils spezifischen Einflüssen zu unterscheiden. Das hat den Vorteil, dass dadurch der Einfluss des Konzeptes Expertise besser beschreibbar wird. Hier wird beispielsweise abbildbar, dass die kunsthistorische Ausbildung, unabhängig vom erworbenen *Wissen* über Kunst, dazu führt, dass Bildern mit einem ganz bestimmten Blickverhalten begegnet wird. Offenbar lernen Kunsthistoriker\_innen im Studium, Bilder auf eine bestimmte Art und Weise zu betrachten. Dieser „Expert\_innenblick“ ist nicht notwendigerweise an umfangreichen Wissen gekoppelt. Es ist durchaus möglich, dass Expert\_innen Kompetenzen dieser Art bereits früh im Studium erwerben, während andere Wissens- und Erfahrungswerte erst später gebildet werden. Die dritte Studie gestattet eine solche Unterscheidung nicht. Jedoch wird hier ein anderes vorteilhaftes Merkmal genutzt: die Entscheidung, Expertise in Form eines kontinuierlichen Punkte-Indexes zu erheben, ermöglicht es abzubilden, ob sich Menschen mit steigender Expertise immer weiter

---

unterscheiden bzw. ab welchem Punkt Expertise einen entscheidenden Unterschied ausmacht und ab wann ein Zugewinn an Expertise *keinen* weiteren Unterschied mehr zur Folge hat.<sup>199</sup>

Die Art der Messung und Auswertung erlaubt also nur Antworten auf sehr spezifische Fragen. Dies zu betonen, ist auch deswegen wichtig, weil es gleichzeitig bedeutet, dass die Ergebnisse unterschiedlicher Studien nicht in vollem Umfang vergleichbar sind, selbst wenn sie im Kern dasselbe Konzept adressieren.

### *Eye-Tracking*

Die Auswertung von Eye-Tracking Messungen weist im Vergleich zur Auswertung von Fragebögen mehrfach-gestaffelte Probleme auf. Diese beginnen zuallererst damit, dass keine Messung perfekt ist. Das heißt, dass die gemessenen Blickpunkte nicht zwangsweise *genau* dort liegen, wo die Befragten auch tatsächlich hingesehen haben. In der Regel wird dieser Messfehler durch einen der Messung vorgeschalteten Kalibrierungsprozess auf ein Minimum reduziert, sodass die Abweichungen nicht mehr als 0,7 Grad des Sehfeldes betragen.<sup>200</sup> Der Versuch, Abweichungen auf 0,7 Grad zu reduzieren, ist dadurch motiviert, dass Fixationen normalerweise einen Radius von etwa 2 Grad des Sehfeldes abdecken. Bei einer Messabweichung von 0,7 Grad kann noch immer davon ausgegangen werden, dass der gemessene Punkt innerhalb des gemeinten Fixationsradius liegt; wodurch die Abweichung für die Interpretation der Daten als irrelevant markiert werden kann. Das Problem hierbei ist jedoch, dass die Genauigkeit der Messung nicht auf der gesamten Bildfläche ident ist. Während Eye-Tracker in der Bildmitte in der Regel den Blick von Befragten sehr exakt aufzuzeichnen vermögen, nimmt die Adäquatheit der Messung zu den Rändern hin immer weiter ab. Der Wert der zulässigen Messabweichung von 0,7 Grad stellt deswegen einen Durchschnittswert dar. Das hat zur Folge, dass die gemessenen Werte in der Mitte des

---

<sup>199</sup> Vgl. auch Diekmann 2007: 240.

<sup>200</sup> Siehe Beschreibung der Eye-Tracking-Methode in Studie drei.

---

gezeigten Bildbeispiels ohne größere Probleme inhaltlich interpretierbar sind; diejenigen Werte hingegen, die in den Bildecken gemessen wurden, geben nicht unbedingt exakt die tatsächlich betrachteten Bilddetails wieder.<sup>201</sup>

Über die Interpretation der reinen Messpunkte hinaus, stellt sich aber ein noch viel größeres Problem: Eye-Tracking Daten werden normalerweise für die Auswertung und Interpretation in *Events* unterschieden. Diese Events, umfassen wie in der dritten Studie beschrieben, Fixationen und Sakkaden, also Haltepunkte und Blicksprünge. Der Eye-Tracker selbst misst diese Events aber *nicht*. Eye-Tracking Daten bestehen im Wesentlichen aus Informationen über die Position des Auges zu einem bestimmten Zeitpunkt, sogenannten Blickpunkten. Der Eye-Tracker, der zur Aufzeichnung in den Studien zwei und drei verwendet wurde, misst mit 120 Hz, was bedeutet, dass alle 8ms Informationen zur aktuellen Position des Auges gespeichert werden. Diese Informationen über die jeweilige Position des gemessenen Auges im Abstand von 8ms bilden die *Rohdaten*. Fixationen und Sakkaden – die eigentlich interessanten Größen – sind jedoch deutlich länger als diese 8ms. Fixationen haben ein ungefähres Minimum von 100-120 ms und Sakkaden haben eine ungefähre Dauer von 30-80 ms. Bevor die gemessenen Daten also überhaupt ausgewertet werden können, müssen die in den Rohdaten enthaltenen Blickpunkte zu „Events“, also zu Fixationen und Sakkaden gebündelt werden.<sup>202</sup>

Es ist wichtig zu betonen, dass diese Bündelung von Blickpunkten bereits eine Interpretationsleistung darstellt. Diese wird jedoch nicht von den Forschenden subjektiv, sondern durch einen mathematischen Algorithmus erbracht. Eine solche automatisierte Erstinterpretation der Rohdaten, die sogenannte Eventklassifikation, kann durch unterschiedliche Algorithmen und unterschiedliche Parameter erfolgen. Beides kann das Ergebnis dieser Erstinterpretation beeinflussen und dazu führen, dass dasjenige Datenset,

---

<sup>201</sup> Für eine detaillierte Besprechung diverser Vorgehensweisen beim Einsatz von Eye-Tracking als Methode siehe Holmqvist et al. 2011.

<sup>202</sup> Ebd.

---

dass schließlich von den Forschenden bearbeitet wird, sich unterscheidende Einträge aufweist.

Die wichtigste Entscheidung bei der Auswahl einer Methode zur Eventklassifikation betrifft das allgemeine Vorgehen. Es gibt im Wesentlichen zwei Möglichkeiten, auf welche zu diesem Zweck zurückgegriffen werden kann: einen geschwindigkeitsbasierten Ansatz und einen streuungsbasierten Ansatz.<sup>203</sup> Geschwindigkeitsbasierte Ansätze berechnen die Distanz zwischen zwei gemessenen Blickpunkten und schließen daraus, wie schnell das gemessene Auge sich bewegt haben muss, um diese Distanz zu überbrücken. Da Fixationen, wie aus dem Namen bereits ableitbar ist, Fixpunkte darstellen, bewegt sich das Auge während einer Fixation kaum. Sakkaden hingegen, die ja *Blicksprünge* darstellen, sind dadurch gekennzeichnet, dass das Auge eine rapide Bewegung vollzieht. Geschwindigkeitsbasierte Ansätze stellen demnach fest, ob die Geschwindigkeit, die notwendig ist, um die Distanz zwischen zwei gemessenen Blickpunkten zu überwinden, einen bestimmten Schwellenwert überschreitet. Überschreitet die Geschwindigkeit den gesetzten Schwellenwert, kann mit großer Sicherheit davon ausgegangen werden, dass die beiden gemessenen Blickpunkte Teil einer Sakkade waren; wird der Wert hingegen unterschritten, handelt es sich mit größerer Wahrscheinlichkeit um Teile einer Fixation. Streuungsbasierte Ansätze bedienen sich eines Radius-Abgleichs. Hier werden die Blickpunkte der Rohdaten dahingehend verglichen, ob sie innerhalb eines bestimmten Radius liegen, der in Regel dem Äquivalent der durchschnittlichen Fixationsgröße entspricht, also etwa zwei 2 Grad des Sehfeldes misst; bei einem Abstand von ca. 90-100 cm zum betrachteten Objekt entsprechen diese 2 Grad ungefähr 120 mm. Da Sakkaden bedingen, dass sich die Position des Auges deutlich verändert, befinden sich zwei gemessene Blickpunkte, die Teil einer Sakkade sind, in der Regel nicht innerhalb eines Radius von der Größe einer Fixation, sondern liegen deutlich weiter voneinander entfernt. Streuungsbasierte Ansätze fassen daher alle gemessenen

---

<sup>203</sup> Holmqvist et al. 2011.

---

Blickpunkte, die innerhalb eines Radius mit Fixationsgröße liegen, zu einer Fixation zusammen.

Beide Ansätze stellen valide Vorgehensweisen dar, haben aber trotzdem sowohl Vor- als auch Nachteile. Geschwindigkeitsbasierte Ansätze sind in der Regel genauer und weniger fehleranfällig als streuungsbasierte Ansätze. Das liegt allerdings nur zum Teil am Ansatz selbst, sondern hauptsächlich an den Voraussetzungen, unter denen dieser zum Einsatz kommen kann. Geschwindigkeitsbasierte Ansätze können nur auf Daten angewendet werden, die mit einem Hochgeschwindigkeitsgerät aufgezeichnet wurden. Das liegt daran, dass die zeitlichen Abstände zwischen den Messpunkten klein genug sein müssen, um für den Algorithmus noch sinnvoll als *Teil von etwas* interpretierbar zu sein. Messabstände von 8ms, wie sie das in Studien zwei und drei verwendete Gerät aufweist, sind zu groß für einen geschwindigkeitsbasierten Ansatz. Das liegt daran, dass die Interpretation von sehr kurzen Sakkaden, welche in etwa eine Dauer von 30ms aufweisen, bei Messabständen von 8ms auf der Basis von nur drei Messpunkten erfolgen müsste, wobei sowohl der erste als auch der letzte dieser drei Messpunkte bereits wieder Teil von Fixationen wären. Die gemessenen Blickpunkte sind damit erstens ihrer Anzahl nach zu wenige und zweitens nicht präzise genug, um sie bedenkenlos einem *bestimmten* Event zuzuordnen. Für Niedriggeschwindigkeitsgeräte sind geschwindigkeitsbasierte Ansätze also nicht geeignet, tatsächlich sind sie in diesem Messspektrum sogar weitaus fehleranfälliger als streuungsbasierte Ansätze. Daten die mit Niedriggeschwindigkeitsgeräten aufgezeichnet wurden, werden daher in der Regel mithilfe von streuungsbasierten Ansätzen aufbereitet.<sup>204</sup>

Fehlinterpretationen treten grundsätzlich bei beiden Methoden der Eventklassifikation auf. Es geschieht immer wieder, ganz gleich wie gut der jeweilig zugrunde gelegte Algorithmus auch sein mag, dass Blickpunkte falsch zugeordnet werden. Das liegt zu einem großen Teil

---

<sup>204</sup> Vgl. auch Holmqvist et al. 2011; Die Entscheidung für ein Niedriggeschwindigkeitsgerät ist – trotz der weniger exakten Eventklassifikationsmethoden – häufig sinnvoll, da Niedriggeschwindigkeitsgeräte keine Fixierung der Versuchsteilnehmenden erfordern (z. B. Kinnstütze) und somit eine freiere und „natürlichere“ Betrachtung erlauben (siehe auch Methodenübersicht).

---

daran, dass die Eventklassifikation nach dem Ausschlussprinzip geschieht. Geschwindigkeitsbasierte Ansätze identifizieren Sakkaden und klassifizieren alles, was nicht als Sakkade interpretiert werden kann, als Fixation. Streuungsbasierte Ansätze wählen den entgegengesetzten Weg: sie identifizieren Fixationen und klassifizieren alles als Sakkade, was gemäß ihren Einstellungen nicht als Fixation interpretiert werden kann. Aufgrund dieses Ausschlussprinzips kann es vorkommen, dass in der jeweils nicht direkt interpretierten Eventklasse fehlerhafte Zuordnungen geschehen. Können beispielsweise bei der Verwendung eines streuungsbasierten Ansatzes eine Reihe von Blickpunkten nicht eindeutig einer Fixation zugeordnet werden, werden sie automatisch als Sakkade interpretiert, und zwar unabhängig davon, ob es sich bei diesen Blickpunkten tatsächlich um eine Sakkade handelt. Das kann dazu führen, dass im Datensatz, der anschließend von Forschenden zur Interpretation des Blickverhaltens zugrunde gelegt wird, auch „Events“ enthalten sind, die eigentlich keine Events sind. Eine manuelle Überprüfung der Eventklassifikation ist daher nach Abschluss der automatischen Klassifikation zwingend erforderlich. Ziel hierbei ist es, alle Events auszuschließen, deren Ausprägung den theoretischen Definitionen von Fixationen oder Sakkaden widerspricht. Beispielsweise können im Fall von streuungsbasierten Ansätzen Blickpunktewolken als Sakkaden interpretiert werden, deren Dauer die üblichen 30-80ms deutlich überschreitet. Solche Events müssen vor der weiteren Bearbeitung der Daten als Fehler markiert werden und dürfen für die Interpretation der Daten nicht weiter berücksichtigt werden. Daraus ergibt sich, dass der Datensatz, der für die eigentliche Auswertung und Interpretation des Blickverhalten herangezogen wird, Lücken enthält. Er bildet also nicht den gesamten Betrachtungsvorgang ab, sondern lediglich Teile davon. Grundsätzlich gilt also: je fehlerhafter die Eventklassifikation, desto unvollständiger die Abbildung des Betrachtungsvorgangs.

Die Fehler der Eventklassifikation sind dabei nicht auf den Algorithmus zurückzuführen. Dieser funktioniert in seiner Natur als mathematische Funktion einwandfrei und setzt punktgenau um, was er soll. Die Fehler betreffen die Wahl von technischen Parametern,

---

welche die Bündelung von Blickpunkten beeinflussen.<sup>205</sup> Im Fall von geschwindigkeitsbasierten Ansätzen sind das die Schwellenwerte, welche die Minimalgeschwindigkeit von Sakkaden festschreiben; in streuungsbasierten Ansätzen hingegen handelt es sich um die Angaben zu Radius und Zeitfenster, auf deren Basis der Algorithmus Reihen von Blickpunkten zu Fixations-Events zusammenfasst. Veränderungen in diesen Schwellenwerten können zu einer abweichenden Eventklassifikation führen. Werden in streuungsbasierten Ansätzen also entweder Radius oder Minimal-Zeitfenster verändert, so können die „erkannten“ Fixationen andere sein als mit abweichenden Einstellungen: ob der Algorithmus Blickpunkte bereits zu einer Fixation zusammenfasst, die in einem Zeitfenster von 100 ms im Radius von 2,5 Grad auf der untersuchten Fläche streuen, oder ob eine solche Klassifikation erst ab einem Zeitfenster von 120 ms und einem Radius von 2 Grad vollzogen wird, ist ein relevanter Unterschied. Im ersten Fall würden auch sehr kurze Fixationen klassifiziert werden. Im zweiten Fall würden dieselben Reihen von Blickpunkten hingegen als Sakkaden interpretiert werden. Dabei ist jedoch keinesfalls eine dieser beiden Einstellungen als *falsch* zu bewerten. Ganz im Gegenteil, beide Einstellungen sind gemäß den bisherigen empirischen Erkenntnissen zur Natur von Fixationen zulässig.

Die Messungenauigkeiten, die Fehleranfälligkeit der Eventklassifikation und die Abhängigkeit der Klassifikationsergebnisse von zugrunde gelegten Einstellungen erschweren die Auswertung und Interpretation von Blickbewegungsdaten. Sie können daher schwerlich als objektive Abbildung von Blickverhalten verstanden werden, sondern müssen viel mehr als interpretativ überformte Näherungswerte begriffen werden. Außerdem sind Blickbewegungsauswertungen, je nach zugrunde gelegten Einstellungen deswegen nicht immer vergleichbar. Diese Probleme lassen sich nicht beheben und ein reflektierter Umgang mit ihnen ist zwingend erforderlich. So ist es zum Beispiel nötig, alle verwendeten Messungs- und Auswertungseinstellungen offenzulegen, um die spezifische Fehleranfälligkeit der Datengrundlage transparent zu machen.

---

<sup>205</sup> Vgl. auch Brinkmann/Commare 2015.



---

Die Studien zwei und drei sind diesen Erfordernissen nachgekommen. Dennoch illustriert insbesondere die dritte Studie die Persistenz der skizzierten Probleme. Die Ergebnisse dieser Untersuchung legen den Schluss nahe, dass die an einem Kunstwerk wahrgenommene Dynamik keinen Einfluss auf das Blickverhalten hat. Tatsächlich konnte jedoch in einer kürzlich durchgeführten Studie ein im Vergleich hierzu (leicht) abweichendes Ergebnis gezeigt werden.<sup>206</sup> Im Rahmen dieser Studie wurden die Blickbewegungen der Befragten mit einem Hochgeschwindigkeitsgerät aufgezeichnet. Die Messung ist damit nicht nur exakter und umfangreicher, die aufgezeichneten Daten können darüber hinaus mittels der weniger fehleranfälligen geschwindigkeitsbasierten Eventklassifikation aufbereitet werden. Die Ergebnisse der Studie zeigen, dass zwar Fixationen nicht wesentlich von der wahrgenommenen Dynamik im Bild beeinflusst werden, doch zeigt sich auch, dass die *Sakkaden* der Betrachtenden sehr wohl schneller, der Blick gewissermaßen also dynamischer wird, wenn das betrachtete Werk als sehr dynamisch wahrgenommen wird. Die dritte hier besprochene Studie ist daher ein gutes Beispiel, für die Probleme die sich aufgrund Messung, Auswertung und Interpretation ergeben können. Sie illustriert außerdem eindrücklich, warum empirische Ergebnisse, selbst physiologische Messungen, nicht pauschal als objektive Abbildung der Realität begriffen werden sollten.

### *Statistik*

Jenseits der methodenspezifischen Probleme und Einschränkungen sind allen Datenauswertungsverfahren, sofern sie Daten einer quantitativen Analyse unterziehen, die allgemeinen Probleme der Statistik gemein. Die Liste der Probleme, die hier auftreten können, ist lang. Manche dieser Probleme wurden schon im Rahmen der

---

<sup>206</sup> Die Publikation der Studie, welche von Hanna Brinkmann, Eugene McSorley, Raphael Rosenberg und Louis Williams an der University of Reading durchgeführt wurde, befindet sich in Vorbereitung. Ein Abstract zum Thema mit dem Titel *Does "action viewing" really exist? The relationship between dynamic paintings and gaze behavior* wurde im Zuge der Visual Science of Art Conference 2018 publiziert (Brinkmann et al. 2018).

---

Methodeneinführung in Kapitel 2 angesprochen. Dazu zählt beispielsweise das Kausalitätsproblem. Statistische Maßzahlen erlauben lediglich den Schluss, dass zwischen zwei untersuchten Aspekten ein Zusammenhang besteht, nicht jedoch die Richtung des Zusammenhangs. Es ist also nicht möglich, aus den Auswertungen zu schließen, ob ein Faktor den anderen bedingt oder ob es sich umgekehrt verhält. Für einen solchen Schluss ist eine inhaltliche Interpretation der numerischen Ergebnisse zwingend erforderlich. Für viele Fragen liegt die Kausalitätskette eines beobachteten Zusammenhangs auf der Hand; so ist es beispielsweise unproblematisch bei einem gefundenen Zusammenhang von Expertise und Komplexitätswahrnehmung anzunehmen, dass Expertise einen Einfluss auf die Wahrnehmung nimmt und nicht umgekehrt. Hingegen für Fragen, die den Zusammenhang zweier Wahrnehmungsinhalte untersuchen, wie beispielsweise Gefallen und Interesse, ist ein solcher Schluss nicht mehr so einfach möglich. Die Lösung eines solchen Problems liegt in der Regel aber nicht in der Adaption des statistischen, sondern in der Anpassung des methodischen Vorgehens (vgl. Kapitel 1.2). Probleme dieser Art sind jedoch keinesfalls die einzigen Auswertungs- und Interpretationsprobleme allgemeiner Natur, mit denen sich quantitativ-empirische Ansätze konfrontiert sehen. Da statistische Auswertungen selbst nicht fehleranfällig sind, sofern sie technisch-korrekt ausgeführt werden, sind die Probleme eher interpretativer Natur – allerdings in ganz anderer Hinsicht als bei der beschriebenen Kausalitätsproblematik.<sup>207</sup> Ich werde zur Veranschaulichung der hier gemeinten Problematik einige wenige allgemeine Beispiele herausgreifen, die die Aussagekraft bzw. den Erklärungsgehalt statistischer Ergebnisse betreffen und für alle drei gewählten Studien von Bedeutung sind.

Als grundlegendes Problem können wir festhalten, dass die Aussagekraft bzw. der Erklärungsgehalt von quantitativ-empirischen Ergebnissen in vielerlei Hinsicht überschätzt werden kann. Das liegt zum einen an der suggestiven Kraft der statistischen Maßzahlen, zum

---

<sup>207</sup> Fehler hinsichtlich der Modell-Daten-Passung, die eine inhaltliche Interpretation der Ergebnisse verunmöglichen, sind von der Diskussion ausgenommen. Fehler solcherart sind zwar möglich, basieren aber auf einem methodischen Unverständnis und sind daher vermeidbar.

---

anderen aber auch schlicht am konventionalen Umgang mit diesen. Statistische Ergebnisse werden in der Regel nur dann als relevant erachtet, wenn sie *statistisch signifikant* sind. Statistische Signifikanz bedeutet dabei, dass ein beobachteter Zusammenhang nicht zufällig zustande gekommen ist, sondern mit einer großen Wahrscheinlichkeit auf einer real-existenten Regelmäßigkeit basiert. Die Regel dabei ist, dass ein bestimmter Zusammenhang, sofern er überzufällig ist, eine Generalisierung der Resultate erlaubt.<sup>208</sup> In den Ergebnissen wird eine solche statistische Signifikanz mit bestimmten Maßzahlen ausgedrückt, beispielsweise dem p-Wert.<sup>209</sup> Im Prinzip sind sowohl Überzufälligkeit als auch Generalisierbarkeit natürlich wünschenswerte Merkmale für die Ergebnisse von empirischen Untersuchungen. Tatsächlich handelt es sich dabei sogar um das vornehmliche Ziel einer empirischen Studie, da hier ja die Suche nach Regelmäßigkeiten und intersubjektiv geteilten Strukturen im Zentrum des Interesses steht. So gesehen sind Ergebnisse, die sich durch statistische Signifikanz auszeichnen, durchaus als bedeutsam zu verstehen. Das Problem besteht hier auch nicht darin, dass der statistischen Signifikanz Bedeutung zugemessen wird, sondern es besteht darin *wie viel* Bedeutung ihr zugemessen wird.

Statistische Signifikanz besagt zunächst nicht mehr und nicht weniger, als dass ein Sachverhalt tatsächlich besteht (und nicht genauso gut auch Zufall gewesen sein könnte). Worüber statistische Signifikanz hingegen keine Auskunft gibt, ist, wie *bedeutsam* ein solcher Zusammenhang ist. Das liegt daran, dass es sich hierbei um eine Maßzahl handelt, die darauf abzielt, zu evaluieren, ob – und wenn ja, wie stark – unterschiedliche Merkmale in der untersuchten Datengrundlage korreliert sind. Dabei ist es für die mathematische Funktion unerheblich, ob die durchgeführte Untersuchungen inhaltlich sinnvoll ist. Studien die dieses Problem eindrücklich demonstrieren, lassen sich in den Sozialwissenschaften viele finden. Eine besonders illustrative Studie konnte beispielsweise zeigen, dass die Schuhgröße einer Person mit der Anzahl ihrer erwachsenen Tanten zusammenhängt.<sup>210</sup> Selbstverständlich

---

<sup>208</sup> Diekmann 2007: 720.

<sup>209</sup> Für eine ausführlichere Diskussion des p-wertes vgl. bspw. Diekmann 2007: 704-723.

<sup>210</sup> Vgl. Diekmann 2007: 67ff.

---

bedingt aber weder die Schuhgröße einer Person die Anzahl ihrer Tanten, noch bedingt die Anzahl der Tanten die Schuhgröße einer Person. Was allerdings sowohl Schuhgröße als auch die Anzahl der erwachsenen Tanten einer Person bedingen kann, ist das Alter der Befragten selbst: Kinder haben kleine Füße und in manchen Fällen auch junge Tanten. Je älter Kinder werden, desto größer werden ihre Füße und desto weniger wahrscheinlich ist es, dass ihre Tanten noch minderjährig sind. Trotz der Tatsache, dass der Zusammenhang zwischen Schuhgröße und der Anzahl der erwachsenen Tanten inhaltlich unsinnig ist, findet er sich in den Daten und ist darüber hinaus statistisch signifikant. Wir können daraus schließen, dass statistische Signifikanz zwar Auskunft darüber gibt, ob ein bestimmter Zusammenhang auf Regelmäßigkeiten beruht, deswegen aber noch lange keinen Schluss darüber erlaubt, ob die getestete These sinnvoll oder korrekt ist.<sup>211</sup>

Ein weiteres Problem, welches der Umgang mit Maßzahlen statistischer Signifikanz häufig nach sich zieht, ist eine Überschätzung ihrer Relevanz in anderer Hinsicht: Nur, weil ein bestimmter Zusammenhang statistisch signifikant *und* inhaltlich sinnvoll ist, ist er nicht gleichzeitig auch von tragender Bedeutung für die Beantwortung der interessierenden Frage. Da die meisten Konzepte als mehrdimensionale Bezugsstrukturen verstanden werden müssen, stellt *ein* beobachteter statistisch signifikanter Zusammenhang auch nur *ein* Merkmal dar, dass zur Erklärung des Ganzen beiträgt. Dass ein beobachteter Zusammenhang statistisch signifikant ist, besagt dabei lediglich, *dass* er für die Gesamterklärung als bedeutsam begriffen werden kann. Nicht gesagt ist damit hingegen, dass er für die Gesamterklärung des interessierenden Phänomens *eine große Rolle* spielt. Wenn sich in der ersten Studie also ein statistisch signifikanter Einfluss von Expertise auf das Komplexitätsempfinden von Kunstwerken zeigt, so ist damit nicht gleichzeitig gesagt, dass Expertise einer der *wichtigsten* Einflussfaktoren ist. Die Größe oder Bedeutsamkeit des

---

<sup>211</sup> Nuzzo 2014; Leek/Peng 2015. Anmerkung: Es gibt zahlreiche weitere Belege für beobachtete signifikante Zusammenhänge, die aus anderen als den angenommenen (und getesteten) Gründen zustande gekommen sind. Siehe hierzu z. B. Nuzzo 2014, oder auch Diekmann 2007: 18-89.

---

Einflusses lässt sich nur in der Zusammenschau unterschiedlicher statistischer Maßzahlen erfassen.<sup>212</sup>

Ein sehr ähnliches Problem findet sich auch in der Interpretation von *statistischen Modellen*. Hier werden, sofern es sich um multivariate Auswertungen handelt, mehrere Einflussfaktoren gleichzeitig untersucht. Die Ergebnisse solcher Untersuchungen zeigen auf, welche der im Modell inkludierten Faktoren die untersuchte Größe (z. B. Komplexitätserfahrung) statistisch signifikant beeinflussen und auch welche Faktoren keinen statistisch signifikanten Einfluss nehmen. Die Modelle inkludieren naturgemäß eine limitierte Anzahl von Faktoren – und schließen damit gleichzeitig eine viel größere Anzahl an Faktoren von der Untersuchung aus. Im Sinne der geisteswissenschaftlichen Kritik könnten wir solche Modelle also grundsätzlich als reduktionistisch beschreiben. Im Grunde genommen ist dies auch tatsächlich zutreffend. Statistische Modelle *sind* Reduktionen. Die, in den jeweiligen Modellen untersuchten Einflussfaktoren, stellen keine erschöpfende Untersuchung der interessierenden Phänomene dar. Es ist demnach nicht möglich, statistische Modelle basierend auf dem Vorhandensein von signifikanten Zusammenhänge als vollständiges Abbild einer mehrdimensionalen Bezugsstruktur zu verstehen.

Dass statistischen Untersuchungen reduktiv sind, ist jedoch nur problematisch, wenn die Aussagekraft dieser Modelle missverstanden wird. Die Modelle selbst sind in dieser Hinsicht völlig transparent, sodass es sich hier tatsächlich um einen „menschlichen Fehler“ im Umgang mit den Ergebnissen statistischer Berechnungen handelt. Transparent sind statistische Modelle insofern, als hier – in Regel durch eigene Kennzahlen – deutlich ausgewiesen wird, wie umfassend ein bestimmtes Modell, die in den Daten vorgefundene

---

<sup>212</sup> Welche Maßzahlen diesbezüglich berücksichtigt werden müssen, hängt zum Teil sehr stark von der jeweiligen statistischen Schätzmethode ab. Ich werde daher auf eine detaillierte Diskussion verzichten. Eine grobe Übersicht zu simpleren statistischen Modellen findet sich bei Leonhart 2009; für eine Übersicht zu komplexeren statistischen Modellen siehe Rabe-Hesketh/Skrondal 2012.

---

Varianz erklärt.<sup>213</sup> Ein gutes Modell erklärt etwa 20-30% der beobachteten Varianz.<sup>214</sup> Dass dies gleichzeitig bedeutet, dass etwa 70-80% des Phänomens *nicht* durch das Modell erklärt werden, wird in der Berechnung berücksichtigt. Dies geschieht in Form eines für die Berechnung zentralen Fehlerterms, zu welchem die Ergebnisse immer in Relation stehen. Die Problematik der Reduktion wird in quantitativen Auswertungsmethoden also nicht nur berücksichtigt, sondern stellt sogar einen integralen Bestandteil der Berechnungen dar. Fehlinterpretationen sind damit *nicht* in den Methoden selbst oder ihren Ergebnissen zu verankern.

Der letzte Punkt, den ich ansprechen möchte, beschreibt ein verwandtes Problem. Statistisch-signifikante Ergebnisse werden häufig unterhinterfragt als generalisierbar begriffen. Das Problem hierbei ist nicht, dass es ein grundlegender Fehlschluss wäre, statistisch-signifikante Ergebnisse als generalisierbar zu begreifen, sondern dass die Einschränkungen dieser Generalisierbarkeit übersehen werden. Statistische Signifikanz impliziert keine beliebige Generalisierbarkeit der Ergebnisse, sondern lediglich eine ganz bestimmte. In welcher Form Ergebnisse generalisierbar sind, hängt im Wesentlichen von der gezogenen Stichprobe, also den befragten Personen ab.<sup>215</sup> Diese sind in den meisten Studien nur für eine jeweils ganz bestimmte Gruppe von Menschen repräsentativ, nämlich jene Gruppe, mit welcher sie entscheidende Merkmale teilen. Ausgehend von der gezogenen Stichprobe, welche die Basis für die Untersuchung bildet, kann nur auf diejenige Grundgesamtheit geschlossen werden, für welche die Stichprobe ein angemessener Repräsentant ist. Wenn wir uns beispielsweise für die Gesamtheit aller Kunstlaien im westlichen Kulturraum interessieren, für die Untersuchung aus praktischen Gründen aber lediglich auf Psychologiestudierende der

---

<sup>213</sup> Je nach Modell und Schätzmethode sind das unterschiedliche Maßzahlen. Eine der gängigsten ist das R<sup>2</sup> bzw. adjustierte R<sup>2</sup>. Für eine genauere Diskussion siehe Leonhart 2009: 327-371.

<sup>214</sup> Ebd.

<sup>215</sup> Unter einer Stichprobe versteht man in der Regel die Teilmenge einer interessierenden Grundgesamtheit, welche nach bestimmten Kriterien ausgewählt wurde. Stichproben werden zumeist zufällig gezogen, um sicher zu stellen das nicht bestimmte Merkmale im Vergleich zu ihrem Vorkommen in der Grundgesamtheit in der Stichprobe überrepräsentiert sind; vgl. Diekmann 2007: 373-433.

---

Universität Wien zurückgreifen, so lässt sich aufgrund der Strichprobenziehungen keine generalisierte Aussage über Kunstläien im westlichen Kulturraum treffen. Strenggenommen dürfte aufgrund einer solchen Stichprobe nur auf die Gesamtheit der Wiener Psychologiestudierenden geschlossen werden.<sup>216</sup>

Die grundsätzliche Generalisierbarkeit von empirischen Ergebnissen ist demnach ein Mythos. Gesicherte Erkenntnisse zu robusten und stichprobenübergreifenden Zusammenhängen lassen sich tatsächlich nur mittels Reproduktionsstudien herstellen. Wie die dritte Studie zeigt, ist die schlichte Tatsache, dass eine Reproduktionsstudie zu vergleichbaren Ergebnissen kommt, jedoch noch immer kein Garant für valide Ergebnisse. Obwohl hier zwei identische Untersuchungen mit verschiedenen Stichproben (und Bildbeispielen) zum gleichen Ergebnis kommen, zeigt sich bei leicht abweichenden Mess- und Auswertungstechniken, dennoch ein in gewisser Hinsicht relativierendes Ergebnis.

### *Publikationsform*

Die meisten der in diesem Abschnitt beschriebenen Probleme lassen sich nicht wirklich lösen, weswegen die Praxis empirisch-arbeitender Forschender einen Umgang damit erfordert. Wie am Rande bereits erwähnt, ist es Usus, die Einschränkungen einer Studie in der Publikation der Ergebnisse möglichst transparent zu machen. Die Kommunikation dieser Einschränkungen erfolgt dabei gemäß gesetzter Konventionen. Dazu gehört die Beschreibung des methodischen Vorgehens wie auch die Nennung bestimmter Maßzahlen. Eine ausführliche Erläuterung der Einschränkungen, welche sich durch die verwendeten Methoden der Datenerhebung, Auswertung und Interpretation ergeben, wie es etwa im vorliegenden Text geschieht, erfolgt in der Regel nicht. Um den (schriftlichen) Äußerungen die relevanten Informationen entnehmen zu können, müssen die Leser\_innen daher zum

---

<sup>216</sup> Vgl. hierzu Henrich/Heine/Norenzayan 2010.

---

einen Vertraut mit den Kommunikationskonventionen sein und zum anderen über das nötige Methodenwissen verfügen.

Während die, teilweise in kodifizierten Randbemerkungen, bereitgestellten Informationen für Leser\_innen aus einer empirisch-arbeitenden Disziplinen kein Problem darstellen, sind diese für Leser\_innen aus den Geisteswissenschaften keineswegs selbsterklärend. Die klassische Publikationsform empirischer Arbeiten kann daher aus geisteswissenschaftlicher Perspektive einen, in methodischer Hinsicht, unreflektierten Eindruck bei der geisteswissenschaftlichen Leserschaft hinterlassen. Dieser Eindruck ist in großen Teilen aber nicht dem tatsächlich als unreflektiert beschreibbaren Umgang mit empirischen Methoden und ihren Ergebnissen anzulasten, sondern der Unübersetzbarkeit der Kommunikationskonventionen. Für Texte, die einer empirisch-arbeitenden kunstwissenschaftlichen Forschung entspringen, ist dies durchaus problematisch. Zwar ist es wichtig, methodisches Vorgehen und sich dadurch ergebende Einschränkungen offenzulegen, andererseits nähme eine Diskussion, wie sie hier vorliegt, in einem Journal-Artikel schlicht zu viel Raum ein. Darüber hinaus sind viele der hier angesprochenen Punkte in anderen Kontexten durchaus als basales Grundlagenwissen zu verstehen, was ihrer wiederholten und expliziten Kommunikation einen etwas belanglosen Beigeschmack verleihen würde.

Die diskutierten Studien haben in dieser Hinsicht zwei unterschiedliche Formate bedient. Während Studie eins und zwei sich an den Konventionen traditionell empirisch-arbeitender Disziplinen orientieren, stellt die dritte Studie eine Mischform dar, in der methodisches Vorgehen und auch einige Auswertungsaspekte detaillierter besprochen werden. Trotz der detaillierteren Methodenbeschreibung bleiben jedoch auch in der dritten Studie noch einige wichtige Aspekte hinter kodifizierten Randbemerkungen vorborgen, sodass wir uns berechtigterweise fragen dürfen, ob die Wahl einer solchen textlichen Mischform überhaupt einen Vorteil für Leser\_innen aus dem geisteswissenschaftlichen Kontext darstellt. Eine echte textliche Lösung dieses Problems ist daher (bisher) nicht auszumachen.



---

### 3.4 Zwischenfazit

Einer der vordergründigen Einwände, welcher von der negativen Kritik gegen empirische Ansätze erhoben wird, betrifft die Übersetzbarkeit geisteswissenschaftlicher Konzepte in empirisch messbare Größen. Dies sei unmöglich, so die Argumentation, weil zum einen jedweder empirische Ansatz von reduktionistischem Denken durchdrungen sei und zum anderen weil empirische Methoden nicht an-sich objektive, valide und generalisierbare Ergebnisse garantieren. Die im Verlauf dieses Kapitels diskutierten Probleme und Einschränkungen empirischer Ansätze zeigen deutlich, dass diese Kritikpunkte tatsächlich nicht einfach von der Hand zu weisen sind.

Weder Datengrundlage noch Ergebnisse von empirischen Studien können als objektive Repräsentation der Realität begriffen werden. Bei näherer Betrachtung zeigt sich vielmehr, dass empirische Daten und Ergebnisse Näherungswerte sind und gewissermaßen eine Abstraktion der Realität präsentieren. Der wesentliche Unterschied zur theoretischen Arbeit der Geisteswissenschaften besteht darin, dass es sich bei empirischen Arbeiten nicht ausschließlich um abstrakte Überlegungen handelt, sondern die abstrakten Überlegungen hier einen Praxisbezug über Näherungswerte aufweisen. Interpretieren wir die Daten und Ergebnisse empirischer Arbeiten als objektive Abbilder der Realität, trifft die negative Kritik zu. In diesem Fall könnten empirische Arbeiten tatsächlich nicht als angemessene Umsetzung geisteswissenschaftlicher Fragestellungen dienen. Interpretieren wir dieselben Daten und Ergebnisse aber in im Sinne von Näherungswerten, stellen sie, wie das vorherige Kapitel demonstriert hat, trotz der (von der negativen Kritik betonten) Probleme einen validen Diskursbeitrag dar.

---

# Diskussion

---

Ziel und Anliegen dieser Arbeit ist es einen begründeten Antwortversuch auf die Frage nach dem Wert und den Voraussetzungen empirischer Ansätze für die Kunstwissenschaften zu entwerfen. Die vorangegangenen Kapitel haben hierzu bereits aufgezeigt, dass empirische Ansätze sowohl Vor- als auch Nachteile bergen; dass die Chancen, die sie bieten, immer auch Probleme mit sich bringen; dass Erwartungshaltung und Realität nicht notwendigerweise zusammenfallen. Aus dieser Landschaft gespaltener Potenziale erwächst ein spannungsgeladenes Verhältnis im Fachdiskurs. Hier stehen sich, wie Ulrich Pfisterer despektierlich beschreibt, „mahnende Propheten“ und „lamentierende Unzufriedene“ gegenüber in einem Streit um die Angemessenheit empirischer Ansätze für den Kontext kunstwissenschaftlicher Frage- und Problemstellungen.

Die von der *positiven Kritik* hervorgehoben Chancen und Vorteile empirischer Herangehensweisen lassen sich an den vorgestellten Studien eindeutig nachvollziehen. An ihrem Beispiel zeigt sich deutlich, dass im Fach relevante Fragen existieren, deren Beantwortung zwingend eines empirischen Zugangs bedarf. Hier können sowohl theoriegebende als auch theorietestende Ansätze von der Anwendung eines empirischen Methodeninstrumentariums profitieren. Obwohl die Chancen und Vorteile empirischer Herangehensweisen aus den vorgestellten Studien hervorgehen, zeichnen sich nach einer ausführlichen Methodenreflexion doch auch erhebliche Einschränkungen ab. Diese relativierenden Einschränkungen bilden (zu einem guten Teil) den Ansatzpunkt und die Basis der *negativen Kritik*. Aus der Perspektive der *negativen Kritik* sind die Beschränkungen und Probleme empirischer Ansätze so gravierend, dass deren Einsatz als unzulässig verstanden wird. Inwiefern die Einschränkungen tatsächlich eine Absage an eine empirisch-arbeitende Kunstwissenschaft rechtfertigen, ist jedoch fraglich. Zwar zeigt eine detaillierte Methodenreflexion auf, dass eine empirische Perspektive auf kunstwissenschaftliche Fragen nicht als der Versuch, objektive Fakten zu schaffen, missverstanden werden darf, doch ist damit gleichzeitig auch deutlich, dass ein solcher Versuch ohnehin scheitern müsste. Einige Aspekte und Argumente der *negativen Kritik* richten sich also gegen vermeintliche

---

Prämissen, die in dieser Form überhaupt nicht bestehen. Es liegen hier vielmehr Missverständnisse über den angemessenen Erwartungshorizont vor.

Obwohl diese Missverständnisse den Kern der *negativen Kritik* bilden, sind den Argumentationslinien dennoch auch Punkte zu entnehmen, die sich weniger auf empirische Ansätze als Methoden beziehen, sondern die inhaltliche Relevanz ihrer Ergebnisse ganz allgemein infrage stellen. Diese Kritikpunkte sind trotz ausführlicher Diskussion und Offenlegung der Möglichkeiten und Grenzen empirischer Ansätze nicht geklärt. Den in Kapitel 2 erarbeiteten Kernaspekten der negativen Kritik folgend, lassen sich hierzu eine Reihe von Fragen an die Ergebnisse empirischer Untersuchungen formulieren. Diese werde ich nachfolgend am Beispiel der vorgestellten Studien und ihrer Methoden konkret nachvollziehen. Im Sinne der zentralen Kritiken – nämlich der Reduktionismus und die Theoriearmut empirischer Methoden sowie die Ahistorizität ihrer Ergebnisse – lassen sich die folgenden Fragen an die diskutierten Studien stellen: Welche inhaltlichen Fragen sind in den jeweiligen Untersuchungen offengeblieben? Sind die Fragen möglicherweise wegen eines unzulässigen Reduktionismus offengeblieben? In welche Weise sind die Ergebnisse in kunstwissenschaftliche Theorien eingebettet und welche Konsequenzen ergeben sich für die Theoriebildung? Sind die Ergebnisse tatsächlich ahistorisch bzw. inwiefern ist ihre Ahistorizität für den vorliegenden Fragenkontext als problematisch zu bewerten? Diese Fragen abschließend soll am Ende dieser Diskussion auch die Publikationsform in ihrer Funktion als realisierter Diskursbeitrag betrachtet werden. Die zentrale Frage, die ich dazu stellen möchte, ist: bezieht sich die negative Kritik tatsächlich auf die jeweiligen Studien selbst – oder ergeben sich die kritisierten Einschränkungen aufgrund der spezifischen Publikationsform?

---

## *Reduktionismus*

Der vordergründigste Kritikpunkt, der empirischen Methoden gegenüber geäußert wird, ist sicherlich der vermeintliche Reduktionismus empirischer Perspektiven. Um diese Bedenken zu adressieren, werde ich im Folgenden an einem Beispiel diskutieren, ob – und wenn ja, welche – Fragen oder Probleme durch die Untersuchung nicht bearbeitet wurden und inwiefern dieses Versäumnis der reduktionistischen Arbeitsweise der verwendeten empirischen Ansätze anzulasten ist. Da die erste der vorgestellten Studien am umfangreichsten ist und darüber hinaus das einzige Beispiel darstellt, dass als theoriegebender Ansatz verstanden werden kann, lege ich die Studie *More Than the Sum of Its Parts: Perceiving Complexity in Painting* für die weitere Diskussion zugrunde.

Offengebliebene Fragen sind in empirischen Untersuchungen in der Regel auf die Konzeptualisierung des Untersuchungsgegenstandes zurückzuführen. Gemeint ist, dass die Operationalisierung, d.h. das Messbar-Machen der untersuchten Konzepte, den Ein- und Ausschluss bestimmter Aspekte des Konzepts erfordert. Das hat zur Folge, dass nur das Zusammenwirken der in der Studie *inkludierten* Aspekte Gegenstand der Betrachtung ist und dass die gewählten Konzeptualisierungen einen unterschiedlichen Grad an Offen- oder Geschlossenheit aufweisen. Wenn wir also unser Augenmerk im Folgenden auf offengebliebenen Fragen richten, dann ist damit nicht jede denkbare (Detail-)Frage gemeint, die *auch* an den Problemraum *Komplexitätserfahrung in der Malerei* gestellt werden könnte. Stattdessen ist damit die Frage gemeint, inwiefern jene Konzeptualisierung, welche für die Untersuchung zugrunde gelegt wurde, *in-sich* das Phänomen unzulässig reduziert und welche weiteren Fragen somit unbeantwortet geblieben sein könnten.

Welche Fragen sind also offengeblieben? Grundsätzlich ließen sich natürlich viele (Detail-)Fragen stellen. Im konkreten Fall der ersten Studie zeigt sich allerdings schnell, dass einige der hier denkbaren und noch unbeantworteten Detailfragen für den Kontext kunstwissenschaftlicher Überlegungen wenig interessant sind. Da der Problemraum

---

grundsätzlich interdisziplinär ist, da er an der Grenze von Kognitionspsychologie und Kunstwissenschaft angesiedelt ist, fallen viele Detailfragen in den „Kompetenzbereich“ der Psychologie. Dazu zählen etwa Fragen, die sich zu spezifischen Aspekten im Ablauf kognitiver Prozesse stellen ließen. Diese Fragen werde ich für die nachfolgende Diskussion nicht berücksichtigen. Stattdessen werde ich im Folgenden exemplarisch eine Frage herausgreifen, die sich aus dem Verwendungskontext des Begriffes *Komplexität* in den Kunstwissenschaften ergibt.

Wie dem kurzen Überblick zur Begriffsverwendung in der kunstwissenschaftlichen Literatur in Kapitel 2.1.1 zu entnehmen ist, wird der Terminus *Komplexität* in der Regel im Zusammenhang von Vergleichen verwendet. Wenn also beispielsweise Hofmann befindet, dass das künstlerische Profil der Brücke "komplexer als das der Fauves"<sup>217</sup> sei, so impliziert dies eigentlich schon, dass *Komplexität als relationaler Ausdruck* zu verstehen ist. Anders gesagt: Komplexität ist keine immanente Eigenschaft von Kunstwerken. Das leitet sich nicht nur deutlich aus der praktischen Verwendung des Begriffs ab, sondern kann auch als Wesen des Komplexitätskonzepts erachtet werden. Damit ist gemeint, dass das Prädikat *komplex* immer schon voraussetzt, dass es weniger komplexe Objekte, Sachverhalte, Strukturen, etc. gibt. Komplexität wäre demnach per definitionem eine relationale Kategorie. Stellen wir also die Frage nach der Komplexität eines bestimmten Objektes, müssen wir notwendig die implizite Frage mitdenken: „Komplex, im Vergleich wozu?“

Das Definitionsangebot, welches in der ersten dargestellten Studie vorgebracht wird, berücksichtigt diesen relationalen Charakter der Komplexität(-serfahrung) nur bedingt. Zwar werden hier diverse Abstufungen bestimmter formaler und semantischer Aspekte in Beziehung zur Komplexitätserfahrung gesetzt, weswegen natürlich grundsätzlich eine Relationalität gegeben ist. Doch handelt es sich hier um abstrakte Kategorien, die unabhängig von einem spezifischen Bewertungskontext diskutiert werden. Dass aber der spezifische

---

<sup>217</sup> Hofmann 1987: 266.

---

Bewertungskontext sogar konstitutiv für die Erfahrung von Komplexität sein mag, für die Frage *Komplex, im Vergleich wozu?*, ist – aus kunstwissenschaftlicher Perspektive – schwerlich von der Hand zu weisen (vgl. *Halo-Effekt*, Kapitel 3.2). Betrachten wir beispielsweise *Die Versuchung des Hl. Antonius* von Max Ernst und vergleichen sein Werk mit Robert Motherwells nahezu zeitgleich entstandenem *Western Air*, so würden wohl die meisten Max Ernsts Werk als das Komplexere beschreiben (siehe Abbildung 1 und 2, Kapitel 2.1). Nun ließe sich dagegen einwenden, dass beide Werke nur wenig verbindet und ein Vergleich somit jeder theoretischen Grundlage oder Motivation entbehrt. Würde man dagegen zwei Werke vergleichen, die sich eindeutig zu einander in Beziehung setzen lassen, wie würde ein intuitives Komplexitätsurteil ausfallen? Beispielsweise ließen sich hierfür etwa die beiden Verkündigungsszenen von Fra Angelico und Filippo Lippi anführen (siehe Abbildung 20). Hier würden wohl die meisten zustimmen, wenn wir Lippis Verkündigungsszene als die Komplexere beschreiben.



Abbildung 20 | links: Fra Angelico (1445). *Verkündigung* (230 x 321 cm), Foto: Artstor, © Museo di San Marco, Florenz; rechts: Lippi, Fra Filippo Tommaso (1450). *Verkündigung* (127 x 116 cm), Foto: Artstor, © Alte Pinakothek, München.

---

Mit Blick auf die beiden ausgewählten Bildbeispielpaare entsteht leicht der Eindruck, dass der intuitive Komplexitätseindruck hauptsächlich auf Unterschiede im formalen Bildaufbau zurückzuführen ist. Tatsächlich ist es besonders auffällig, dass im komplexer erscheinenden Beispiel die Anzahl der dargestellten Elemente deutlich höher ist. Zudem sind Bildelemente weniger deutlich voneinander abgegrenzt. Durch die Kombination dieser beiden Merkmale entsteht der Eindruck eines wenig übersichtlichen Bildraumes. Die wesentlichen Charakteristika komplexer Strukturen sind hier also deutlicher auf der formalen Ebene ausgeprägt als in dem gegenübergestellten Beispiel: Eine nicht exakt bestimmbare Anzahl an einzelnen Zeichen steht hier gleichzeitig in unauflöslich-wechselseitiger Beziehung zueinander.

Der Eindruck, dass die formale Struktur eines Bildes der absolute Primat der Komplexitätswahrnehmung sei, ist jedoch ein durch Auswahl und Zusammenstellung der Bildbeispielpaare induzierter Trugschluss. Die Gegenüberstellung von Bildbeispielen führt dazu, dass jene Bildaspekte, welche zwischen den jeweiligen Werken eindeutig unterschiedlich ausgeprägt sind, überbetont werden. Es ist also durch die Gegenüberstellung von Fra Angelicos und Lippis jeweiligen Verkündigungsszenen ganz *besonders* deutlich, *wie* übersichtlich der Bildraum in Fra Angelicos Werk gestaltet ist. Hätten wir beide Arbeiten ohne diese Gegenüberstellung jeweils unabhängig voneinander betrachtet, wären wir vielleicht nicht zu dem Schluss gekommen, Lippis *Verkündigung* wäre unübersichtlich(er) – tatsächlich zeichnet sich das Werk ja auch durch einen sehr klaren geometrischen Aufbau aus. In einem anderen Vergleich wäre das Formale womöglich weder so prominent erschienen noch notwendigerweise der Primat der Komplexitätswahrnehmung gewesen. Selbverständlich wäre es ebenso denkbar, dass man einen Eindruck von der Komplexität eines Bildes hauptsächlich wegen seiner semantischen Merkmale hat.

Diese sich durch den Bewertungskontext vermittelte Relationalität der Komplexitätserfahrung wird in der vorgestellten Studie nicht thematisiert. Tatsächlich verhält es sich sogar so, dass diese Art der Relationalität bewusst von der Untersuchung



---

ausgeklammert wurde. Im Text aller dreier Studien findet sich am Rande der Hinweis, die Bilder wären im Versuchsaufbau in zufälliger Reihenfolge gezeigt worden. Diese Aussage stellt dabei keineswegs eine methodische Formalität dar. Vielmehr versteckt sich dahinter eine konzeptionelle Entscheidung – nämlich die kontextuell-gebundene Relationalität von Erfahrungen von der Untersuchung auszuschließen. Durch die Randomisierung der gezeigten Bildbeispiele ist sichergestellt, dass die am Versuch teilnehmenden Personen dieselben Kunstwerke in unterschiedlichen Reihenfolgen sehen. Da die finalen Berechnungen immer aus durchschnittlichen Näherungswerten bestehen, ist somit sichergestellt, dass es keine *systematische* Verzerrung aufgrund dieser speziellen Relationalität in den Ergebnissen gibt. Gleichzeitig ist damit allerdings die Betrachtung der kontextuell-gebundenen Relationalität der einzelnen Erfahrung ausgeschlossen: Ob – und wenn ja, in welcher Form – die tatsächliche Erfahrung durch ihren Kontext geprägt ist, lässt sich aufgrund dieses methodischen Vorgehens nicht mehr sagen.

Diese Operationalisierung der Komplexitätserfahrung lässt sich durchaus als Reduktionismus auffassen. Dies gilt umso mehr, als es weder in methodischer noch mathematischer Hinsicht eine Notwendigkeit darstellt, die Untersuchung von kontextuell-gebundener Relationalität auszuklammern. Mit einer anderen Auswertungsentscheidung, oder auch einem anderen Studiendesign, wäre es durchaus möglich gewesen, den Einfluss von davor gesehenen Bildbeispielen auf ein aktuelles Bild zu erfassen. Das Versäumnis ist demnach nicht einem grundsätzlichen Reduktionismus empirischer Ansätze zuzuschreiben, sondern lediglich einem in den gewählten Studien zum Tragen kommenden. Allerdings muss man festhalten: obwohl es grundsätzlich möglich ist, weniger reduktionistisch vorzugehen, ist die gewählte Operationalisierung doch sowohl in methodischer als auch mathematischer Hinsicht als konsensfähiger Usus zu sehen.<sup>218</sup> Sofern die kontextuell-gebundene Relationalität von Erfahrungen nicht explizit Gegenstand einer Untersuchung ist, wird sie in der Regel durch die beschriebenen Vorgehensweisen von der Betrachtung ausgeschlossen.

---

<sup>218</sup> Vgl. auch Diekmann 2007: 465.

---

Die sich zwingend stellende Anschlussfrage lautet nun natürlich: Inwiefern ist der Reduktionismus der Studie für die Integration der Ergebnisse in den geisteswissenschaftlichen Diskurs problematisch? Grundsätzlich ist hier zunächst anzumerken, dass die vorgestellten Untersuchungen alle mit einem Setting arbeiten, in dem Gegenüberstellungen von Kunstwerken insofern stattfinden, als jeweils Reihen von Werken nacheinander betrachtet werden. Die sich durch diese Gegenüberstellungen ergebende Relationalität einzelner Erfahrungen ist weiter zwar nicht Gegenstand der Untersuchung, ihrer Relevanz wird aber dennoch Rechnung getragen. Durch die Randomisierung der gezeigten Bildbeispiele und die sich am Durchschnitt orientierenden Berechnungen wird nicht der konkrete Kontext der Bewertung ins Zentrum des Interesses gerückt, sondern der Allgemeine. Damit ist gemeint, dass die Ergebnisse ausdrücklich implizieren, dass sie im Vergleich zu einer *Bandbreite* an konkreten Kontexten zu verstehen sind. Anders gesagt: Der in dieserart Untersuchung Berücksichtigung findende Bewertungskontext, stellt sich durch einen ganzen Korpus unterschiedlicher Bildbeispiele dar. Damit ist eher ein Bewertungshorizont als ein spezifischer Bewertungsrahmen gemeint. So gesehen ist der vermeintliche Reduktionismus der vorgestellten Studien möglicherweise überhaupt nicht als Reduktionismus im eigentlichen Sinne zu verstehen, sondern nur als bestimmte Perspektive auf das gesetzte Problem.

Dies wird insbesondere an der ersten Studie deutlich. Hier ist, obschon nicht ausdrücklich formuliert, die implizite Folgefrage *Komplex, im Vergleich wozu?* insofern Teil der Untersuchung, als der Vergleichswert nicht durch einen einzelnen Widerpart gestellt wird, sondern durch eine Bandbreite an Kunstwerken. Eine solche, sich eher an übergeordneten Kategorien orientierende Perspektive auf das gestellte Problem ist, was das die Erfassung des Phänomens anbelangt, wünschenswert. Eine Beschränkung der Untersuchung auf die kontextuell-gebundene Relationalität von Erfahrungen, die zulasten der allgemeinen Relationalität in Bezug auf einen größeren Horizont geht, wäre hingegen problematisch. Problematisch deswegen, weil selbst im gesetzten Fall, dass zwei Werke in direktem Bezug stehen und somit die Erfahrung des jeweils anderen wechselseitig beeinflussen, doch immer

---

noch ein allgemeiner Bewertungshorizont präsent ist, vor welchem die jeweils in Beziehung stehenden Werke erfahren und evaluiert werden. Für den Fall, dass ein bestimmtes Werk nicht in direktem Bezug zu einem anderen erfahren wird, ist die notwendige Präsenz eines allgemeineren Horizontes noch deutlich nachzuvollziehen. Am Rande ist hierzu noch zu bemerken, dass es durchaus denkbar wäre beide Relationalitäts-Aspekte abzudecken: es handelt sich hierbei, wie bereits erwähnt, nicht um eine Entweder-oder-Frage. Diesen Punkt abschließend lässt sich jedoch festhalten, dass trotz der prinzipiellen Möglichkeit weitere Relationalitäts-Aspekte der Komplexitätserfahrung zu integrieren, der durch die methodischen und mathematischen Entscheidungen bedingte Reduktionismus der vorgestellten Studie nicht zu einem grundlegenden Problem führt. Ganz im Gegenteil, die dadurch realisierte Konzeptualisierung deckt die im Wesentlichen wichtigen Aspekte hinreichend ab.

### *Das Verhältnis von Messung und Theorie*

Neben dem Vorwurf eines grundsätzlichen Reduktionismus steht – mit diesem eng verknüpft – auch die Frage nach dem Verhältnis von Theorie und Messung empirischer Arbeiten im Fokus der Kritik. Wenn Geimer also anmerkt, dass empirische Ansätze im Wesentlichen auf der „Suche nach der >biologischen Verankerung< der Kunst“ seien und sich hier die Annahme ausdrücken würde, dass „Verfahren der Naturwissenschaften (wie Messung und Experiment) per se solides und generalisierbares Wissen erzeugen“, so impliziert dies den Vorwurf einer Gleichsetzung von Bedeutung mit Referentialität.<sup>219</sup> Die Ergebnisse

---

<sup>219</sup> Vollständiges Zitat Peter Geimer: „Die Suche nach der >biologischen Verankerung< der Kunst beruht auf dem Vorurteil, dass Verfahren der Naturwissenschaften (wie Messung und Experiment) per se solides und generalisierbares Wissen erzeugen und auch für die Forschungsgegenstände der Kunstgeschichte zuständig sind – eine Vorstellung, die durch wissenschaftliche Arbeiten längst relativiert ist. Sie erinnert aber auch an eine Praxis, die Roland Barthes bereits von Jahrzehnten zu recht als >Mythologisierung< kritisiert hat: die Verwandlung kulturell und historisch verfasster Phänomene in vermeintliche >Natur<, um ihnen den Anschein unhinterfragbarer Evidenz zu geben.“ (Geimer 2015: 340.); siehe auch Seite Kapitel 2.4.

---

(naturwissenschaftlicher) Messungen, so der Kritikpunkt, würden in empirischen Arbeiten *an-sich* Bedeutung tragen; der durch sie ausgedrückten Referentialität würde nicht angemessen Rechnung getragen. In empirischen Ansätzen, so die Annahme, werden Denken und beobachtbares Verhalten eng geführt.

Wie bereits in Kapitel 2.5 ausführlich diskutiert, handelt sich hierbei schlicht um ein Missverständnis vielen empirischen Studien gegenüber. Empirische Ansätze basieren fundamental auf einer theoretischen Einbettung, da sie in der Regel die im Kontext theoretischer Auseinandersetzungen entwickelten Thesen systematisch auf ihren Praxisgehalt überprüfen. Außerdem erfolgt auch die Interpretation von (numerischen) Ergebnissen empirischer Messungen immer in Bezug zu Theorie. Dies gilt auch für die vorgestellten Studien: theoretische Überlegungen waren stets der Ausgangspunkt. Für die erste Studie gilt dies umso mehr, als das Ziel hier sogar war, das Komplexitätskonzept in der Malerei (neu) zu definieren. Das Ziel selbst war theoretisch, auch wenn das methodische Vorgehen empirisch war. Trotz dieser notwendigen und zentralen Präsenz theoretischer Überlegungen bleiben die vorgestellten Studien in zweierlei Hinsicht kritisierbar: zum einen kann durchaus hinterfragt werden, ob die gewählte theoretische Rahmung als hinreichend verstanden darf; zum anderen lässt sich der Kritikpunkt auch stärker im Bezug zur Frage nach der Abbildung der Referentialität von beobachtetem Verhalten oder Strukturen lesen. Die Frage, ob – und wenn ja, inwiefern – diesem Verhältnis Rechnung getragen wurde, kann grundsätzlich und unabhängig von dem soeben skizzierten Missverständnis an empirische Arbeiten gestellt werden.

Zur Frage nach der Angemessenheit der theoretischen Rahmung lässt sich nur bedingt eine positive Antwort geben: einerseits ist es richtig, dass, wie soeben angemerkt, die getesteten Größen in allen drei Studien aus Thesen hervorgegangen sind, die entweder in der Literatur direkt geäußert wurden oder sich aus dieser ableiten lassen; andererseits ist die Diskussion stark auf empirisch-beobachtbare Merkmale ausgerichtet. Im Text der Studien fehlt eine Diskussion von Alternativen, die jeweiligen Thesen zu deuten und auch eine umfassende

---

Begründung, warum die schlussendlich gewählte Konzeptualisierung zugrunde gelegt wurde. Außerdem findet auch kein Rückbezug der Ergebnisse an den geisteswissenschaftlichen Theoriediskurs statt, welcher über die reine Interpretation der Ergebnisse hinausgeht. Gewissermaßen lässt sich hier also tatsächlich von einer Theoriearmut sprechen. Aufgrund der fehlenden Diskussion alternativer theoretischer Deutungen kann zudem der Vorwurf eines in theoretischer Hinsicht reduktionistischen Vorgehens nicht ohne weiteres zurückgewiesen werden.

Zur Frage nach der Abbildung der Referentialität von beobachtetem Verhalten lässt sich eine deutlich positivere Antwort formulieren. In keiner der Studien wird beobachtetes Verhalten absolut gesetzt; im Gegenteil, die gemessenen Einheiten, seien sie nun physiologischer Natur (Blickbewegungen) oder Berichte über bewusste Erfahrung (Fragebögen), werden ganz explizit als *Indikatoren* für bestimmte Prozesse verstanden, welche selbst nicht beobachtbar sind. So werden etwa gemessene Blickbewegungen nicht als Abbildung von „Sehen“ verstanden, sondern lediglich als Indikatoren von Aufmerksamkeit. Ähnliches gilt für die Angaben, die die Versuchsteilnehmenden zu ihren jeweiligen Bilderfahrungen gemacht haben. Die Beobachtung beispielsweise, dass Kunstexpert\_innen dazu neigen, Kunstwerke als weniger komplex einzustufen als Laien, wird nicht im Sinne einer unhinterfragten Expert\_innen-Kompetenz oder -Eigenschaft gefasst, sondern als Ausdruck anders gelagerter kognitiver Strukturen. Während man den vorgestellten Studien also durchaus Theoriearmut vorwerfen darf, insbesondere was eine ausführliche Diskussion der Einbettung der Thesen und Ergebnisse angeht, trifft der Vorwurf einer Gleichsetzung von Bedeutung mit Referentialität nicht zu.

---

### *Ahistorizität*

Es mag zunächst nicht offensichtlich sein, warum die vermeintliche Ahistorizität von Ergebnissen empirischer Untersuchungen für die vorliegende Frage überhaupt als relevant verstanden werden sollte. Im Falle der ersten Studie beispielsweise ging es doch hauptsächlich darum, den Begriff Komplexität für den Kontext der Kunst zu definieren. Das Ziel war es also die Verwendung des Begriffs in Form eines Theoriewerkzeuges zu hinterfragen. Damit ist das Ziel vordergründlich ein methodisches, welches darauf abzielt zeitgenössischen Forscher\_innen ein Instrument zur Verfügung zu stellen, dass zu einer reflektierten Haltung beiträgt und sowohl Wahrnehmung als auch Beschreibungen von Kunst umfangreich(er) erklärt. Verstehen wir die Fragestellung jedoch nicht ausschließlich als methodisch motiviert, sondern auch mit inhaltlichem Bezug, so stellt sich die Frage nach der Reichweite der Ergebnisse. Lassen sich die über Komplexitätswahrnehmung formulierten Erkenntnisse allgemeiner fassen? Ist eine Übertragung des Konzeptes auch auf historische Kontexte möglich? Dürfen wir annehmen, dass die Komplexitätserfahrung bei der Auseinandersetzung mit Kunst bei zeitgenössischen und historischen Betrachtenden auf denselben Parametern beruht? Stellen wir diese Fragen, so wird das Problem der Ahistorizität empirischer Studien evident.

Letztlich lässt sich für dieses Problem keine definitive Lösung finden. Es ist schlicht unmöglich historische Betrachtende zu ihren Komplexitätseindrücken zu befragen. Historische Vergleichsdaten gibt es ebenfalls nicht. Gesicherte Rückschlüsse auf historische Kontexte lassen sich aus den Ergebnissen der Studien also nicht ziehen. Die Ahistorizität der Ergebnisse ist dennoch nicht in dem Umfang problematisch, wie es zunächst scheinen mag. Zum einen dürfen wir durchaus annehmen, dass es gewisse *ethnologische und anthropologische Konstanten* gibt, die über Zeit *wenig Veränderung* erfahren. Basale Funktionen und Determinanten des kognitiven Systems können darüber hinaus sogar als

---

*natürliche Konstanten* begriffen werden, welche über die Zeit *keine Veränderung* erfahren.<sup>220</sup>

Wie bestimmte kognitive Prozesse ablaufen, kann also teilweise als über die Zeit unveränderlich begriffen werden. Auch diejenigen Faktoren, welche diese Prozesse beeinflussen, dürfen als transhistorisch begriffen werden. So ist es beispielsweise plausibel, dass Aspekte formaler Bildkomplexität, wie etwa Unübersichtlichkeit, immer das Komplexitätsempfinden von Betrachtenden beeinflusst haben. Auch Kategorien semantischer Komplexität, wie etwa die Wahrnehmung semantischer Tiefe, wird wohl zu jeder Zeit den Eindruck von Mehrdimensionalität nach sich ziehen und dazu führen, dass eine größere Anzahl an Inhalten (oder Details) in der Auseinandersetzung mit einem Werk verhandelt werden. Ähnliches gilt für die Ergebnisse von Studie zwei und drei: es ist beispielsweise nicht unplausibel, dass das Vorhandensein von vielen gleichzeitig dargestellten Bildelementen, zu jeder Zeit die Konsequenz hatte, dass der Blick von Betrachtenden weniger gezielt auf nur bestimmte Einzelemente gerichtet wird.

Trotz der hohen Plausibilität, dass hier tatsächlich von transhistorischen Konstanten ausgegangen werden kann, ist es nicht unproblematisch die Ergebnisse ohne weitere Einschränkung zu generalisieren. Dies ist deswegen nicht ohne weiteres möglich, weil – im Bezug auf die erste Studie – das Verständnis der die Komplexität determinierenden Faktoren variabel sein kann. Zum einen hat sich bereits anhand der Untersuchung zeitgenössischer Betrachtender gezeigt, dass nicht jede und jeder bestimmte Bildbeispiele als im gleichen Umfang unübersichtlich empfindet. Auch die Erfahrung semantischer Tiefe lässt sich den Ergebnissen der ersten Studie zu Folge nicht als intersubjektiv einheitlich verstehen. Zu sehr hängt eine Erfahrung von Vorerfahrungen ab. Diese Vorerfahrungen bilden den Vergleichsrahmen, den Horizont, vor welchem jede neue Erfahrung stattfindet. War eine Person in ihrer Vergangenheit also großteils mit verhältnismäßig übersichtlich-gestaltetem

---

<sup>220</sup> Ein vergleichbares Argument findet sich auch in den jüngsten Arbeiten Klaus Speidels. Diese sind noch nicht publiziert, aber in Form von mündlichen Vorträgen bereits in den bestehenden Fachdiskurs eingebracht worden. Ein publiziertes Abstract zum Thema findet sich Online auf der Website der *Art & Science Conference on Empirical Methods in Art History and Visual Studies* 2017 (siehe Speidel 2017).

---

Bildmaterial konfrontiert, so wird ihre „Toleranzschwelle“ für Unübersichtlichkeit eventuell niedriger liegen als im Falle einer Person, die sich bereits häufig mit unübersichtlichem Bildmaterial auseinandergesetzt hat. Was hier zunächst für zeitgenössische Betrachtende gilt, muss natürlich auch in gleichem Maße für historische Betrachtende angenommen werden. Es ist darüber hinaus anzunehmen, dass der grundlegende Ausgangswert für die Faktoren, welche die Komplexitätserfahrung bedingen, in historischen Kontexten ein jeweils anderer ist, als wir ihn etwa für zeitgenössische Betrachtende anlegen würden. Damit ist gemeint, dass es denkbar ist, dass Faktoren wie Übersichtlichkeit oder auch semantische Tiefe zu unterschiedlichen Zeitpunkten der Geschichte etwas anderes bedeutet haben.

Für die Frage nach der Übertragbarkeit der Ergebnisse auf historische Kontexte lässt sich also festhalten, dass es zumindest sehr wahrscheinlich ist, dass die in den Studien identifizierten Faktoren, welche Einfluss auf die Komplexitätserfahrung und deren Konsequenzen für die Bildbetrachtung nehmen, als transhistorisch zu verstehen sind.<sup>221</sup> Dabei sind allerdings nur die abstrakten Kategorien tatsächlich übertragbar, die konkreten Bezugswerte hingegen nicht. Das hat zur Folge, dass zwar der Begriff als solcher übertragbar ist, nicht aber die auf Bewertungen zeitgenössischer Betrachtender beruhenden durchschnittlichen Komplexitätswerte, die im Anhang der ersten Studie für 90 Kunstwerke gelistet sind.

### *Publikationsform*

Die vorherigen Abschnitte haben aufgezeigt, dass die vorgestellten Studien bei genauerer Betrachtung der negativen Kritik weitestgehend standhalten. Das jeweilige Vorgehen in den Studien ist nur bedingt reduktionistisch oder ahistorisch. Außerdem ist die Anwendung einer auf Strukturen bedachten Perspektive auf die Fragestellung nicht nur hilfreich, sondern

---

<sup>221</sup> Vgl. auch Speidel 2017.



---

zwingend erforderlich, um relevante Aspekte zu adressieren und sinnvolle Antworten auf ungeklärte Fragen der Rezeptionsästhetik geben zu können.

Dennoch sind nicht alle Aspekte, welche von der negativen Kritik an empirischen Ansätzen unterstrichen wurden, für die vorgestellten Studien von geringer Bedeutung. Insbesondere das Verhältnis von Theorie und Messung kann als problematisch begriffen werden. Einige als relevant zu bewertende Aspekte sind durch die vorgestellten Studien nicht hinreichend abgedeckt. Dazu gehören etwa eine Diskussion alternativer Interpretationen, welche die letztendlich getroffene Wahl kontextualisiert und begründet oder auch ein expliziter Rückbezug der Ergebnisse auf den kunstwissenschaftlichen Diskurs. In dieser Hinsicht sind die Studien zwei und drei schwächer als die erste Studie. Diese Kritikpunkte beziehen sich jedoch bei genauerer Betrachtung nicht auf die Studien an sich, sondern eindeutig auf den sie rahmenden Text.

Die Texte zu den drei Studien orientieren sich hinsichtlich Form, Struktur und Argumentationsstrategie stark an den Konventionen traditionell empirisch-arbeitender Disziplinen, wie der Psychologie. Tatsächlich sind zwei der Studien aufgrund ihres interdisziplinären Charakters und der empirischen Vorgehensweise in Zeitschriften publiziert worden, die der psychologischen Ästhetik zugeordnet werden müssen. In diesem Kontext ist es üblich Publikationen möglichst knapp zu halten. Eine kurze Einführung, die präzise auf die später empirisch untersuchten Größen zugespitzt ist, und Informationen über alternative Interpretationen oder gröbere Zusammenhänge/Einbettungen ausklammert, ist hier explizit erwünscht. Auch die Diskussion ist in diesem Publikationsformat weitestgehend entlang der in der knappen Einleitung aufgezeigten Thesen zu orientieren. Es handelt sich hierbei also gewissermaßen mehr um einen Ergebnisbericht, denn um einen eigenständigen theoretisch-inhaltlichen Beitrag.

Die Publikationsform traditionell empirisch-arbeitender Disziplinen stellt damit einen deutlichen Bruch mit den Erwartungen dar, welche an geisteswissenschaftliche Texte gestellt

---

werden. Während die Publikationen traditionell-empirischer Arbeiten in der Regel kleine Beiträge zu größeren Diskursen sind, die durch das Feld fortlaufend aktualisiert werden, gleichen geisteswissenschaftliche Publikationen häufig Monumentalwerken, die eine in sich abgeschlossene Bearbeitung eines bestimmten Themas darstellen. Das für den Fall geisteswissenschaftlicher Publikationsformen eine breitere Kontextualisierung und eine deutlichere Rückbindung nicht nur zu den gesetzten Thesen, sondern auch zum größeren Theoriekontext angestrebt wird, ist kaum verwunderlich, immerhin handelt es sich zumeist um Arbeiten, die nicht einfach nur zum Ziel haben, einen Beitrag zur Theoriebildung zu leisten. Vielmehr soll eine bestimmte Nische möglichst erschöpfend behandelt sein.

Empirische Untersuchungen mit interdisziplinärem Charakter, die im Kontext der Kunstwissenschaften angestellt werden, müssten in ihrem Publikationsformat die Konventionen beider Forschungstraditionen bedienen. Nur so ließe sich der Kritik hinsichtlich des Verhältnisses von empirischer Messung und theoretischer Diskussion bzw. Einbettung entgegenwirken. Dies scheint jedoch nicht ohne weiteres möglich zu sein. Zum einen sind viele geisteswissenschaftliche Zeitschriften, die einer solchen Mischform Raum geben könnten, trotz der hochschulpolitischen Forderung nach mehr Interdisziplinarität noch immer zu stark auf klassische Themen und Methoden ausgerichtet, als dass sie empirische Untersuchungen überhaupt publizieren würden; zum anderen eignet sich ein großer Anteil von empirischen Untersuchungen nur bedingt zur Publikation in Form von Monographien, wo ein solches Problem nicht gegeben wäre. Die hier verhandelten Themen sind oft nicht hinreichend umfassend, da sie, wie gesagt, nur einen kleinen Beitrag zu einem größeren Diskurs darstellen. Eine Publikation in monographischer Form scheint daher kaum gerechtfertigt.

Das Problem ließe sich jedoch auch von einer anderen Seite betrachten. Vielleicht sollten wir nicht die Anpassung von Publikationsformen und –formaten empirischer Untersuchungen an bestehende Konventionen fordern, sondern die sich durch den Einsatz neuer methodischer Herangehensweisen eröffnenden Möglichkeit und Perspektiven sowie ihre Konsequenzen

---

ernstnehmen. Das hätte zur Folge, dass an empirische Arbeiten nicht länger der Anspruch gestellt würde, einer klassisch-geisteswissenschaftlichen Arbeit in Umfang und Reichweite zu entsprechen. Empirische Arbeiten könnten in der Folge als Diskursbeiträge bewertet und verstanden werden, was sich auch in einem für diesen Zweck geeignetem Publikationsformat niederschlagen dürfte. Gleichzeitig würde eine solche, eher an den traditionell-empirische Wissenschaften angelehnte Publikationsstrategie unterstreichen, dass empirische Arbeiten eben *nicht* mit klassischen Ansätzen und Arbeiten gleichzusetzen sind. Sie stehen nicht für sich und bedürfen notwendig eines vorangegangenen Diskurses, an den sie angegliedert werden.

---

### *Schlusswort*

Die im Verlaufe der Arbeit diskutierten Punkte zusammenfassend kann festgehalten werden, dass sich die Möglichkeiten, die von der positiven Kritik an empirischen Ansätzen in der Kunstwissenschaft hervorgehoben werden, tatsächlich realisieren lassen. Je nachdem wie eine empirische Studie geplant und umgesetzt wird, hält sie den Kritikpunkten der negativen Kritik im weiteren Sinne stand. Es bleibt dennoch festzuhalten, dass in der Praxis empirischen Arbeitens nicht immer das volle Potenzial der methodischen Ansätze ausgeschöpft wird. So ließen sich beispielsweise bestimmte als konzeptuell-reduktionistisch bewertbare Umsetzungen spezifischer Messgrößen dahingehend erweitern, dass nicht nur mehr Einzelaspekte, sondern auch deren wechselseitige Bezüge besser und vollständiger abgebildet wären. So gesehen sind die von der negativen Kritik betonten Probleme empirischer Ansätze durchaus berechtigt. Sie sollten sich allerdings weniger pauschal und allgemein auf die jeweiligen Methoden beziehen, sondern explizit die Praxis von empirisch-arbeitenden Forschenden in den Fokus nehmen. Denn es ist diese Praxis, welche bedingt, dass eine empirische Perspektive auf kunstwissenschaftliche Fragen häufig so wirkt als würde sie einer klassischen (hermeneutischen) geisteswissenschaftlichen Perspektive diametral gegenüberstehen. Ein solcher Schluss ist aber nur bedingt korrekt: Zwar ist es richtig, dass die unterschiedlichen Ansätze unterschiedliche erkenntnistheoretische Prämissen haben, dass das grundsätzliche Vorgehen hier jeweils ein völlig anderes ist, doch ist ganz dezidiert auch bei empirischen Herangehensweisen *nicht* von einer „Essentialisierung“ der Kunstwissenschaft zu sprechen, wie Peter Geimer befürchtet. Empirische Ansätze sind vielmehr so zu verstehen, dass überall dort, wo grundsätzliche Beobachtbarkeit gegeben ist, die Chance genutzt wird bzw. genutzt werden *kann*, den bestehenden Diskurs um einen Praxisbezug zu bereichern.

Die Empirie der Kunstwissenschaften, oder vielleicht auch der Geisteswissenschaften im Allgemeinen, ist damit in gewisser Hinsicht eine andere als die Empirie in den Natur-, Computer- und Sozialwissenschaften. In der Geistes- und/oder Kunstwissenschaft sind

---

empirische Ansätze immer Ergänzung, immer Teil einer Schleife, eines iterativ gestalteten Forschungsprozesses, der grundlegend auf einem Theorie- und Methodenmix basiert. Empirische Zugänge sind hier also Teil eines großen Geflechtes. Sie basieren auf der scheinbar paradoxen Integration von Hermeneutik und Empirie. Das zeigt sich schon allein daran, dass der rahmende Forschungsprozess, in welchen empirische Herangehensweisen eingegliedert sind, notwendigerweise hermeneutischer Natur ist. Da die meisten für kunstwissenschaftliche Überlegungen relevanten Konzepte diesem hermeneutischen Diskurs entnommen sind, sind die Gegenstände einer empirisch-arbeitenden Kunstwissenschaft hermeneutischen Ursprungs. Auch die Interpretation der Ergebnisse kann in keinem anderen Rahmen als dem Ursprungsrahmen der untersuchten Konzepte stattfinden, weswegen empirische Ergebnisse nach Abschluss der empirischen Untersuchung an den hermeneutischen Diskurs rückgebunden werden (können). Der interpretative Charakter geistes- und kunstwissenschaftlicher Forschung geht demnach durch die Integration empirischer Perspektiven nicht verloren – ganz im Gegenteil, empirische Untersuchungen lassen sich vielmehr als Einschübe, als Haltpunkte im hermeneutischen Diskurs verstehen, die das Potenzial bieten, diesen sowohl zu hinterfragen als auch zu befeuern.<sup>222</sup>

Diesem aus Natur und Praxis empirischer Methoden folgenden Schluss lässt sich bezüglich des dargestellten Konfliktes der *positiven* und *negativen* Kritik an empirischen Ansätzen folgern, dass hier ein noch grundlegendes Missverständnis besteht als bisher aufgezeigt. Die Missverständnisse liegen hier nicht nur im Detail, sondern in der grundsätzlichen Bewertung empirischer Herangehensweisen: Empirische Forschung als das verstanden, was sie wirklich ist, ist im Wesentlichen ein *methodischer* Ansatz. Das zu betonen, ist deswegen wichtig, weil mit dem Begriff *Methode* im geistes- und kunstwissenschaftlichen Kontext oft etwas anderes gemeint ist, als es an dieser Stelle der Fall ist. *Methode* bedeutet hier nicht, dass der Ansatz konstitutiv ist für die *theoretische* Perspektive, welche auf das untersuchte Material angelegt wird. Empirische Methoden haben also in theoretischer Hinsicht nicht eine

---

<sup>222</sup> Vgl. auch Drucker/Helmreich/Lincoln/Rose 2015.

---

„Essentialisierung“ der Kunstwissenschaft zum Ziel. Alle Formen der Kunstwissenschaft, ob soziologisch, rezeptionsästhetisch, feministisch, kulturvergleichend, etc., bleiben unabhängig von einem eventuellen Rückgriff auf ein empirisches Methodenrepertoire genau das, was sie sind. Sie werden nicht zu einer ominösen „empirischen Kunstwissenschaft“, denn der Begriff einer empirischen Kunstwissenschaft ist letztlich kaum mehr als eine Beschreibung eines Vorgehens, ähnlich dem „quellenbasierten Arbeiten“ oder der „bildvergleichenden Analyse“. Empirische Methoden sind also *nicht* Methoden im geisteswissenschaftlichen Sinne, sondern eher als ein Werkzeug bzw. eine Vorgehensweise zu verstehen. Dies ermöglicht ihre Eingliederung in den hermeneutischen Rahmenprozess und kann letztlich als Ansatzpunkt für einen Brückenschlag zwischen den Positionen der „mahrenden Propheten“ und „lamentierenden Unzufriedenen“ dienen.



# REFERENZEN

(exklusive der Referenzen der Studien)

---

**Aitken 1974**

P. P. Aitken, Judgments of pleasingness and interestingness as functions of visual complexity, in: Journal of Experimental Psychology, 103, 1974, S. 240-244.

**Anderson 2009**

Chris Anderson, The End of Theory. The Data Deluge Makes the Scientific Method Obsolete, in: Wired Magazine 2009: <https://www.wired.com/2008/06/pb-theory/>.

**Arora/Elgammal 2012**

Ravneet Singh Arora/ Ahmed Elgammal, Towards automated classification of fine-art painting style: A comparative study, in: Pattern Recognition (ICPR), 21, 2012, S. 3541-3544.

**Atteslander/ Kneubühler 1975**

Peter Atteslander/ Hans-Ulrich Kneubühler, Verzerrungen im Interview: zu einer Fehlertheorie der Befragung, Opladen 1975.

**Bachmann-Medick 2006**

Doris Bachmann-Medick, Cultural Turns. Neuorientierungen in den Kulturwissenschaften, Reinbek bei Hamburg 2006.

**Bal/Bryson 1991**

Mieke Bal/ Norman Bryson, Semiotics and Art History, in: The Art Bulletin, 73(2), 1991, S. 174-208.

**Balzer/Schröder/Schäfer 2011**

Helmut Balzer/ Marion Schröder/ Christian Schäfer, Wissenschaftliches Arbeiten: Ethik, Inhalt & Form wiss. Arbeiten, Handwerkszeug, Quellen, Projektmanagement, Präsentation, Witten 2011.



---

**Barker et al. 2012**

Elton Barker/Chris Bissell/Lorna Hardwick/Allan Jones/Mia Ridge/John Wolffe, Colloquium Digital technologies. Help or hindrance for the humanities?, in: Arts & Humanities in Higher Education 11(1-2), 2012, S. 185-200.

**Baumeister 1971**

Willi Baumeister, Die Natur der abstrakten Kunst, in: Götz Adriani (Hg.), Baumeister. Dokumente, Texte, Gemälde, Köln 1971.

**Baumeister 1952**

Willi Baumeister, „Über das Eigengesetzliche in der Kunst“, in: Westfälische Rundschau, 8.7.1952, Auszüge Online: <http://www.willi-baumeister.org/de/content/zitate-von-willi-baumeister> (Letzter Abruf am 13.10.17).

**Baxandall 1972**

Michael Baxandall, Painting and Experience in 15th Century Italy. A Primer in the Social History of Pictorial Style, Oxford 1972.

**Beck 1995**

Andreas Beck, Bildanalyse in der Kunst, in: Friedrich H. W. Heuck/ Eckard Macherauch (Hg.), Forschung mit Röntgenstrahlen. Bilanz eines Jahrhunderts (1895–1995), Springer 1995, S. 609-623.

**Bell/Ommer 2017**

Peter Bell/ Björn Ommer, Kunst messen, Pixel zählen? Die Zusammenarbeit zwischen Kunstgeschichte und Computer Vision oszilliert zwischen quantitativen und hermeneutischen Methoden, in: Marcel Schweiker, Joachim Hass et a. (Hg), Messen und Verstehen in der Wissenschaft. Interdisziplinäre Ansätze. Heidelberg 2017, S. 227–238.

---

**Bell/ Ommer 2016a**

Peter Bell/ Björn Ommer, Visuelle Erschließung. Computer Vision als Arbeits- und Vermittlungstool, in: Andreas Bienert (Hg.), EVA Berlin 2016, Elektronische Medien & Kunst, Kultur und Historie. Berlin 2016, S. 67–73.

**Bell/Ommer 2016b**

Peter Bell/ Björn Ommer, Digital Connoisseur? How Computer Vision Supports Art History, in: Stefan Albl/ Alina Aggujaro (Hg), „Il metodo del conoscitore. Approcci, limiti, prospettive. Rom 2016, S. 291–301.

**Bell/Ommer 2015**

Peter Bell/ Björn Ommer, Training Argus, in: Kunstchronik. Monatsschrift für Kunstwissenschaft, Museumswesen und Denkmalpflege, 68, 2015, S. 414-420.

**Benjamin 2012**

Walter Benjamin, Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit: drei Studien zur Kunstsoziologie, Frankfurt am Main 2012.

**Berlyne 1970**

Daniel E. Berlyne, Novelty, complexity, and hedonic value, in: Perception & Psychophysics, 8, 1970, S. 279-286.

**Berlyne 1963**

Daniel E. Berlyne, Complexity and incongruity variables as determinants of exploratory choice and evaluative ratings, in: Canadian Journal of Psychology, 17, 1963, S. 274-290.

---

**Berlyne et al. 1968**

Daniel E. Berlyne/ John C. Ogilvie/ L. C. C. Parham, The dimensionality of visual complexity, interestingness, and pleasingness, in: *Canadian Journal of Psychology*, 22, 1968, S. 376-387.

**Betz et al. 2010**

Juliane Betz, Martina Engelbrecht, Christoph Klein, Raphael Rosenberg, Dem Auge auf der Spur: Eine historische und empirische Studie zur Blickbewegung beim Betrachten von Gemälden, in: *IMAGE*, 11, Januar 2010:

<http://www.gib.unituebingen.de/image/ausgaben?function=fnArticle&showArticle=159> .

**Bin et al. 2017**

Wang Bin/ Yu Hui/ Guy Bovyn/Joost M.A. Caen, Identification of pigments from 'Zhenhai Temple': a Qing Dynasty export oil painting, in: *Sciences of conservation and archaeology*, 29 (2), 2017, S 82-88.

**Bindemann et al. 2005**

Markus Bindemann/ A. Mike Burton/ Ignace T. C. Hooge/ Rob Jenkins/ Edward H. F. de Haan, Faces retain attention, in: *Psychonomic Bulletin & Review*, 12, 2005, S. 1048–105.

**Bishop 2018**

Claire Bishop, Against Digital Art History, in: *International Journal for Digital Art History* 3, 2018: <https://doi.org/10.11588/dah.2018.3.49915>.

**Boehm 2008**

Gottfried Boehm, Die Form des Formlosen—Abstrakter Expressionismus und Informel, in: *Fondation Beyeler (Hg.), Action Painting: Jackson Pollock, Ostfildern 2008*, S. 38-46.

---

**Boehm 1994**

Boehm, Gottfried, Die Wiederkehr der Bilder, in: Gottfried Boehm (Hg.), Was ist ein Bild?, München 1994, S. 11-38.

**Bohnsack 2008**

Ralf Bohnsack, Rekonstruktive Sozialforschung – Einführung in qualitative Methoden, Opladen 2008.

**Büttner 2003**

Frank Büttner, Das Paradigma ‘Einfühlung’ bei Robert Vischer, Heinrich Wölfflin und Wilhelm Worringer, in: Christian Drude und Hubertus Kohle (Hg.), 200 Jahre Kunstgeschichte in München, München/Berlin 2003, S. 82-93.

**Bushart 2007**

Magdalena Bushart, ‘Form’ und ‘Gestalt.’ Zur Psychologisierung der Kunstgeschichte um 1900, in: Otto G. Oexle (Hg.), Krise des Historismus – Krise der Wirklichkeit, Göttingen 2007, S. 147-179.

**Brieber/Nadal/Leder 2015**

David Brieber/ Marcos Nadal/ Helmut Leder, In the white cube: Museum context enhances the valuation and memory of art, in: Acta Psychologica, 154, 2015, S. 36-42.

**Brieber et al. 2014**

David Brieber/ Marcos Nadal/ Helmut Leder/ Raphael Rosenberg, Art in time and space: context modulates the relation between art experience and viewing time. PloS one, 9(6), 2014: e99019.

---

**Brink 2000**

Claudia Brink, Der Bilderatlas: Mnemosyne in Warburg's Gesammelte Schriften, II.1, Berlin 2000.

**Brinkmann 2017**

Hanna Brinkmann, The Cultural Eye. Eine empirische Studie zur kulturellen Varianz in der Kunstwahrnehmung, Wien 2017: <http://othes.univie.ac.at/46664/>.

**Brinkmann/Commare 2015**

Hanna Brinkmann/ Laura Commare, Blick\_folgen: Zur Visualisierung von Augenbewegungen bei der Kunstbetrachtung, in: Visual Past 2(1), 2015, S. 389-405.

**Brinkmann et al. 2018**

Hanna Brinkmann/Eugene McSorley/Raphael Rosenberg/ Louis Williams, *Does "action viewing" really exist? The relationship between dynamic paintings and gaze behavior*, in: Abstracts der Visual Science of Art Conference 2018, S. 11: [http://www.vsac2018.eu/wp-content/uploads/2018/08/VSAC2018\\_Booklet.pdf](http://www.vsac2018.eu/wp-content/uploads/2018/08/VSAC2018_Booklet.pdf).

**Brinkmann et al. 2014**

Hanna Brinkmann/ Laura Commare/ Helmut Leder/ Raphael Rosenberg, Abstract Art as a Universal Language?, in: Leonardo 47 (3), 2014, S. 256-257.

**Caviness 2001**

Madeline H. Caviness, Art in the Medieval West and Its Audience, Aldershot 2001.

**Clausberg 2011**

Karl Clausberg, Aktualgenese & barocke Bewegtheit. Vor-Gestalten der Neuro-Ästhetik, in: Kunstgeschichte. Open Peer Reviewed Journal, Online: <http://www.kunstgeschichte-ejournal.net/259/> (Letzter Abruf am 05.10.17).

---

**Clausberg 1999**

Karl Clausberg, Neuronale Kunstgeschichte. Selbstdarstellung als Gestaltungsprinzip, Wien 1999.

**Collenberg-Plotnikov 2013**

Bernadette Collenberg-Plotnikov, 'Zoologen und Physiker als die berufensten Forscher in Sachen der Aesthetik'? – Zur Bestimmung der experimentellen Ästhetik in der Allgemeinen Kunstwissenschaft, in: Zeitschrift für Ästhetik und Allgemeine Kunstwissenschaft, 58 (1), 2013, S. 11-34.

**Commare 2011**

Laura Commare, Social Tagging als Methode zur Optimierung Kunsthistorischer Bilddatenbanken – Eine empirische Analyse des Artigo-Projekts, in: Kunstgeschichte. Open Peer Reviewed Journal: urn:nbn:de:bvb:355-kuge-160-9.

**Commare/Brinkmann 2017**

Laura Commare/ Hanna Brinkmann, Warum ‚Anything goes‘ der Goldstandard sein sollte – Überlegungen zu Methodentradition und empirischen Forschungsansätzen in den Kunstwissenschaften, in: Julia Rüdiger, Anna Sauer und Christina Bartosch (Hg.), Newest Art History – Wohin geht die jüngste Kunstgeschichte? Tagungsband zur 18. Tagung des Verbandes österreichischer Kunsthistorikerinnen und Kunsthistoriker, Online 2017, S.161-180:[https://voekk.at/sites/default/files/downloads/tagungsbaende/Newest%20Art%20History\\_VoeKK-Tagungsband.pdf](https://voekk.at/sites/default/files/downloads/tagungsbaende/Newest%20Art%20History_VoeKK-Tagungsband.pdf).

**Commare/Brinkmann 2016**

Laura Commare/ Hanna Brinkmann, Aesthetic Echoes in the Beholder's Eye? Empirical evidence for the divergence of theory and practice in the perception of abstract art, in: Michael Zimmermann (Hg.), Vision in Motion. Streams of Sensation and Configurations of Time, Berlin 2016, S. 221-234.

---

**Davidson 2008**

Cathy N. Davidson, Humanities 2.0: Promise, Perils, Predictions, in: PMLA 123(3), 2008, S.707-717.

**Davis 2015**

Whitney Davis, Knowing art historically, in: Zeitschrift für Kunstgeschichte, 78 (3/4), 2015, S. 341-344.

**Day 1967**

H. Day, Evaluations of subjective complexity, pleasingness and interestingness for a series of random polygons varying in complexity, in: Perception & Psychophysics, 2, 1967, S. 281-286.

**Diekmann 2007**

Andreas Diekmann, Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen, Berlin 2007.

**Diekmann/Bell 2016**

Lisa Diekmann /Peter Bell, Die Kunst als Ganzes. Heterogene Bilddatensätze als Herausforderung für die Kunstgeschichte und die Computer Vision, in: Konferenzabstrakts der 3. Tagung des Verbands Digital Humanities im deutschsprachigen Raum e.V., S. 118: <http://www.dhd2016.de/abstracts/votr%C3%A4ge-004.html>.

**Dressen 2017**

Angela Dressen, *Grenzen und Möglichkeiten der digitalen Kunstgeschichte und der Digital Humanities – eine kritische Betrachtung der Methoden*, in: kunsttexte 4, 2017: <https://edoc.hu-berlin.de/bitstream/handle/18452/19402/Angela%20Dressen%20-%20final.pdf> (S.1-17).

---

**Drucker 2013**

Johanna Drucker, Is there a “digital” art history? in: Visual Resources 29(1-2), 2013, S. 5-13.

**Drucker 2009**

Johanna Drucker, Speclab. Digital Aesthetics and Projects in Speculative Computing, Chicago 2009.

**Drucker/Helmreich/Lincoln/Rose 2015**

Johanna Drucker/ Anne Helmreich /Matthew Lincoln/ Francesca Rose, Digital art history. The American scene [07.12.2015], in: Perspective 2, 2015: <http://perspective.revues.org/6021>.

**Eco 1977**

Umberto Eco, Zeichen. Einführung in einen Begriff und seine Geschichte, Frankfurt a. M. 1977.

**Eisenman 1967**

R. Eisenmann, Complexity-simplicity: I. Preference for symmetry and rejection of complexity, in: Psychonomic Science, 8(4), 1967, S. 169-170.

**Eisenman/Gillens 1968**

R. Eisenmann/ H. Gillens, Preferences for complexity-simplicity and symmetry-asymmetry, in: Perceptual and Motor Skills, 26, 1968, S. 888-890.

**Elkins 1997**

Elkins, James: What are we seeing, exactly, in: Kathleen Cohen/James Elkins/Marilyn Aronberg Lavin/Nancy Macko/Gary Schwartz/Susan L. Siegfried/Barbara Maria Stafford, A



---

Range of Critical Perspectives. Digital Culture and the Practices of Art and Art History, in: The Art Bulletin 79(2), 1997, S. 187-216, darin S. 191-198.

**Farley/Weinstock 1980**

Frank H. Farley/ Camilla A. Weinstock, Experimental aesthetics: Children's complexity preference in original art and photoreproductions, in: Bulletin of the Psychonomic Society, 15, 1980, S. 194-196.

**Floridi 2015**

Luciano Floridi, Die 4. Revolution: Wie die Infosphäre unser Leben verändert, Berlin 2015.

**Frohne et al. 2015**

Ursula Frohne/Johannes Grave/Jeffrey F. Hamburger/ Michael F. Zimmermann, Editorial, in: Zeitschrift für Kunstgeschichte, 78(1), 2015, S. 1-5.

**Gagliardi/ Petridis 2015**

Susan Elizabeth Gagliardi/Constantine Petridis, Senoufo Unbound: Dynamics of Art and Identity in West Africa, in: African Arts, 48 (1), 2015, S. 6-23.

**Geimer 2016**

Peter Geimer, Jede Ähnlichkeit wäre rein zufällig. Was haben eine AfD-Politikerin und eine gemalte Herzogin gemeinsam? Beide lächeln! Londons Tate Gallery arbeitet mit Technik-Schnickschnack an der Selbstaufgabe wissenschaftlicher Kritik, in: Frankfurter Allgemeine Zeitung, Nr. 274, 23.11.2016, S. 9.

**Geimer 2015**

Peter Geimer, Jenseits der turns. Zwei Beobachtungen, in: Zeitschrift für Kunstgeschichte, 78 (3/4), 2015, S. 336-340.

---

**Gombrich 1966**

Ernst H. Gombrich, Norm and Form. Studies in the art of the Renaissance, London 1966.

**Grau 2015**

Oliver Grau, Ohne Interpretation ist Technologie wertlos (Interview), in: Kunstgeschichte aktuell, 3, 2015.

**Greenberg 1986**

Clement Greenberg, The collected essays and criticism, Vol. 2, Arrogant Purpose, 1945–1949, John O’Brian (Hg.), Chicago 1986.

**Guthmüller/Klein 2006**

Marie Guthmüller/ Wolfgang Klein (Hg.), Ästhetik von unten. Empirisches und ästhetisches Wissen, Tübingen 2006.

**Haftmann 1959**

Werner Haftmann, Malerei nach 1945, Einführung von Werner Haftmann, in: II. documenta ’59, Kunst nach 1945, Ausstellungskatalog Kassel, Band 1: Malerei, Köln 1959, S. 11-19.

**Haftmann 1954**

Werner Haftmann, Malerei im 20. Jahrhundert. Eine Entwicklungsgeschichte, München 1954.

**Hasenmueller 1978**

Christine Hasenmueller, Panofsky, Iconography, and Semiotics, in: The Journal of Aesthetics and Art Criticism, 36(3), 1978, S. 289-301.

---

**Hayles 2012**

N. Katherine Hayles, *How We Think. Digital Media and Contemporary Technologies*. Chicago/London 2012.

**Henrich/Heine/Norenzayan 2010**

Joseph Henrich/ Steven J. Heine/ Ara Norenzayan, *The weirdest people in the world?*, in: *Behavioral and Brain Sciences* 33, 2010, S. 61-135.

**Hofmann 1987**

Werner Hofmann, *Grundlagen der Modernen Kunst. Eine Einführung in ihre symbolischen Formen*, Stuttgart 1987.

**Holmqvist et al. 2011**

Kenneth Holmqvist/ Marcus Nyström/ Richard Andersson/ Richard Dewhurst/ Halszka Jarodzka/ Joost van de Weijer, *Eye Tracking. A comprehensive guide to methods and measures*, Oxford 2011.

**Hoormann 2007**

Anne Hoormann, *Die Entfernung der Inhalte aus dem Bild. Zur Konstruktion des "unschuldigen Sehens" im Abstrakten Expressionismus*, in: Dieter Burdorf, Mechthild Fend, Bettina Uppenkamp (Hg.), *Medium und Material. Zur Kunst der Moderne und der Gegenwart*, München 2007, S. 262-275.

**Hoppe 2001**

Hoppe, Stephan, *Die Fußnoten des Modells. CAD-Modelle als interaktive Wissensräume am Beispiel des Altenberger-Dom-Projektes*, in: Marcus Frings (Hrsg.), *Der Modelle Tugend. CAD und die neuen Räume der Kunstgeschichte*, Weimar 2001, S. 87 - 102.

---

**Huemer 2016**

Christian Huemer, Mapping Markets. Der Getty Provenance Index auf dem Prüfstand [Vortrag am 06.12.2016 im Rahmen des Panels „Collections, Collecting and the Art Market. Digital perspectives for Art Historical Research I“ auf der 3rd DHA Conference in Wien].

**Ilsink 2016**

Matthijs Ilsink/ Jos Koldeweij/ Ron Spronk/ Luuk Hoogstede/ Robert G. Erdmann/ Rik Klein Gotink/ Hanneke Nap/ Daan Veldhuizen, Hieronymus Bosch, Painter and Draughtsman, Catalogue Raisonné, Brussels 2016.

**Imdahl 1980**

Max Imdahl, Giotto. Arenafresken. Ikonographie – Ikonologie – Ikonik, München 1980.

**Itti/ Koch 2000**

Laurent Itti / Christof Koch, A saliency-based search mechanism for overt and covert shifts of visual attention, in: Vision Research, 40(10-12), 2000, S. 1489-1506.

**Jakesch/Leder 2016**

Martina Jakesch/ Helmut Leder, The Qualitative Side of Complexity: Testing Effects of Ambiguity on Complexity Judgments, in: Psychology of Aesthetics, Creativity and the Arts, 9(3), 2016, S. 200-205.

**Jannidis/Kohle/Rehbein 2017**

Fotis Jannidis/ Hubertus Kohle/ Malte Rehbein (Hrsg.), Digital Humanities. Eine Einführung. Stuttgart 2017.

**Jarzombek 1994**

Mark Jarzombek, De-Scribing the Language of Looking: Wölfflin and the History of Aesthetic Experientialism, in: Assemblage, 23, 1994, S. 28-69.

---

### **Joyeux-Prunel 2016**

Béatrice Joyeux-Prunel, Voir plus grand, plus loin, plus long. Le numérique au service d'une « histoire totale » de l'art, in: Zeitschrift für Kunstgeschichte, 79(3), 2016, S. 299–304.

### **Kaiser 2014**

Maximilian Kaiser, „Struktur, Netzwerk, Diskurs. Anatomie einer Künstlervereinigung“, in: Agnes Husslein-Arco (Hg.), Hagenbund. Ein europäisches Netzwerk der Moderne (1900-1938), Ausst.-Kat. Österreichische Galerie Belvedere, München 2014, S. 103-112.

### **Kandinsky 1926**

Wassily Kandinsky, Punkt und Linie zu Fläche, München 1926.

### **Kandinsky 1911**

Wassily Kandinsky, Über das Geistige in der Kunst. Insbesondere in der Malerei, München 1911.

### **Kemp 1992**

Wolfgang Kemp, Der Betrachter ist im Bild, Kunstwissenschaft und Rezeptionsästhetik, Berlin 1992.

### **Kemp 1986**

Wolfgang Kemp, Kunstwerk und Betrachter: Der rezeptionsästhetische Ansatz, in: Hans Belting/ Heinrich Dilly/ Wolfgang Kemp/ Willibald Sauerländer/ Martin Warnke (Hrsg.), Kunstgeschichte. Eine Einführung, Berlin 1986, S. 203-221.

---

**Kindt/Schmidt 1979**

Walther Kindt/ Siegfried J. Schmidt, Motivationen und Aspekte einer empirischen Literaturwissenschaft, in: Siegfried J. Schmidt (Hg.), *Empirie in Literatur- und Kunstwissenschaft*, München 1979, S. 7-42.

**Kirchberg/Tröndle 2015**

Volker Kirchberg/ Martin Tröndle, *The Museum Experience: Mapping the Experience of Fine Art*, in: *Curator – The Museum Journal*, 58 (2), 2015, S. 1-25.

**Klinke 2018**

Harald Klinke, *Datenanalyse in der Digitalen Kunstgeschichte. Neue Methoden in Forschung und Lehre und der Einsatz des DHVLab in der Lehre*, in: Harald Klinke (Hrsg.), *#DigiCampus Digitale Forschung und Lehre in den Geisteswissenschaften*, readbox unipress / Open publishing LMU 2018, S. 19-34.

**Klinke 2016**

Harald Klinke, *Digitale Kunstgeschichte*, in: *Akademie Aktuell* 1, 2016, S. 68 – 71.

**Knorr-Cetina 1999**

Karin Knorr-Cetina, *Epistemic Cultures. How the sciences make knowledge*, Cambridge/Mass 1999.

**Kohle 2018**

Hubertus Kohle, *Die Geisteswissenschaften und das Digitale. Ein Quantensprung oder business as usual?* In Harald Klinke (Hrsg.), *#DigiCampus Digitale Forschung und Lehre in den Geisteswissenschaften*, readbox unipress / Open publishing LMU 2018, S. 9-18.

---

**Kohle 2016**

Hubertus Kohle, Kunstgeschichte und *Digital Humanities*. Einladung zu einer Debatte, in: Zeitschrift für Kunstgeschichte, 79(2), 2016, S. 151-154.

**Kohle 2013**

Hubertus Kohle, Digitale Bildwissenschaft, Verlag Werner Hülsbusch 2013.

**Kohle 2011**

Hubertus Kohle, Kunstgeschichte goes Social Media. Laien optimieren eine Bilddatenbank – mit einem digitalen Spiel, in: Aviso. Zeitschrift für Wissenschaft und Kunst in Bayern, 3, 2011, S. 38-43.

**Kohle 2008**

Hubertus Kohle, Ordnung und Umbruch in der Wissenschaft der Kunstgeschichte: Der Einfluss des Digitalen, in: Hubach, Hanns (Hrsg.), Reibungspunkte: Ordnung und Umbruch in Architektur und Kunst; Festschrift für Hubertus Günther, Petersberg 2008, S. 141-144.

**Kohle 1997**

Hubertus Kohle, Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende. Berlin 1997.

**Kohle/Brassat 2009**

Hubertus Kohle/ Wolfgang Brassat, Methoden-Reader Kunstgeschichte. Texte zur Methodik und Geschichte der Kunstwissenschaft, Köln 2009.

**Krause/Reichle 2014**

Celia Krause/ Ruth Reiche, Ein Bild sagt mehr als tausend Pixel? Über den Einsatz digitaler Methoden in den Bild- und Objektwissenschaften, in: Kunstgeschichte. Open Peer Reviewed Journal, 2014.

---

**Kreitler et al. 1974.**

Shulamith Kreitler/ Edward Zigler/ Hans Kreitler, The complexity of complexity, in: *Human Development*, 17, 1974, S. 54-73.

**Krupinski/Locher 1988**

Elizabeth Krupinski/ Paul Locher, Skin conductance and aesthetic evaluative response to nonrepresentational works of art varying in symmetry, in: *Bulletin of the Psychonomic Society*, 26, 1988, S. 355-358.

**Küster 2008**

Ulf Küster, Action Painting—Mythos und Realität, in: Fondation Beyeler (Hg.), *Action Painting: Jackson Pollock, Ostfildern 2008*, S. 12-19.

**Langton et al. 2008**

Stephen R. H. Langton/ Anna S. Law/ A. Mike Burton/ Stefan R. Schweinberger, Attention capture by faces, in: *Cognition*, 107, 2008, S. 330–342.

**Lanzoni/Brain/Young 2012**

Susan Lanzoni/ Robert Brain/ Allan Young (Hg), Sonderheft: The Varieties of Empathy in Science, Art, and History, in: *Science in Context*, 25 (3).

**Latour/Hermant 1998**

Bruno Latour/ Emilie Hermant, *Paris ville invisible*. Paris 1998.

**Laursen/Hoff 2006**

Brett Paul Laursen/ Erika Hoff, Person-centered and variable-centered approaches to longitudinal data, in: *Merrill-Palmer Quarterly*, 52(3), 2006, S. 377-389.



---

**Leder/Nadal 2014**

Helmut Leder/ Marcos Nadal, Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode—Developments and challenges in empirical aesthetics, in: *British Journal of Psychology*, 105(4), 2014, S. 443-464.

**Leder et al. 2014**

Helmut Leder/ Gernot Gerger/ David Briber/ Norbert Schwarz, What makes an art expert? Emotion and evaluation in art appreciation, in: *Cognition and Emotion*, 28(6), 2014, S. 1137-1147.

**Leder et al. 2004**

Helmut Leder/ B ... Belke/ A... Oeberst/ Do ....Augustin, A model of aesthetic appreciation and aesthetic judgements, in: *British Journal of Psychology* 95(4), 2004, S. 489-508.

**Leek/Peng 2016**

Jeffrey T. Leek/ Roger D Peng, Opinion: reproducible research can still be wrong: Adopting a prevention approach, in: *PNAS*, 112 (6), 2016, S. 1645-1646.

**Lengyel 2011**

Dominik Lengyel, Ein Stadtmodell von Pergamon - Unschärfe als Methode für Darstellung und Rekonstruktion antiker Architektur, in: Lars Petersen/ Ralf von den Hoff (Hrsg.), *Skulpturen in Pergamon - Gymnasion, Heiligtum, Palast*, Freiburg 2011, S. 22-26.

**Leonhart 2009**

Rainer Leonhart, *Lehrbuch Statistik*, Bern 2009.

**Locher 2001**

Hubert Locher, *Kunstgeschichte als historische Theorie der Kunst 1750 – 1950*, München 2001.

---

**Manovich 2015**

Lev Manovich, Data science and digital art history, in: International Journal for Digital Art History, 1, 2015, S. 14-35.

**Manovich 2014**

Lev Manovich, Trending. Verheißungen und Herausforderungen der Big Sozial Data, in: Ramón Reichert (Hg.), BIG DATA. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 65-84.

**Manovich 2011**

Lev Manovich, Style Space: How to compare image sets and follow their evolution. Software Studies Initiative 2011: <http://lab.softwarestudies.com/2011/08/style-space-how-to-compare-image-sets.html?m=1> (Letzter Abruf am 18.01.2018).

**Manovich 2002**

Lev Manovich, The Language of New Media, Cambridge/Mass 2002.

**Marmor 2016**

Max Marmor, Art History and the Digital Humanities, in: Zeitschrift für Kunstgeschichte, 79(2), 2016, S. 151-158.

**Massaro et al. 2012**

Davide Massaro/ Federica Savazzi/ Cinzia Di Dio/ David Freedberg/ Vittorio Gallese/ Gabriella Gilli/ Antonella Marchetti, When Art Moves the Eyes: A Behavioral and Eye-Tracking Study, in: PLoS ONE, 7(5), 2012: e37285.

---

**Matell/Jacoby 1971**

Michael S. Matell/ Jacob Jacoby, Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity, in: Educational and Psychological Measurement 31(3), 1971, S. 657-674.

**Messinger 1998**

S. M. Messinger, Pleasure and Complexity: Berlyne revisited, in: The Journal of Psychology, 132, 1998, S. 558-560.

**Mitrović 2013**

Branko Mitrović, Visuality After Gombrich: the Innocence of the Eye and Modern Research in the Philosophy and Psychology of Perception, in: Zeitschrift für Kunstgeschichte, 76(1), 2013, S. 71 – 89.

**Mitchell 1994**

William Mitchell, The Pictorial Turn, in William Mitchell (Hg), Picture Theory, Chicago 1994, S. 11-34.

**Molnar/ Ratsikas 1987**

François Molnar/ Demitrios Ratsikas, Some aesthetical aspects of visual exploration, in: J. K. O'Regan and A. Levy-Schoen (Hg), Eye Movements from Physiology to Cognition, Amsterdam 1987, S. 363-374.

**Monroy/Bell/Ommer 2014**

Antonio Monroy/ Peter Bell/ Björn Ommer Morphological Analysis for Investigating Artistic Images, in: Image and Vision Computing, 32, 2014, S. 414-423.

---

**Moravánszky 2012**

Ákos Moravánszky, The Optical Construction of Urban Space: Hermann Maertens, Camillo Sitte and the Theories of 'Aesthetic Perception,' in: The Journal of Architecture, 17(5), 2012, 655-666.

**Morris 1939**

Charles M. Morris, Esthetics and the Theory of Signs. The journal of Unified Science, 8(1), 1939, S. 131-150.

**Morris 1938**

Charles M. Morris, Foundation of the Theory of Signs. Chicago: 1938.

**Müller-Tamm/Schmidgen/Wilke 2014**

Jutta Müller-Tamm/ Henning Schmidgen/ Tobias Wilke (Hg.), Gefühl und Genauigkeit. Empirische Ästhetik um 1900, München 2014.

**Munsinger/Kessen 1964**

H. Munsinger/ W. Kessen, Uncertainty, structure, and preference. In: Psychological Monographs: General and Applied, 78, 1964, S. 1-24.

**Murphy/Jako/Anhalt 1993**

Kevin R. Murphy/ Robert A. Jako/ Rebecca L. Anhalt, Nature and consequences of halo error: A critical analysis, in: Journal of Applied Psychology, 78(2), 1993, S. 218-225.

**Nadal 2007**

Marcos Nadal, Complexity and aesthetic preference for diverse visual stimuli, psych. Diss. (unpubl.), Palma 2007, Online:

<http://ibdigital.uib.es/greenstone/collect/tesisUIB/archives/TDX-0404.dir/TDX-0404108-112455.pdf> (Letzter Abruf am 09.10.2017).

---

**Nadal et al. 2010**

Marcos Nadal/ Enric Munar/ Gisèle Marty/ Camilo J. Cela-Conde, Visual Complexity and Beauty Appreciation: Explaining the Divergence of Results, in: *Empirical Studies of the Arts*, 28(2), 2010, S. 173-191.

**Neperud/Marschalek 1988**

R. W. Neperud/ D. G. Marschalek, Information and affect bases of aesthetic response, in: *Leonardo*, 21, 1988, S. 305-312.

**Newman 1992**

Barnett Newman, *The Sublime is Now* (1947), in: John P. O'Neill (Hg.), *Selected Writings and Interviews*, Berkeley and Los Angeles: 1992, S. 170-175.

**Nicki 1972**

R. M. Nicki, Arousal increment and degree of complexity incentive, in: *British Journal of Psychology*, 63, 1972, S. 165-171.

**Nicki/Lee/Moss 1981**

R. M. Nicki/ P. L. Lee/ V. Moss, Ambiguity, cubist works of art, and preference, in: *Acta Psychologica*, 49, 1981 S. 27-41.

**Nicki/Moss 1975**

R. M. Nicki/ V. Moss, Preference for non-representational art as a function of various measures of complexity, in: *Canadian Journal of Psychology*, 29, 1975, S. 237-249.

**Nuzzo 2014**

Regina Nuzzo, Scientific method: statistical errors. P values, the 'goldstandard' of statistical validity, are not as reliable as many scientists assume, in: *Nature*, 506, 2014, S. 150–152.

---

**Onians 2007**

John Onians, *Neuroarthistory: From Aristotle and Pliny to Baxandall and Zeki*, New Haven 2007.

**Osborne/ Farley 1970**

John W. Osborne/ Frank H. Farley, The relationship between aesthetic preference and visual complexity in abstract art, in: *Psychonomic Science*, 19, 1970, S. 69-70.

**Panofsky 1983**

Erwin Panofsky, *Meaning in the Visual Arts* (1955), Chicago 1983.

**Pächt 1977**

Otto Pächt, *Methodisches zur kunsthistorischen Praxis. Ausgewählte Schriften*, München 1977.

**Parsons 2005**

Sarah Parsons, What Lies Beyond the Slide Library? Facing the Digital Future of Art History, in: *RACAR. Revue d'art canadienne* 30(1/2), 2005, S. 114-125.

**Pfisterer 2016**

Ulrich Pfisterer, Hans im Glück: Kunstgeschichte heute, in: *Zeitschrift für Kunstgeschichte*, 79 (1), 2016, S. 3-9.

**Pias 2003**

Claus Pias, Das digitale Bild gibt es nicht. Über das (Nicht-) Wissen der Bilder und die informatische Illusion [08.05.2003], in: *Zeitenblicke* 2(1). 2003:  
<http://www.zeitenblicke.historicum.net/2003/01/pias/index.html>.

---

**Pierce 1998**

Charles Sanders Pierce, Pragmatism (1907), in: Pierce Edition Project (Hg.), The Essential Peirce, Volume 2, Selected Philosophical Writings (1893-1913), Bloomington 1998, S. 398-433.

**Pollock 2014**

Griselda Pollock, Computers Can Find Similarities Between Paintings – but Art History Is About So Much More, 2014: <http://theconversation.com/computers-can-find-similarities-between-paintings-but-art-history-is-about-so-much-more-30752>.

**Rabe-Hesketh/Skrondal 2012**

Sophia Rabe-Hesketh/ Anders Skrondal, Multilevel and Longitudinal Modeling Using Stata, College Station 2012.

**Reiss/Reiss 1997**

Michael Reiss/ Gilfe Reiss, Ocular dominance: some family data, in: Laterality: Asymmetries of Body, Brain and Cognition, 2 (1), 1997, S. 7-16.

**Reitstätter 2015**

Luise Reitstätter, Die Ausstellung verhandeln. Von Interaktionen im musealen Raum, Bielefeld 2015.

**Rieder/Röhle 2012**

Bernhard Rieder/Theo Röhle, Digital Methods: Five Challenges, in: David M. Berry (Hg.), Understanding Digital Humanities, Springer 2012, S. 67-84.

**Ro/ Friggel/ Lavie 2007**

Tony Ro/ Ashley Friggel/ Nilli Lavie, Attentional biases for faces and body parts, in: Visual Cognition, 15, 2007, S. 322–348.

---

### **Rosenberg 2016**

Raphael Rosenberg, Bridging Art History, Computer Science and Cognitive Science: A Call for Interdisciplinary Collaboration, in: Zeitschrift für Kunstgeschichte, 79(3), 2016, S. 305-314.

### **Rosenberg 2011**

Raphael Rosenberg, Dem Auge auf der Spur. Blickbewegungen beim Betrachten von Gemälden – historisch und empirisch, in: Jahrbuch der Heidelberger Akademie der Wissenschaften für 2010, Heidelberg 2011, S. 76-89.

### **Rosenberg/Klein 2015**

Raphael Rosenberg/ Christoph Klein, The Moving Eye of the Beholder. Eye-Tracking and the Perception of Paintings, in: J. P. Huston, M. Nadal, F. Mora, L. Agnati & C. J. Cela-Conde (Hg.), Art, Aesthetics and the Brain, Oxford 2015, S. 79-108.

### **Rosenberg 1952**

Harold Rosenberg, The American Action Painters, in: ARTnews, 51 (8), 1952.

### **Rushkoff 2014**

Douglas Rushkoff, Present Shock. When Everything Happens Now, New York 2014.

### **Saklofske 1975**

D. H. Saklofske, Visual aesthetic complexity, attractiveness and diverse exploration, in: Perceptual and Motor Skills, 41, 1975, S. 813-814.

### **Saleh et al. 2016**

Babak Saleh/ Kanako Abe/ Ravneet Singh Arora/ Ahmed Elgammal, Toward automated discovery of artistic influence, in: Multimedia Tools and Applications, 75(7), 2016, S. 3565-3591.



---

**Saussure 2001**

Ferdinand de Saussure, Grundfragen der allgemeinen Sprachwissenschaft (1916), Charles Bally/ Albert Sechehaye (Hg.). New York 2001.

**Scherz 2017**

Sabine Scherz, Kunstgeschichte berechnet: Interdisziplinäre Bilddatenanalyse crowdgesourcter Annotationen, München 2017.

**Schich et al. 2017**

Maximilian Schich/ Christian Huemer/ Piotr Adamczyk/ Lev Manovich/ Yang-Yu Liu, Network Dimensions in the Getty Provenance Index, 2017: arXiv preprint arXiv:1706.02804.

**Schulze 1990**

Gerhard Schulze, Die Transformation sozialer Milieus in der Bundesrepublik Deutschland. Lebenslagen, Lebensläufe, Lebensstile. in: Soziale Welt, Sonderband 7, 1990, S 409–432.

**Schwarz et al. 1985**

Norbert Schwarz/ Hans-J. Hippler/ Brigitte Deutsch/ Fritz Strack, Response scales: Effects of category range on reported behavior and comparative judgments, in: Public Opinion Quarterly, 49(3), 1985, S. 388-395.

**Specker et al. *in press***

Eva Specker/ Raphael Rosenberg/ Helmut Leder/ Jane Boddy/ Hanna Brinkmann/ Michael Forster/ Matthew Pelowski, Introducing the VAIK. A new and validated way to measure Art interest and Art Knowledge, PACA, in Vorbereitung.

---

**Speidel 2017**

Klaus Speidel, Can there be experimental reception research in art-history? Taking the skeptic seriously, in: Konferenzabstrakts der 3rd Art & Science Conference on Empirical Methods in Art History and Visual Studies, 2017, Online:  
<https://artandscience.univie.ac.at/2017/abstracts/>

**Thalwitzer/Brinkmann/Rosenberg 2015**

Mario Thalwitzer/ Hanna Brinkmann/ Raphael Rosenberg, Using a concealed, calibration free eye-tracker for studying art perception in the museum, in: Ulrich Ansorge et al. (Hg.), Abstracts of the 18th European Conference on Eye Movements, Journal of Eye Movement Research, 2015, 8 (4), S. 140.

**Thiel 2012**

Thomas Thiel, Eine empirische Wende für die Geisteswissenschaften? in: Frankfurter Allgemeine Zeitung, 24.07.2012, Online: <http://www.faz.net/aktuell/feuilleton/forschung-und-lehre/digital-humanities-eine-empirische-wende-fuer-die-geisteswissenschaften-11830514.html>.

**Thoma/Lavie 2013**

Volker Thoma/ Nilli Lavie, Perceptual load effects on processing distractor faces indicate face-specific capacity limits, in: Visual Cognition, 21, 2013, S. 1053–1076.

**Thompson 1973**

Stuart R. Thompson, Complexity in Perception: An Explanation in Phenomenological Terms by Psychology, with some Parallels in Aesthetic Philosophy, in: Studies in Art Education, 14(2), 1973, S. 3-14.

---

**Thorndike 1920**

Edward L. Thorndike, A constant error in psychological ratings, in: *Journal of Applied Psychology*, 4(1), 1920, S. 25-29.

**Tinio/Leder 2009**

Pablo Tinio/ Helmut Leder, Just how stable are aesthetic features? Symmetry, complexity, and the jaws of massive familiarization, in: *Acta Psychologica*, 130, 2009, S.241-250.

**Tröndle et al. 2014**

Martin Tröndle/Steven Greenwood /Volker Kirchberg/Wolfgang Kirchberg, An Integrative and Comprehensive Methodology for Studying Aesthetic Experience in the Field: Merging Movement Tracking, Physiology and Psychological Data, in: *Environment and Behavior*, 46(1), 2014, S. 102-135.

**Tröndle/ Tschacher 2012**

Martin Tröndle/ Wolfgang Tschacher, The Physiology of Phenomenology: The Effects of Artworks, in: *Journal of Empirical Studies of the Arts*, 30(1), 2012, S. 79-117.

**Tschacher et al. 2012**

Wolfgang Tschacher / Steven Greenwood / Volker Kirchberg/ Stéphanie Wintzerith/ Karen van den Berg / Martin Tröndle, Physiological correlates of aesthetic perception in a museum, in: *Journal of Psychology of Aesthetics, Creativity, and the Arts*, 6, 2012, S. 96-103.

**Underwood/Foulsham 2006**

Geoffrey Underwood/ Tom Foulsham, Visual saliency and semantic incongruency influence eye movements when inspecting pictures, in: *The Quarterly Journal of Experimental Psychology*, 59(11), 2006, S. 1931-1949,

---

**Vahland 2010**

Kia Vahland, Die Vermessung der Kunst. Monet beruhigt, Richter macht nervös: Mit einem Datenhandschuh sollen Körpergefühle beim Museumsrundgang erfasst werden - doch Medizin ist nicht alles, in: Süddeutsche Zeitung, 17.05.2010, Online: <http://www.sueddeutsche.de/kultur/st-gallen-kunst-trifft-medizin-die-vermessung-der-kunst-1.90610> (Letzter Abruf am 05.10.17).

**Vasold 2011**

Georg Vasold, Ernst Brücke und die Anfänge der Wiener Schule der Kunstgeschichte, in: *Austriaca. Cahiers universitaires d'information sur l'Autriche* (Université Rouen), 72, 2011, S. 101-116.

**Vitz 1966**

Paul C. Vitz, Preference for different amounts of visual complexity, in: *Behavioral Science*, 11, 1966, S. 105-114.

**Wintzerith/van den Berg/Tröndle 2011**

Stéphanie Wintzerith/ Karen van den Berg / Martin Tröndle, Spuren der Aneignung – Mapping Museums, in: Gerhard Kilger und Wolfgang Müller-Kuhlmann (Hg.), *Szenografie in Ausstellungen und Museen V - Raum und Wahrnehmung. Bewegte Räume*, Essen 2011, S. 102-111.

**Wölfflin 1915**

Heinrich Wölfflin, *Kunstgeschichtliche Grundbegriffe. Das Problem der Stilbildung in der neueren Kunst*, München 1915.

**Wohlwill 1968**

Joachim F. Wohlwill, Amount of stimulus exploration and preference as differential functions of stimulus complexity, in: *Perception & Psychophysics*, 4, 1968, S. 307-312.

---

**Wright/Ward 2008**

Richard D Wright/ Lawrence M. Ward, Orienting of attention, New York: 2008.

**Yarbus 1967**

Alfred L. Yarbus, Eye movements and Vision, New York 1967.

**Zeki 1999**

Semir Zeki, Inner vision: An exploration of Art and the Brain, Oxford 1999.

**Zeki 1998**

Semir Zeki, Art and the Brain, in: Daedalus, 127(2), 1998, S. 71-103.

**Zimbardo 1995**

Philip G. Zimbardo, Psychologie, Berlin 1995.

**Zimmer 1997**

Jürgen Zimmer, Kunstforschung – Kunstwissenschaft – Kunstgeschichte →  
CyberArtHistory?, in: AKMB-news 3(3),1997, S. 3-12.

**Zorich 2012**

Diane M. Zorich, Transitioning to a Digital World. Art History, Its Research Centers, and Digital Scholarship. A Report to the The Samuel H. Kress Foundation and The Roy Rosenzweig Center for History and New Media, George Mason University [Juni 2012]:  
[http://www.kressfoundation.org/research/transitioning\\_to\\_a\\_digital\\_world/](http://www.kressfoundation.org/research/transitioning_to_a_digital_world/).

---

# ABBILDUNGSVERZEICHNIS

---

**Abbildung 1 | S. 86**

Ernst, Max (1945). *Die Versuchung des Hl. Antonius* (108 x 128 cm), Foto: Artstor, © Wilhelm-Lehmbruck Museum, Duisburg.

**Abbildung 2 | S. 87**

Motherwell, Robert (1946). *Western Air* (182,9 x 137,2 cm), Foto: Artstor, © Museum of Modern Art, New York.

**Abbildung 3 | S. 91**

Verteilung der durchschnittlichen Komplexitätseinschätzungen, eigene Darstellung.

**Abbildung 4 | S. 92**

Durchschnittliche Komplexitätsbewertung pro Bild, eigene Darstellung.

**Abbildung 5 | S. 93**

Werkbeispiele mit großen Abweichungen (über 1,5 Ratingpunkte) in den durchschnittlichen Komplexitäts-einschätzungen von Laien und Expert\_innen im niedrigen bis mittleren Komplexitätsspektrum (oben von links: Monet, Claude (1897). *Morning on the Seine near Giverny* (89.9 x 92.7 cm), Foto: Artstor, © Metropolitan Museum of Art, New York; van Gogh, Vincent (1889). *Portrait eines jungen Bauerns mit Hut* (65 cm × 54 cm), Foto: Artstor, © Collection E. Sforzi, Florenz; unten: Moran, Thomas (1884). *The Much Resounding Sea* (63.9 x 158.2 cm), Foto: Artstor, © National Gallery of Art, Washington DC.).

**Abbildung 6 | S. 94**

Werkbeispiel mit großen Abweichungen (über 1,5 Ratingpunkte) in den durchschnittlichen Komplexitäts-einschätzungen von Laien und Expert\_innen im hohen Komplexitäts-spektrum (Prost, Frans Jansz (1665). *Brasilianische Landschaft* (56.2 × 83.5 cm), Foto: Artstor, © Detroit Institute of Arts, Detroit.).

**Abbildung 7 | S. 96**

Werkbeispiele mit großen Abweichungen in den durchschnittlichen Komplexitäts-einschätzungen von Laien und Expert\_innen (oben: Tizian (1514). *Himmlische und irdische Liebe* (118 × 279 cm), Foto: Artstor, © Galleria Borghese; unten von links: Crespi, Giuseppe Maria (1700). *Kentaur Chiron bringt dem jungen Achilles das Bogenschiessen bei* (125 × 123cm), Foto: Artstor, © Kunsthistorisches Museum Wien; Memling, Hans (1475). *Martyrium des Hl. Sebastian* (67,4 x 67,7 cm), Foto: Artstor, © Musées Royaux des Beaux-Arts, Brüssel.).

**Abbildung 8 | S. 97**

Abstrakte Werkbeispiele mit großen Abweichungen in den durchschnittlichen Komplexitätseinschätzungen von Laien und Expert\_innen (von links: Kline, Franz (1956). *Mahoning* (204.2 × 255.3 cm), Foto: Artstor, © Whitney Museum of Art, New York; Marc, Franz (1914). *Kämpfende Formen* (91 x 131,5 cm), Foto: Artstor, © Pinakothek der Moderne, München.).

---

**Abbildung 9 | S. 101**

Standardabweichungen der durchschnittlichen Komplexitätsbewertung von Expert\_innen (geordnet nach der Höhe der Standardabweichungen), eigene Darstellung.

**Abbildung 10 | S. 102**

Standardabweichungen der durchschnittlichen Komplexitätsbewertung von Expert\_innen unterteilt in drei Abschnitte, eigene Darstellung.

**Abbildung 11 | S. 103**

Beispielfrage aus dem Fragebogen der Studie „*More Than the Sum of Its Parts. Perceiving Complexity in Painting*“, eigene Darstellung.

**Abbildung 12 | S. 106**

Standardabweichungen der durchschnittlichen Komplexitätsbewertung der *Expert\_innen* pro Bild (Bilder geordnet nach der durchschnittlichen Bewertung *aller* Befragten), eigene Darstellung.

**Abbildung 13 | S. 107**

Ernst, Max (1945). *Die Versuchung des Hl. Antonius* (108 x 128 cm), Foto: Artstor, © Wilhelm-Lehmbruck Museum, Duisburg.

**Abbildung 14 | S. 109**

Pollock, Jackson (1952). *Convergence* (241,9 x 399,1 cm), Foto: Artstor, © Albright-Knox Art Gallery.

**Abbildung 15 | S. 112**

Klimt, Gustav (1883). *Das Märchen* (85 x 116 cm), Foto: Artstor, © Wien Museum.

**Abbildung 16 | S. 128**

Bildmaterial der Studie „*Abstract Art as a Universal Language?*“ (**Bildpaar 1:** Kandinsky, Wassily (1935). *Kontrast mit Begleitung* (97,1 x 162,1 cm), Foto: Artstor, © Guggenheim Museum New York & Vallotton, Félix (1895). *Straßenszene*, Foto: Artstor, Privat; **Bildpaar 2:** Pollock, Jackson (1952). *Convergence* (241,9 x 399,1 cm), Foto: Artstor, © Albright-Knox Art Gallery & Delacroix, Eugène (1837). *The Battle of Taillebourg* (489 x 454 cm), Foto: Artstor, © Musée de l'Histoire de France (Versailles); **Bildpaar 3:** Kandinsky, Wassily (1926). *Several Circles* (140,4 x 140,7 cm), Foto: Artstor, © Guggenheim New York & Bosschert, Ambrosius (1606). *Flowers in a Glass* (35,6 x 29,3 cm), Foto: Artstor, © Cleveland Museum of Art; **Bildpaar 4:** Motherwell, Robert (1946-47). *Western Air* (182,9 x 137,2 cm), Foto: Artstor, © The Museum of Modern Art New York & Cezanne, Paul (1879-82). *Stilleben mit Karaffe, Milchkanne, Kaffeetasse und Orange* (26,67 x 34,92 cm), Foto: Artstor, © Dallas Museum of Art; **Bildpaar 5:** Götz, Karl Otto (1953). *Bild vom 5.2.1953* (125 x 90 cm), Foto: Artstor, © Staatliche Museen zu Berlin & Nolde, Emil (1910). *Tänzerin in rotem Kleid* (34,8 x 28,8 cm), Foto: Artstor, © Kunsthalle in Emden.



---

**Abbildung 17 | S. 156**

Bildmaterial der Studie „*Aesthetic Echoes in the Beholder’s Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art*” (**Kunstwerke Untersuchung 1**, von oben und links: Kandinsky, Wassily (1935). *Kontrast mit Begleitung*, Foto: Artstor, © Guggenheim Museum New York; Motherwell, Robert (1946-47). *Western Air* (182.9 x 137.2 cm), Foto: Artstor, © The Museum of Modern Art New York; Kandinsky, Wassily (1926). *Several Circles* (140.4 x 140.7 cm), Foto: Artstor, © Guggenheim New York; Pollock, Jackson (1952). *Convergence* (241.9 x 399.1 cm), Foto: Artstor, © Albright-Knox Art Gallery; Götz, Karl Otto (1953). *Bild vom 5.2.1953* (125 x 90 cm), Foto: Artstor, © Staatliche Museen zu Berlin. **Kunstwerke Untersuchung 2**, von oben und links: Pollock, Jackson (1950). *Number 3* (122 x 145 cm), Foto: Artstor, © Saint-Louis Art Museum; Motherwell, Robert (1969). *The Garden Window* (152.4 x 124.5), Foto: Artstor, © Modern Art Museum of Fort Worth; Francis, Sam (1958). *Untitled* (275.6 x 192.4 cm), Foto: Artstor, © Norton Simon Museum of Art; Motherwell, Robert (1941). *The Little Spanish Prison* (69.2 x 43.5 cm), Foto: Artstor, © Museum of Modern Art New York; Kandinsky, Wassily (1933). *Ausgleichrosa* (92 x 73 cm), Foto: Artstor, © Musée national d'Art moderne, Centre Georges Pompidou.).

**Abbildung 18 | S. 171**

Nachstellung Laborsituation im *Labor für empirische Bildwissenschaft* (CReA) am Institut für Kunstgeschichte der Universität Wien mit SMI SensoMotoric Instruments Eye-Tracker, Foto © Karl Pani und René Steyer.

**Abbildung 19 | S. 171**

Beispiel Labornachstellung für ein Eye-Tracking-Experiment mit Kinnstütze, Foto © <http://translationalneuromodeling.org/forschungsambulanz/eye-tracking/> (20.08.18).

**Abbildung 20 | S. 218**

links: Fra Angelico (1445). *Verkündigung* (230 x 321 cm), Foto: Artstor, © Museo di San Marco, Florenz; rechts: Lippi, Fra Filippo Tommaso (1450). *Verkündigung* (127 x 116 cm), Foto: Artstor, © Alte Pinakothek, München.

---

Ich habe mich bemüht, sämtliche Inhaber der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit eingeholt. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir.

---

# TABELLENVERZEICHNIS

---

**Tabelle 1 | S. 108**

Einzelauflistung der Komplexitätsbewertungen der zehn Expert\_innen für Max Ernsts *Die Versuchung des Hl. Antonius* (1945), eigene Darstellung.

**Tabelle 2 | S. 110**

Einzelauflistung der Komplexitätsbewertungen der zehn Expert\_innen für Jackson Pollocks *Convergence* (1952), eigene Darstellung.

**Tabelle 3 | S. 113**

Einzelauflistung der Komplexitätsbewertungen der zehn Expert\_innen für Gustav Klimts *Das Märchen* (nicht datiert), eigene Darstellung.

**Tabelle 4 | S. 129**

Begriffspaare semantisches Differential, eigene Darstellung.

---

# ANHANG

---

STELLUNGNAHME ZUR ARBEITSSTEILUNG I

*More Than the Sum of Its Parts. Perceiving Complexity in Painting*

Laura Commare/ Raphael Rosenberg/ Helmut Leder, *More Than the Sum of Its Parts. Perceiving Complexity in Painting*, in: *Psychology of Aesthetics, Creativity and the Arts*, 2018, Advance online publication: 10.1037/aca0000186.

Der Text mit dem Titel *More Than the Sum of Its Parts. Perceiving Complexity in Painting* ist in Zusammenarbeit mit Raphael Rosenberg und Helmut Leder entstanden und wurde 2018 in der Zeitschrift *Psychology of Aesthetic, Creativity and the Arts* (PACA) publiziert.

*Fragestellung*: Laura Commare, Raphael Rosenberg, Helmut Leder

*Theoretische Erarbeitung und Einbettung*: Laura Commare

*Studiendesign*: Laura Commare, Raphael Rosenberg, Helmut Leder

*Datenerhebung*: Laura Commare

*Datenauswertung*: Laura Commare

*Interpretation der Ergebnisse*: Laura Commare

*Text*: Laura Commare

---

STELLUNGNAHME ZUR ARBEITSSTEILUNG II

*Abstract Art as a Universal Language?*

Hanna Brinkmann/ Laura Commare/ Raphael Rosenberg/ Helmut Leder, *Abstract Art as a Universal Language?*, in: *Leonardo*, 47(3), 2014, S. 256-257.

Der Text mit dem Titel *Abstract Art as a Universal Language?* ist in Zusammenarbeit mit Hanna Brinkmann, Raphael Rosenberg und Helmut Leder entstanden und wurde 2014 in der Zeitschrift *Leonardo* publiziert.

*Fragestellung:* Laura Commare, Hanna Brinkmann, Raphael Rosenberg

*Theoretische Erarbeitung und Einbettung:* Laura Commare, Hanna Brinkmann

*Studiendesign:* Laura Commare, Hanna Brinkmann, Raphael Rosenberg, Helmut Leder

*Datenerhebung:* Laura Commare, Hanna Brinkmann

*Datenauswertung:* Laura Commare

*Interpretation der Ergebnisse:* Laura Commare, Hanna Brinkmann

*Text:* Laura Commare, Hanna Brinkmann

---

STELLUNGNAHME ZUR ARBEITSSTEILUNG III

*Aesthetic Echoes in the Beholder's Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art?*

Laura Commare/ Hanna Brinkmann, Aesthetic Echoes in the Beholder's Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art, in: Michael Zimmermann (Hg), Vision in Motion. Streams of Sensation and Configurations of Time, Berlin 2016, S. 221-234.

Der Text mit dem Titel *Aesthetic Echoes in the Beholder's Eye? Empirical Evidence for the Divergence of Theory and Practice in the Perception of Abstract Art* ist in Zusammenarbeit mit Hanna Brinkmann entstanden und wurde 2016 in dem, von Michael Zimmermann herausgegebenen, Sammelband *Vision in Motion. Streams of Sensation and Configurations of Time* publiziert.

*Fragestellung:* Laura Commare, Hanna Brinkmann

*Theoretische Erarbeitung und Einbettung:* Laura Commare, Hanna Brinkmann

*Studiendesign:* Laura Commare, Hanna Brinkmann

*Datenerhebung:* Laura Commare, Hanna Brinkmann

*Datenauswertung:* Laura Commare

*Interpretation der Ergebnisse:* Laura Commare, Hanna Brinkmann

*Text:* Laura Commare

---

## STELLUNGNAHME ZUR ARBEITSSTEILUNG IV

### *Warum ‚Anything goes‘ der Goldstandard sein sollte – Überlegungen zu Methodentradition und empirischen Forschungsansätzen in den Kunstwissenschaften*

Laura Commare/ Hanna Brinkmann, Warum ‚Anything goes‘ der Goldstandard sein sollte – Überlegungen zu Methodentradition und empirischen Forschungsansätzen in den Kunstwissenschaften, in: Julia Rüdiger, Anna Sauer und Christina Bartosch (Hg.), *Newest Art History – Wohin geht die jüngste Kunstgeschichte?* Tagungsband zur 18. Tagung des Verbandes österreichischer Kunsthistorikerinnen und Kunsthistoriker, Online 2017, S.161-180.

Der Text mit dem Titel *Warum ‚Anything goes‘ der Goldstandard sein sollte – Überlegungen zu Methodentradition und empirischen Forschungsansätzen in den Kunstwissenschaften* ist in Zusammenarbeit mit Hanna Brinkmann entstanden und wurde 2017 im Tagungsband zur 18. Tagung des Verbandes österreichischer Kunsthistorikerinnen und Kunsthistoriker *Newest Art History – Wohin geht die jüngste Kunstgeschichte?* publiziert. Es handelt sich bei dem Text um ein Positionspapier zum Gebrauch und Nutzen empirischer Methoden in den Kunstwissenschaften. Der Aufsatz ist unmittelbar aus der Arbeit an dieser Dissertationsschrift hervorgegangen. Der zweite Teil des Aufsatzes bildet in leicht abgewandelter Form das Kapitel 1.2 der vorliegenden Arbeit (*Allgemeine Klassifikationsmerkmale empirischer Ansätze*).

*Problemstellung:* Laura Commare, Hanna Brinkmann

*Theoretische Erarbeitung und Einbettung Teil I:* Laura Commare, Hanna Brinkmann

*Theoretische Erarbeitung und Einbettung Teil II:* Laura Commare

*Text:* Laura Commare



# EXPERTISEFRAGEBOGEN



## Fragebogen Kunstinteresse

Wir möchten in unseren Studien die individuellen ästhetischen Vorlieben und den Prozess des ästhetischen Erlebens untersuchen. Dabei ist uns bewusst, dass ästhetische Vorlieben und Einschätzungen wesentlich vom Interesse und Wissen über Kunst, Design, etc. beeinflusst werden. Deshalb bitten wir Sie, diesen Fragebogen auszufüllen.

Das Ausfüllen dieses Fragebogens wird ca. 10 Minuten dauern. Ihre Angaben und persönlichen Daten werden vertraulich behandelt und nur zu wissenschaftlichen Zwecken verwendet.

Vielen Dank für Ihre Teilnahme!

### Teil A

#### Wie sehr interessieren Sie sich für Kunst?

- überhaupt nicht interessiert     sehr wenig interessiert     wenig interessiert     mäßig interessiert     eher interessiert     interessiert     sehr interessiert

#### Wie oft gehen Sie im Durchschnitt in Kunstmuseen/Kunstgalerien?

- eigentlich nie     einmal im Jahr     einmal im Halbjahr     einmal in drei Monaten     einmal im Monat     einmal in zwei Wochen     einmal in der Woche

#### Wie häufig lesen Sie Bücher, Zeitschriften oder Kataloge über Kunst?

- eigentlich nie     sehr selten     eher selten     manchmal     eher häufig     häufig     sehr häufig

#### Wie häufig sehen Sie sich Abbildung von Kunstwerken an (Kunstabildbände, Internet, etc.)?

- eigentlich nie     sehr selten     eher selten     manchmal     eher häufig     häufig     sehr häufig

#### Wie häufig sprechen Sie mit Freunden/Bekanntem/Verwandten über Kunst?

- eigentlich nie     sehr selten     eher selten     manchmal     eher häufig     häufig     sehr häufig

---

**Teil B**

Im nächsten Abschnitt finden Sie eine Tabelle mit Namen von Künstlern. Geben Sie bitte an, ob Ihnen der Name des Künstlers bekannt ist. Falls möglich geben Sie bitte auch an, mit welcher Stilrichtung der Künstler hauptsächlich in Verbindung gebracht wird. Falls nicht möglich, markieren Sie bitte die entsprechende Zelle mit einem „X“ (Kreuz).

Künstlerna <del>m</del> e	Bekannt	Kunstrichtung/Stil
Henri Matisse	<input type="checkbox"/> ja <input type="checkbox"/> nein	
Salvador Dalí	<input type="checkbox"/> ja <input type="checkbox"/> nein	
Pablo Picasso	<input type="checkbox"/> ja <input type="checkbox"/> nein	
Roy Lichtenstein	<input type="checkbox"/> ja <input type="checkbox"/> nein	
Pieter Aertsen	<input type="checkbox"/> ja <input type="checkbox"/> nein	
Tizian	<input type="checkbox"/> ja <input type="checkbox"/> nein	
Sandro Botticelli	<input type="checkbox"/> ja <input type="checkbox"/> nein	
Duccio di Buoninsegna	<input type="checkbox"/> ja <input type="checkbox"/> nein	
Jackson Pollock	<input type="checkbox"/> ja <input type="checkbox"/> nein	
Neo Rauch	<input type="checkbox"/> ja <input type="checkbox"/> nein	

---

**Teil C**

Anschließend sehen Sie sich bitte der Reihe nach die Abbildungen in der beigelegten Mappe an. Geben Sie zunächst an, ob Ihnen das jeweilige abgebildete Kunstwerk bekannt ist. Falls möglich geben Sie bitte auch an, von welchem Künstler/welcher Künstlerin das jeweilige Kunstwerk stammt und mit welcher Stilrichtung das Kunstwerk hauptsächlich in Verbindung gebracht wird. Zusätzlich bitten wir Sie (falls möglich) um eine kurze ikonographische/inhaltliche Analyse (in Stichworten). Falls nicht möglich, markieren Sie bitte die entsprechende Zelle mit einem „X“ (Kreuz).

	<b>KünstlerIn:</b>
<b>Nr. 1</b>	
<b>Bekannt:</b> <input type="checkbox"/> ja <input type="checkbox"/> nein	<b>Kunstrichtung/Stil:</b>
	<b>Ikonographie:</b>
	<b>KünstlerIn:</b>
<b>Nr. 2</b>	
<b>Bekannt:</b> <input type="checkbox"/> ja <input type="checkbox"/> nein	<b>Kunstrichtung/Stil:</b>
	<b>Ikonographie:</b>
	<b>KünstlerIn:</b>
<b>Nr. 3</b>	
<b>Bekannt:</b> <input type="checkbox"/> ja <input type="checkbox"/> nein	<b>Kunstrichtung/Stil:</b>
	<b>Ikonographie:</b>
	<b>KünstlerIn:</b>
<b>Nr. 4</b>	
<b>Bekannt:</b> <input type="checkbox"/> ja <input type="checkbox"/> nein	<b>Kunstrichtung/Stil:</b>
	<b>Ikonographie:</b>

<b>Nr. 5</b>	<b>KünstlerIn:</b>
<b>Bekannt:</b> <input type="checkbox"/> ja <input type="checkbox"/> nein	<b>Kunstrichtung/Stil:</b>
	<b>Ikonographie:</b>
<b>Nr. 6</b>	<b>KünstlerIn:</b>
<b>Bekannt:</b> <input type="checkbox"/> ja <input type="checkbox"/> nein	<b>Kunstrichtung/Stil:</b>
	<b>Ikonographie:</b>

---

#### Teil D

Zum Abschluss bitten wir Sie noch ein paar Fragen zu Ihrer Person zu beantworten.

**Wie alt sind Sie?** \_\_\_\_\_  Männlich  
 Weiblich

**Falls zutreffend, bis zu welchem Abschluss haben Sie Kunstgeschichte studiert?**

- Trifft nicht zu
- (noch) ohne Abschluss
- Bachelor
- Master/Magister
- Promotion
- Habilitation

**Gibt es Epochen innerhalb der Bildenden Kunst, die Sie besonders interessieren bzw. mit welchen Sie am meisten Erfahrung haben? (Mehrfachantworten möglich)**

- Nein, gibt es nicht
- Steinzeit/Vorgeschichte
- Antike
- Mittelalter
- Renaissance/Manierismus
- Barock/Rokoko
- 19. Jahrhundert
- Moderne
- Gegenwartskunst
- Außereuropäische Kunst

**Gibt es eine bestimmte Kunstgattung die Sie mehr interessiert als andere?**

- Nein, gibt es nicht
- Zeichnung
- Malerei
- Skulptur
- Architektur
- Filmfotografie
- Andere

---

**Vielen Dank!**

---

## ABSTRACT (D)

### „Das Ganze ist mehr als die Summe seiner Teile“

#### Chancen und Herausforderungen empirischer Zugänge zu den Kunstwissenschaften

Im Verlauf des letzten Jahrzehnts konnte eine verstärkte Zuwendung zu empirischen methodischen Ansätzen in der kunstwissenschaftlichen Forschung verzeichnet werden. Diese Ansätze werden jedoch keineswegs ausschließlich positiv bewertet. Häufig stehen sie großer Skepsis gegenüber und werden als fachfremde Über- und Eingriffe seitens der Natur- und Sozialwissenschaften in den geisteswissenschaftlichen Hegemonialbereich verstanden. Die Möglichkeiten empirischer Forschung, geisteswissenschaftliche Konzepte hinreichend abzubilden werden vielfach in Zweifel gezogen; wodurch gleichzeitig die Sinnhaftigkeit einer empirisch-arbeitenden Kunstwissenschaft verworfen wird.

Ziel dieser Arbeit ist es, einen begründeten Antwortversuch auf die Frage nach dem Wert und den Voraussetzungen empirischer Ansätze für die Kunstwissenschaften zu entwerfen. Im Konkreten werden hierzu zunächst die gegenüber empirischen Ansätzen geäußerten positiven und negativen Kritikpunkte herausgearbeitet und kondensiert, um sie anschließend im Einzelnen einer Prüfung an konkreten Fallbeispielen zu unterziehen. Bei den zugrunde gelegten Fallbeispielen handelt es sich um drei empirische Studien, die der Rezeptionsästhetik zuzuordnen sind und somit im weiteren Sinne in den Bereich der empirischen Bildwissenschaft fallen. Am Beispiel dieser Studien werden argumentativ Möglichkeiten und Grenzen der jeweils gewählten empirischen Zugänge diskutiert. Die Diskussion ist dabei in zwei Schwerpunkte unterteilt: eine inhaltliche Diskussion der – mit empirischen Mitteln erzielten – Ergebnisse sowie eine methodische Diskussion, welche die jeweiligen Vorgehensweisen kritisch beleuchtet.

---

ABSTRACT (E)

“More than the Sum of Its Parts”

Opportunities and Challenges of Empirical Approaches to the Art Sciences

Over the last decade we have been observing a new trend in the humanities and art sciences towards the systematic use of empirical methods. Yet, empirical approaches to the art sciences are not without critique. Empirical studies are often faced with skepticism and seen as outside-subject interference of the natural and social sciences. The possibilities of empirical research to adequately address holistic concepts are doubted; with the consequence that empirical approaches to questions of art are frequently granted no sense of purpose.

The aim of this thesis is to draft a reasoned response to the question on the value and conditions of an empirical art science. To this end, I discuss chances as well as challenges that may be attributed to empirical approaches and investigate their practical relevance by reverting to three specific case studies pertaining to the field of (empirical) reception research. The discussion follows two focal points: a content discussion of the results obtained by empirical means, and a methodical discussion, which critically examines the respective approaches of the selected case studies.