
3	Gehirn, Psychologie und Recht	87
3.1	Willensfreiheit, Verursachung und Verantwortlichkeit	91
3.1.1	Kognitive und Neuropsychologie	93
3.1.2	Pragmatismus in Psychologie und Recht	95
3.1.3	Willensfreiheit im 19. Jahrhundert	98
3.1.4	Willensfreiheit im 20. und 21. Jahrhundert	101
3.1.5	Determination und Verantwortung	107
3.1.6	Beispiele aus dem Recht	109
3.2	Berühmte (un)verantwortliche Gehirne	112
3.2.1	Der berühmteste Patient	116
3.2.2	Phrenologen im Gerichtssaal	118
3.2.3	Gehirnschäden und Verbrechen	122
3.2.4	Exkurs: Rechtswissenschaftliche Rezeption	124
3.2.5	Exkurs: Krankheit und Schuld	127
3.2.6	Keine Eins-zu-eins-Determination	132
3.2.7	Normales Beweismittel vor Gericht	133
3.3	Zusammenfassung	134

Kategorie der Verantwortlichkeit dazu verhält. Dieses Wissen wird in Abschnitt 3.2 aufgegriffen, wo wir neurologische und rechtliche Aspekte von bekannten Fällen von Gehirnverletzungen diskutieren. Nach einer Zusammenfassung in Abschnitt 3.3 machen wir dann in Kapitel 4 mit anwendungsnahen normativen Fragen speziell zur Gehirnentwicklung weiter. Damit liegt der Schwerpunkt in diesem Kapitel auf der zweiten Kategorie von Chandlers Systematisierung. Danach wird die erste eine stärkere Rolle spielen. Außerdem verfolgt mein Ansatz eine Würdigung der psychologischen und neurowissenschaftlichen Kenntnisse im Sinne eines *kritischen* Neurorechts.

3.1 Willensfreiheit, Verursachung und Verantwortlichkeit

Die Frage, wie wir unser Verhalten erklären, zieht sich wie kaum eine andere durch die westliche Philosophiegeschichte. Schon in Platons (428/427–348/347 v. u. Z.) Dialog *Phaidon* wurde sie diskutiert. In diesem sprach Sokrates (469–399 v. u. Z.) kurz vor seinem Tod – er musste wegen angeblicher Blasphemie und Aufwiegelung der Jugend entweder in die Verbannung gehen oder aus dem Schierlingsbecher trinken – zum letzten Mal mit seinen Schüler*innen. Dort behandelte er die Sichtweise seines früheren Lehrers Anaxagoras (ca. 499–428 v. u. Z.), ein materialistisch denkender Naturphilosoph: Wie viel Sinn ergibt auf die Frage, *warum* Sokrates im Gefängnis sei, die Antwort: „weil die Sehnen und Knochen sich so und so bewegt haben“? Aus heutiger Sicht würden wir das „Reduktionismus“ nennen und dem „Handeln aus Gründen“ gegenüberstellen (Dretske, 1988; Schleim, 2023c). Gemäß Letzterem ist Sokrates im Gefängnis, weil er dazu verurteilt wurde und sich der Strafe stellte.

Die primäre Wissenschaft vom Verhalten ist traditionell erst einmal die Psychologie. Eine ihrer bedeutendsten Schulen im 20. Jahrhundert, der Behaviorismus, trug diesen Anspruch schon im Namen. Was später und bis heute in der Philosophie als „eliminativer Materialismus“ diskutiert wurde (zum Beispiel Churchland, 1981),¹⁴ hatten führende Behavioristen wie John B. Watson (1878–1958) oder Burrhus F. Skinner (1904–1990) tatsächlich schon um Jahrzehnte vorweggenommen: Diese For-

14 *Eliminativer Materialismus* wird heute meist als die philosophische Position verstanden, dass das Vokabular unserer Alltags- oder sogar der kognitiven Psychologie traditionell entstanden ist, aber keine Entsprechung in der Wirklichkeit hat; dieses Vokabular würde durch neurowissenschaftliche Durchbrüche in der Zukunft ersetzt und in diesem Sinne „eliminiert“. Paul M. Churchland erinnerte mich 2014 im persönlichen Gespräch daran, dass er die These ursprünglich für *intentionale* Zustände aufstellte. Das sind Zustände, die sich auf etwas in der Welt beziehen, wie Absichten, Überzeugungen, Wünsche und so weiter. Da diese für die rechtlichen Fragen zentral sind, muss hier auf die alternativen subjektiven Erfahrungszustände (phänomenales Erleben) nicht näher eingegangen werden.

scher*innen wollten nämlich alles, was sich nicht objektivieren lässt, aus der wissenschaftlichen Psychologie ausschließen. Und das waren insbesondere die vermuteten *internen* psychologischen Vorgänge. Introspektion, den Versuch, die Psyche von innen heraus zu erklären, lehnte Watson als „mentale Gymnastik“ vollständig ab; auch Bewusstsein hatte demnach keinen Platz in der Wissenschaft (Watson, 1913/1994).

Skinner hielt es vor allem für problematisch, Verhalten mit psychologischen Konstrukten kausal zu erklären (Skinner, 1953). Das wären beispielsweise Sätze wie: „Der Autofahrer ignorierte das Stoppschild, weil er vor Anfang des Fußballspiels zu Hause sein *wollte*“, „Die Frau stand mit der *Absicht* in der Warteschlange, als eine der Ersten eine Konzertkarte zu bekommen“, oder „Kim blieb wegen der *Angst* im Haus“. Dass es in dieser Denkwelt keinen Platz für Willensfreiheit und andere zentrale Begriffe für unser Menschenbild gab, drückte Skinner schon im Titel seiner damals viel gelesenen Aufsatzsammlung aus: *Beyond Freedom and Dignity* (Skinner, 1971).

Anstatt aus internen Vorgängen wollten Behavioristen Verhalten aus der Konditionierung durch Eigenschaften von Belohnung und Bestrafung in der Umgebung erklären, den sogenannten positiven beziehungsweise negativen Verstärkern. Diese Eigenschaften ließen sich ihrer Meinung nach objektiv beschreiben und messen, eben durch Manipulation der Umwelt und Beobachtung des Verhaltens. Bis heute besteht diese Denkschule beispielsweise im *Journal of the Experimental Analysis of Behavior* und anderen Fachzeitschriften fort (siehe auch Reinecker, 2015; Vargas, 2020). Allerdings hat dieser Ansatz – abgesehen von seiner Kontraintuitivität und Unvereinbarkeit mit dem vorherrschenden Menschenbild – auch grundlegende theoretische Probleme:

So kann es beim Menschen aus praktischen Gründen keine hinreichende Kontrolle über die Vorbedingungen geben, um in einem starken Sinn wie etwa in der Physik Vorhersagen von Verhalten aus allgemeinen Gesetzen abzuleiten. Die Geschichte unserer Konditionierungen ist auf unseren individuellen Lebenswegen einfach zu komplex. Und der für die Behavioristen zentrale Begriff der Verstärkung wurde für seine Zirkularität kritisiert: nämlich als das, was ein bestimmtes Verhalten wahrscheinlicher macht. Die Erklärung, was das Verhalten wahrscheinlicher macht, ist aber wiederum der Verstärkungseffekt (Westmeyer, 1973, S. 53f.; aber siehe Moore, 2008, S. 123f.).

Diese Gedanken zum Behaviorismus sind nicht als Lektion in Psychologiegeschichte vorangestellt. Vielmehr zeigt sich hier eine frappierende Analogie zur neueren Herausforderung des Rechts durch die Neurowissenschaften: Diese besteht in der Willensfreiheitsdebatte – wie wir gleich vertiefen werden – in der Ablehnung der etablierten Konzepte von Verantwortung und Schuld. Behaviorismus und eliminativer Materialismus taten das vor Jahrzehnten schon viel grundlegender, nämlich durch die Darstellung verbreiteter Annahmen über unser psychologisches Innenleben als eine Art Fiktion. Für Sachverhalte des Rechts sind insbesondere *Absichten*

von entscheidender Bedeutung: Wurde ein Mensch zum Beispiel fahrlässig getötet, weil Sicherheitsmaßnahmen vernachlässigt wurden, handelte es sich um eine Notwehrsituation, war es eine Tötung aus einem spontan entstandenen Konflikt heraus oder ein geplanter Mord „aus niedrigen Beweggründen“ (§ 211 Abs. 2 StGB), wo die *Gründe* sogar im Gesetzestext stehen?

Insofern ist die kognitive Psychologie, mit der wir uns im Folgenden beschäftigen, prinzipiell näher an den etablierten rechtlichen Kategorien als der Behaviorismus. Dennoch hat die Psychologie das hier angerissene Problem ihrer Gegenstandsfrage in über 100 Jahren nicht überwinden können (vgl. Anderson, 2015; Galliker, 2016, Kapitel 7.3), was sich eines Tages auf andere Bereiche einschließlich des Rechts auswirken könnte. Demgegenüber wurden Behavioristen wie Skinner nicht müde, die Bedeutung der Untersuchung *echten* Verhaltens in *echten* Situationen und nicht nur in Laborversuchen zu betonen (Moore, 2008, S. 352). Wir werden jetzt sehen, wie weit wir mit der Annahme innerer psychologischer Vorgänge kommen.

3.1.1 Kognitive und Neuropsychologie

Bekanntermaßen wurde der Behaviorismus durch die „kognitive Revolution“ – eigentlich war es eher ein „kognitives Wiederaufleben“ – in der Psychologie abgelöst (vgl. Fancher & Rutherford, 2017; Wertheimer & Puente, 2020). Warum sich *unterschiedliche* Personen in *derselben* Umgebung unterschiedlich verhalten, wollte man schließlich doch unter Rückgriff auf unser psychologisches „Innenleben“ erklären. Die in den kognitiven Modellen mit Pfeilen verbundenen Kästen, die für diese Vorgänge stehen, handelten dieser Strömung mitunter den Namen „Boxologie“ ein. Da das für sich noch keine ontologische Fundierung bot, also keine grundlegende Antwort auf die Frage, *was psychologische Prozesse eigentlich sind*, orientierten sich viele Psycholog*innen im 20. Jahrhundert an der Biologie: Daraus entstanden biologische, Evolutions- und Neuropsychologie.

Wir sahen allerdings schon in Kapitel 1.1, dass die Biologie der Neuropsychiatrie seit über 200 Jahren nicht das gewünschte ontologische Fundament beschert hat: Bis heute gibt es zum Beispiel keine zuverlässigen diagnostischen Biomarker – und zwar für *keine einzige* der mehreren Hundert psychischen Störungen des Diagnosehandbuchs DSM (APA, 2022). Nun kann man argumentieren, dass die Entitäten der Psychiatrie – eben die klassifizierten Störungen – abstrakte diagnostische Konstrukte sind, die sich aus einfacheren psychischen Vorgängen zusammensetzen, und dass es für diese einfacheren Vorgänge eine ontologische Basis in der Neurobiologie gibt (vgl. Schleim, 2022a, 2023a).

Dem ist aber nicht so, jedenfalls nicht für die abstrakteren kognitiven, emotionalen und sozialen Vorgänge, die wir ab Kapitel 2 auf rechtliche Fragen bezogen. Dass sich psychische Konstrukte nicht einfach in physiologischen Prozessen fundieren lassen, stellt die Neuropsychologie schon seit über 100 Jahren vor Herausforderungen (Anderson, 2015; Cacioppo & Tassinari, 1990). Erst kürzlich zogen erfahrene Experimentalpsychologen den Schluss: „Niemand weiß, was Aufmerksamkeit ist“ (Hommel et al., 2019; dt. Übers.). Im Vergleich hiermit ist Watsons behavioristische Kritik an der Bewusstseinsforschung am Anfang des 20. Jahrhunderts interessant. Dazu ein Originalzitat:

„Es scheint die Zeit gekommen zu sein, in der die Psychologie jeden Bezug auf das Bewusstsein aufgeben muss; in der sie sich nicht länger vormachen muss, dass sie mentale Zustände zum Gegenstand der Beobachtung macht. Wir haben uns so sehr in spekulative Fragen über die Elemente des Geistes, die Natur bewusster Inhalte [...] verstrickt, dass ich als experimenteller Forscher das Gefühl habe, dass mit unseren Prämissen und den Arten von Problemen, die sich daraus entwickeln, etwas nicht stimmt. Es gibt keine Garantie mehr, dass wir alle dasselbe meinen, wenn wir die heute in der Psychologie gebräuchlichen Begriffe verwenden.“

(Watson, 1913/1994, S. 249; dt. Übers.)

Genau 110 Jahre später brach in der heute zweifellos etablierten Bewusstseinsforschung ein vehementer Streit darüber aus, ob eine der führenden Bewusstseinstheorien „Pseudowissenschaft“ ist (Lenharo, 2023a, 2024). Die Details dieser Debatte führen hier zu weit. Allerdings können wir feststellen, dass führende Forscher*innen auf diesem Gebiet rätseln, inwiefern Bewusstsein auf der einen Seite von Aufmerksamkeit und auf der anderen vom Gedächtnis abzugrenzen ist (Koch, Massimini, Boly & Tononi, 2016). Andere Neurowissenschaftler*innen verglichen die gängigen Theorien des Bewusstseins und kamen zum Schluss, dass diese vielleicht nicht einmal den Status einer „Theorie“ verdienen (Signorelli, Szczotka & Prentner, 2021). Zudem könne man nicht beurteilen, welcher dieser Ansätze Bewusstsein am besten erklärt, weil keine Einigkeit darüber herrsche, was so eine Erklärung leisten solle beziehungsweise was „erklären“ hier bedeute.

Andere erfahrene Bewusstseinsforscher unterschieden sechs verschiedene *Erklärungsziele* – wie zum Beispiel Inhalt des Bewusstseins, Emotion, phänomenale Eigenschaften oder Selbstbewusstsein – und acht verschiedene *Funktionen* – wie zum Beispiel Aufmerksamkeit, Metakognition, Selbst oder Arbeitsgedächtnis (Northoff & Lamme, 2020). Sie diskutierten, wie sich acht zurzeit verfolgte Ansätze in diesen Aspekten voneinander unterscheiden. Und solche Differenzen sind nicht nur theoretischer Natur: In Abhängigkeit von ihren Annahmen fanden Forscher ganz andere Gehirnregionen, die für Bewusstsein zentral sein sollen, nämlich die frontalen

Regionen auf der einen Seite gegenüber parietal-temporalen Netzwerken auf der anderen (Koch et al., 2016; Northoff & Lamme, 2020).

Wenn man Watsons Forderung von 1913 zum Maßstab nimmt, dass Forscher*innen sich über die Bedeutung zentraler Begriffe einig sein müssen, sieht es für Teile von Psychologie und kognitiver Neurowissenschaft problematisch aus. Tatsächlich wird die fehlende begriffliche und theoretische Integration für die anhaltende Krise der Psychologie mitverantwortlich gemacht (Eronen, 2024; Groeben & Westmeyer, 1981; Hutmacher & Franz, 2025).

Die grundlegende Frage, mit welchen Entitäten es die Disziplin überhaupt zu tun hat, gilt nicht nur für die Psychiatrie (Hyman, 2021; Kendler, Zachar & Craver, 2011; Schleim, 2022a; Vintiadis, 2024). Sie gilt ebenso für die Psychologie. Und für Letztere sollte sie eher zu klären sein, da beispielsweise die zentrale Definition psychologisch-psychiatrischer Störungen aus dem Diagnosehandbuch DSM auf Störungen der Kognition, Emotionsregulation und des Verhaltens verweist, die eine Dysfunktion psychologischer oder biologischer Prozesse widerspiegelt (APA, 2022). Demnach baut die Psychiatrie auf der Psychologie auf – nicht umgekehrt. Angesichts der anhaltenden Krise in Psychologie und Psychiatrie erfährt zurzeit nicht nur der eingangs erwähnte Ansatz der „4E Cognition“ mehr Aufmerksamkeit. Auch in der etwas in Vergessenheit geratenen phänomenologischen Psychologie versucht man beispielsweise mit neuen Vorschlägen, die Disziplin gewissermaßen aus sich selbst heraus zu begründen (zum Beispiel Wendt, 2024).

Weder können noch müssen wir dieses Problem hier lösen. Schließlich hat man in praktischen Kontexten wie dem Recht, der Psychotherapie oder Psychiatrie, anders als vielleicht in wissenschaftlicher Forschung und Philosophie, nicht ewig Zeit für das Finden von Antworten. Wie es schon in der am Anfang des Buches zitierten und bis heute für die Verwendung wissenschaftlichen Wissens vor Gericht maßgeblichen Entscheidung des U.S. Supreme Court aus dem Jahr 1993, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, hieß: „Wissenschaftliche Schlussfolgerungen sind einer ständigen Überprüfung unterworfen. Das Recht hingegen muss Streitigkeiten endgültig und schnell lösen.“ Es ist also kein Zufall, dass das Recht *pragmatische* Kategorien verwendet.

3.1.2 Pragmatismus in Psychologie und Recht

Ich habe an anderer Stelle dafür argumentiert, dass Ansätze zum Beschreiben und Verstehen des Menschen drei Bedingungen erfüllen müssen: Sie müssen nützlich, kohärent und sinnvoll sein (Schleim, 2023c, S. 40f.). *Nützlichkeit* bedeutet, dass sie sich in der Praxis bewähren; *Kohärenz* bedeutet, dass ihre Bestandteile so gut wie mög-

lich zusammenpassen und sich so wenig wie möglich widersprechen; und *Sinn* bedeutet, dass sie sich in unsere kulturelle Tradition und Sprachpraxis einbetten lassen.

Vor diesem Hintergrund stellt sich unmittelbar die Frage, ob sich kontraintuitive Ansätze wie der Behaviorismus, eliminative Materialismus oder auch ein naturalistischer Reduktionismus überhaupt konsistent vertreten lassen: Schließlich beinhalten alle diese Positionen, dass unsere Sprachpraxis grundlegend falsch ist. Für den Behavioristen sind die Eigenschaften positiver und negativer Verstärkung in der Umgebung entscheidend; für materialistische beziehungsweise reduktionistische Positionen ist die Ebene der Physik oder wenigstens der Neurobiologie das Fundament. In den Sprachen dieser Disziplinen gibt es aber keine Definition von Gehalt, Sinn und Wahrheit.

Streng genommen, doch aus prinzipiellen Gründen, kann somit kein Vertreter dieser radikalen Sichtweisen gehalt- und sinnvolle wahre Sätze äußern, ohne sich selbst zu widersprechen. Dieses grundlegende theoretische Problem passt interessanterweise dazu, dass sich die alltagspsychologische Sichtweise in Recht und Moral seit Jahrhunderten hartnäckig hält. Das werden auch die in diesem Kapitel eingehender analysierten Fälle deutlicher zeigen.

Hierfür ist insbesondere die Willensfreiheitsdebatte ein so gutes wie einschlägiges historisches und aktuelles Beispiel. Schließlich haben Vertreter*innen der eliminativen oder reduktionistischen Sicht immer wieder unser gängiges, pragmatisches und normatives Menschenbild für wissenschaftlich widerlegt erklärt. Einige leiteten daraus gleich eine Revolution des Strafrechts ab. Natürlich steht es auch ihnen frei, einen alternativen Ansatz zu vertreten, der nützlich, kohärent und sinnvoll ist. Dafür reicht es aber nicht, einfach Selbstbeschreibungen als „mentale Gymnastik“ (zum Beispiel Watson, 1913/1994), psychische Vorgänge als „nichts als Gehirnprozesse“ (zum Beispiel Crick, 1994) oder bewusste Entscheidungen als „Illusion“ (zum Beispiel Wegner, 2002) abzutun.

Anders als oft behauptet, ist ein „freier Wille“, was auch immer das sein mag, weder ein Grundpfeiler unseres Rechts noch unserer Moral. Diese sehen uns im Allgemeinen als *minimal rationale Subjekte* an: Das sind Personen, die, erstens, in der Regel Richtig und Falsch im Sinne der herrschenden normativen Ordnung voneinander unterscheiden können und, zweitens, über die Selbstkontrolle verfügen, nach dieser Einsicht zu handeln. Daraus ergibt sich unsere Verantwortlichkeit, sofern nicht ein bestimmter Ausnahmezustand vorliegt (Bigenwald & Chambon, 2019; Morse, 2007, 2023; Penney, 2012). Moralphilosoph*innen sprechen in ähnlicher Weise davon, dass Menschen *ansprechbar für Gründe* sind (zum Beispiel Hirstein et al., 2018). Die nach deutschem Strafrecht relevanten Kategorien sind die Einsichts- und Steuerungsfähigkeit (§§ 17, 20 StGB).

Wer aber zum Beispiel eine schwere geistige Behinderung hat, eine starke Psychose oder einen intensiven Drogenrausch erfährt, schlafwandelt oder mit hinrei-

chendem Nachdruck zu etwas gezwungen wird, wird in der Regel für in solchen Situationen begangene Taten *nicht* verantwortlich gemacht. Wenn eine Person jedoch eine große Gefahr für andere darstellt, kennen Rechtsordnungen meist eine Möglichkeit als Alternative zur Bestrafung. Dann werden Menschen beispielsweise psychiatrisch untergebracht, bis die Gefährlichkeit gebannt ist. Denn es gilt: Ohne Verantwortung keine Schuld und ohne Schuld keine Strafe – während Rechtsstaaten verhältnismäßige Schritte zur Eindämmung von Gefahren für Leib und Leben unternehmen müssen. Übrigens kann nicht nur Handeln, sondern auch Unterlassen im genannten Sinn moralisch und rechtlich relevant sein. Das wäre etwa dann der Fall, wenn jemand einer anderen Person in Not nicht hilft, obwohl er es könnte und sich der Notlage bewusst ist.

In diesem normativen System ist ein „freier Wille“ kein notwendiger Baustein. Das kann man an den beiden vorherigen Absätzen erkennen.

Was oben „minimal rationale Subjekte“ genannte wurde, bezeichneten Claus Roxin (1931–2025) und Luís Greco in ihrem maßgeblichen Kommentar zum deutschen Strafrecht als „normative Ansprechbarkeit“ eines Täters (Roxin & Greco, 2020, § 20 Rn 28). Hier geht es um den gerade schon angeführten § 20 StGB zur Schuldunfähigkeit wegen seelischer Störungen. Im Einklang mit der im vorliegenden Buch vertretenen Sichtweise hielten die beiden Strafrechtler es bei der Beurteilung der Einsichts- und Steuerungsfähigkeit durch Sachverständige für ein „Missverständnis, dass hier eine Äußerung zur Willensfreiheit [...] verlangt werde. In Wirklichkeit geht es aber nicht darum, sondern nur um ein Urteil darüber, ob und ggf. inwieweit der Täter im Tatzeitpunkt normativ ansprechbar war“ (Roxin & Greco, 2020, § 20 Rn 28). Und bei der Diskussion des Schuldbegriffs führten sie aus:

„Empirisch feststellbar ist die prinzipielle Fähigkeit zur Selbststeuerung und damit die gegebene normative Ansprechbarkeit (damit wird nicht geleugnet, dass im Grenzbereich dieser Ansprechbarkeit Wertungen einfließen, doch ist dies bei allen Rechtsbegriffen so). Normativ zugeschrieben wird dagegen [...] die aus diesem Befund abgeleitete Möglichkeit zu rechtmäßigem Verhalten.“

(Roxin & Greco, 2020, § 19 Rn 46)

Das Recht auf der einen Seite und Psychologie, Psychiatrie und Neurowissenschaften auf der anderen können in diesem Sinne zusammenarbeiten und müssen sich nicht widersprechen. Um diese Brücke von der empirischen auf die normative Ebene zu bauen, ziehe ich es vor, von *Willensprozessen* zu sprechen (Schleim, 2023c). Nach meinem Dafürhalten gibt es nämlich kein Ding „Wille“, das frei wäre oder nicht; auch heute noch tun sich Philosophen schwer damit, zu erklären, was ein Wille überhaupt sein soll (zum Beispiel Hieronymi, 2022). Mit anderen Worten und in Erinnerung an Kapitel 2.4: Ich will „Wille“ nicht verdinglichen. Stattdessen spreche ich von Willens-

vorgängen, die typischerweise durch bestimmte Merkmale charakterisiert sind, zum Beispiel Prozesse des Planens und Abwägens in einem bedeutungsvollen Kontext unter Berücksichtigung der eigenen Präferenzen und Wünsche. Wir werden gleich sehen, dass diese Sichtweise empirisch gestützt werden kann und Willensfreiheit dann keine Frage des „Alles oder Nichts“ ist, sondern eines „Mehr oder Weniger“.

Damit halte ich meine Ausführungen für mit dem in der Psychologie verbreiteten *Rubikonmodell der Handlungsphasen* von Heinz Heckhausen (1926–1988) und Peter M. Gollwitzer vereinbar. Die vier unterschiedenen Phasen sind das Abwägen, Planen, Durchführen und abschließende Bewerten. Bezeichnenderweise kommt im hierfür klassischen Zitationsort ein substantivierter Wille gar nicht vor, außer im (veralteten) Namen der Willenspsychologie (Heckhausen & Gollwitzer, 1987, S. 102). Im Vorwort des Tagungsbands anlässlich Heckhausens 60. Geburtstags heben die Autoren nicht nur die besondere Bedeutung selbstbestimmten Handelns und des Vorausplanens von Handlungsschritten hervor. Vielmehr teilen sie auch eine Beobachtung, die mit Blick auf die 40 folgenden Jahre der Willensfreiheitsdiskussion bezeichnend ist:

„In den einzelnen Abschnitten des Buches zeigen [die Autor*innen] nicht nur auf, daß auch in ihren Disziplinen der Wille als Begriff an wissenschaftlicher Reputation verlor und außer Gebrauch geriet. Es wird auch deutlich, wo und wie die Willensproblematik, und sei es untergründig, dennoch wirksam blieb oder erneut aufgegriffen wurde. Während z.B. ein so klassisches Feld wie die ‚Willenserziehung‘ fast ganz aus der Pädagogik verschwand, glaubt neuerdings ein scheinbar so weit abliegendes Fach wie die Hirnphysiologie, Willensphänomene und deren neuroanatomische Verankerung dingfest gemacht zu haben.“

(Heckhausen, Gollwitzer & Weinert, 1987, S. vii)

Um den sehr schön formulierten Gedanken des „neuroanatomischen Dingfestmachens“ drehen sich viele der neurowissenschaftlichen Themen hier im Buch. Für das Beispiel Willensfreiheit werden diese Bemühungen mit Blick auf die aktuelle Forschung in Kapitel 3.1.4 noch einmal aus Sicht des Rubikonmodells bewertet.

3.1.3 Willensfreiheit im 19. Jahrhundert

Wie wir oben sahen, ist die Frage nach der Erklärung unseres Verhaltens im Grunde so alt wie die Philosophie. Der Konflikt zwischen eher naturwissenschaftlichen und eher psychologischen oder geisteswissenschaftlichen Ansätzen wurde im Zuge von Neuzeit, Aufklärung und Moderne größer. Schließlich erlangte man immer genaueres Wissen über die physiologischen Grundlagen unseres Wahrnehmens und Bewegens. René Descartes (1596–1650) entwickelte aufgrund seiner empirischen

Studien, einschließlich Sektionen von Tieren, die Vorstellung von Lebewesen als Automaten; allein bei uns Menschen komme noch eine immaterielle und denkende Seele hinzu, die in Interaktion mit der Zirbeldrüse (Epiphyse) den Körper steuere (Descartes, 1649).

Im 19. Jahrhundert übertrug der Biologe Thomas H. Huxley (1825–1895) den Gedanken der Automaten vollständig auf den Menschen und begründete damit den *Epiphänomenalismus* (Huxley, 1874). Demnach nimmt das Bewusstsein die Vorgänge in der Welt zwar wahr, doch greift es nicht selbst in das Weltgeschehen ein. Volition (Willensvorgänge, von lat. *volo*, ich will) war für Huxley eher ein Gefühl oder im Sinne von Handlungsfreiheit zu verstehen, also dass man tun kann, was man tun will:

„Nach allem, was wir über die Vorgänge im Nervensystem wissen, müssen wir glauben, dass eine bestimmte molekulare Veränderung im zentralen Teil des Nervensystems auf eine uns völlig unbekannte Weise jenen Bewusstseinszustand hervorruft, den wir Empfindung nennen. [...] Andere molekulare Veränderungen führen zu Zuständen der Lust und des Schmerzes und zu der Emotion, die wir in uns selbst als Willenskraft bezeichnen. Ich habe keinen Zweifel daran, dass dies die Beziehung zwischen den physischen Prozessen des Tieres und seinen mentalen Prozessen ist. Daraus folgt zwangsläufig, dass diese Bewusstseinszustände in keiner Weise mit den Bewegungen der Körpermuskeln in Zusammenhang stehen können. Die Willensäußerungen der Tiere werden einfach Gefühlszustände sein, die ihren Handlungen vorausgehen.“
(Huxley, 1874, S. 365; dt. Übers.)

Huxley veranschaulichte das mit dem Beispiel eines Frosches, den man auf der Hand halte: Wenn man die Handfläche bewege, würde dieser automatische Bewegungen ausführen, um die Balance zu halten. Diese Bewegungen seien aber keine Ausdrücke von Willensvorgängen. Wäre der Frosch ein Philosoph, sinnierte Huxley, würde er vielleicht *denken*, dass er seine Bewegungen selbst bewusst verursache. Doch das sei ein Irrtum. Diese Diskussion über die kausale Rolle des Bewusstseins und mögliche Willenstäuschungen hält bis in unsere Zeit an (zum Beispiel Schleim, 2021; Wegner, 2002). Darauf werden wir noch näher eingehen.

Auch im deutschsprachigen Raum zweifelte man an der Willensfreiheit. Ein besonders prägnantes Beispiel hierfür ist der damals viel gelesene Physiologe Carl Vogt (1817–1895), der dazu schrieb:

„Der freie Wille existiert nicht und mit ihm nicht eine Verantwortlichkeit und eine Zurechnungsfähigkeit, wie sie die Moral und die Strafrechtspflege und Gott weiß wer noch uns auferlegen wollen. Wir sind in keinem Augenblicke Herren über uns selbst, über unsere Vernunft, über unsere geistigen Kräfte, so wenig als wir Herren sind darüber, daß unsere Nieren eben absondern oder nicht absondern sollen. Der

Organismus kann nicht sich selbst beherrschen, ihn beherrscht das Gesetz seiner materiellen Zusammensetzung. Was wir in einem Augenblicke denken, ist das Resultat der augenblicklichen Stimmung, der augenblicklichen Zusammensetzung unseres Gehirnes [...].“ (Vogt, 1852, S. 445–446)

Auffälligerweise hielt Vogt nicht nur Selbstkontrolle und Willensfreiheit für unmöglich, sondern zog er daraus bereits weitreichende Schlüsse für Recht und Moral. Schon zu seiner Zeit, vor über 170 Jahren, waren dafür die Konzepte von Verantwortlichkeit und Zurechnungsfähigkeit zentral. Das rief wiederum bedeutende Zeitgenossen aus Physiologie und Rechtswissenschaften auf den Plan, die seinen Aussagen und denen anderer Materialisten vehement widersprachen.

Zwei Beispiele dafür sind der einflussreiche Anatom und Rektor der Universität Wien Josef Hyrtl (1810–1894) sowie der renommierte Physiologe Emil du Bois-Reymond (1818–1896). Ersterer behauptete, für den Materialismus gebe es seit der Antike keine neuen Argumente, doch die Medien würden solche sensationalistischen Behauptungen gerne aufgreifen (Hyrtl, 1864/1897). Letzterer formulierte sein berühmtes „*ignorabimus*“, dass man die Grundlage von Bewusstsein und freiem Willen *nie* vollständig naturwissenschaftlich werde erklären können. Damit griff er ältere Gedanken des Philosophen Gottfried Wilhelm Leibniz (1646–1716) auf und nahm David Chalmers’ „hartes Problem des Bewusstseins“ vorweg (Chalmers, 1995; du Bois-Reymond, 1872; Schleim, 2022b).

Es ist aber schwierig, aus dieser Art von Argumenten *prinzipielle* Schlussfolgerungen zu ziehen. Sie hängen nämlich größtenteils davon ab, wie optimistisch oder pessimistisch wir die Möglichkeiten der zukünftigen Naturwissenschaften sehen. Es sei jedoch daran erinnert, dass auch Karl R. Popper (1902–1994) rund 100 Jahre nach du Bois-Reymond vom „Schuldscheinmaterialismus“ sprach (Popper & Eccles, 1985, S. 96f.). Damit meinte er die Angewohnheit von Vertretern materialistischer Positionen, die naturwissenschaftliche Lösung von Problemen wie dem des Bewusstseins oder der Willensfreiheit immer wieder aufs Neue für die nahe Zukunft in Aussicht zu stellen. So verlor erst kürzlich mit Christof Koch einer der führenden Bewusstseinsforscher unserer Zeit eine Wette gegen den Philosophen David Chalmers, Bewusstsein innerhalb von 25 Jahren neurobiologisch zu erklären (Lenharo, 2023b).

Für unsere Zwecke hier ist von Bedeutung, dass es schon auf weitreichende Aussagen wie die von Carl Vogt im 19. Jahrhundert Antworten von rechtswissenschaftlicher Seite gab, beispielsweise vom damals einflussreichen Berliner Strafrechtler Franz von Liszt (1851–1919). Dieser schrieb in seinem maßgeblichen, letztlich in 25 Auflagen erschienenem Lehrbuch:

„Das Strafrecht bedarf mithin zu seiner Grundlegung nicht der Annahme einer ursachenlosen Selbstbestimmung, einer dem Kausalgesetz entrückten Willensfrei-

heit. Es genügt vielmehr die von keiner Seite ernstlich bestrittene Annahme, dass alles menschliche Handeln psychisch (nicht mechanisch) verursacht, also durch Vorstellungen bestimmt, determiniert, motiviert ist.“ (von Liszt, 1900, S. 69)

Von Liszt stellte damit fest, dass das Strafrecht weder einen unverursachten Willen oder eine Seele voraussetzt, noch im Widerspruch zum Gedanken der Kausalität steht. Vielmehr komme es darauf an, Handlungen als *durch psychologische Vorgänge verursacht* anzusehen. Seine Abgrenzung von „mechanischen“ Ursachen wirft natürlich die Frage auf, was man unter diesem Begriff versteht. Die philosophische Forschung hierzu hält an (zum Beispiel Craver, 2007; Machamer, Darden & Craver, 2000). Demnach können mechanistische Erklärungen auch psychologische und soziale Faktoren beinhalten und müssen diesen nicht widersprechen.

Wie wir sahen, gibt es mindestens seit dem 19. Jahrhundert Hinweise darauf, dass das bestehende Strafrecht mit psychologischen Erklärungen unseres Verhaltens vereinbar ist – und bis heute ist es nicht zu der von Vogt und anderen geforderten normativen Revolution gekommen. In Abschnitt 4.2 werden wir die tatsächliche Umsetzung eines „Neurostrafrechts“ im frühen 21. Jahrhundert analysieren, doch schließen wir hier erst die Diskussion der Willensfreiheitsdebatte ab.

3.1.4 Willensfreiheit im 20. und 21. Jahrhundert

Es gab also schon ähnliche und weitreichende Aussagen über Willensfreiheit im 19. Jahrhundert wie die, die wir aus unserer Zeit kennen. Dabei sollten wir auch Sigmund Freud (1856–1939) nicht vergessen: Er wies im frühen 20. Jahrhundert darauf hin, dass wir weniger bewusste Kontrolle über unsere psychischen Vorgänge haben, als oft angenommen wird (Freud, 1917/1947; Schleim, 2012). Dabei dachte er einerseits an sexuelle Impulse und andererseits an die begrenzte Kontrolle unserer unbewussten Vorgänge.

Man könnte meinen, dass so eine Diskussion irgendwann zu einer Schlussfolgerung kommt – jedenfalls dann, wenn sie nicht nur von Philosoph*innen geführt wird, sondern sich empirische Forscher*innen darin einschalten. Wie wir im vorherigen Abschnitt sahen, hat Carl Vogt schon vor über 170 Jahren den freien Willen zu einem Ding der Unmöglichkeit erklärt und weitreichende moralische wie rechtliche Umwälzungen gefordert. Natürlich ist heute sehr viel mehr Wissen über den Körper und psychische Vorgänge verfügbar als zu Vogts oder Huxleys Zeiten. Doch an der Dynamik der neuen Debatte fällt auf, dass sie gar nicht mit der Veröffentlichung der viel zitierten Libet-Experimente in den frühen 1980ern aufkam. Diese wurden von manchen als Beleg für die unbewusste Determination unserer Entscheidungen verstanden, worauf wir gleich noch näher eingehen werden. Vielmehr fiel der starke Anstieg

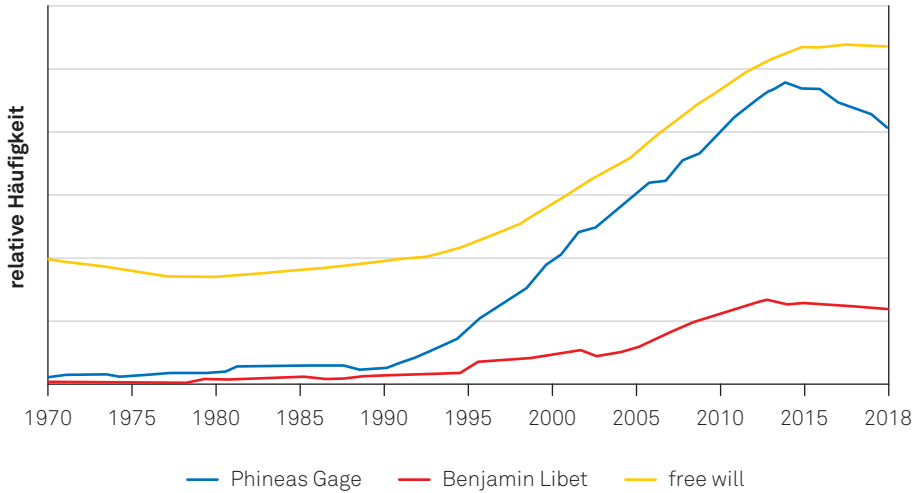


Abbildung 3-1: Wir sehen die relative Häufigkeit der Begriffe „free will“ (gelb), „Benjamin Libet“ (rot) und „Phineas Gage“ (blau) im Laufe der Zeit in englischsprachigen Büchern. Seit den 1990ern, der „Dekade des Gehirns“, kam es zu einem deutlichen Anstieg. Auf Phineas Gage kommen wir später zurück. Wie schon zuvor werden die Linien in unterschiedlichen Skalen dargestellt, sollten also nicht direkt miteinander verglichen werden (gelb: 10⁻⁶ Prozent; rot und blau: 10⁻⁸ Prozent). Quelle: Google Ngram

in die Zeit der „Dekade des Gehirns“ sowie des Aufkommens von Neuroethik und Neurolrecht (Abbildung 3-1).

Der Neurowissenschaftler Benjamin Libet (1916–2007) war vor allem ein Pionier auf dem Gebiet der Bewusstseinsforschung. Allerdings wagte er einen Blick in das Innere der „Blackbox“, unser Gehirn, als noch die Behaviorist*innen wie Skinner das Sagen hatten. Später würde er beschreiben, dass er darum keine Finanzierung für seine Forschung fand und man ihm sogar aus karrierestrategischen Gründen ganz davon abriet (Libet, 2004, S. 11). Wir erinnern uns, dass man Bewusstsein damals nicht für ein ernsthaftes wissenschaftliches Thema hielt. Libet führte seine bahnbrechenden Versuche mit der Elektroenzephalografie dennoch aus. Und auch wenn seine Versuchspersonen nur eine kleine, spontane Handbewegung tätigen sollten, sollte ihn das weltberühmt machen. Das geschah jedoch auf eine andere Weise, als er sich das selbst vorgestellt hatte.

Denn wie auch der Titel seines Buchs *Mind Time: The Temporal Factor in Consciousness* bezeugt (Libet, 2004), ging es ihm selbst ursprünglich weder um die Willensfreiheit noch um deren Widerlegung. Vielmehr untersuchte er mit neurowissenschaftlichen Verfahren die zeitliche Dynamik von Bewusstseinsprozessen. Andere Forscher*innen würden hervorheben, dass dem Bewusstsein eine *unbewusste* Hirnreaktion vorausgehe: das sogenannte Bereitschaftspotenzial. Es ist darum

interessant, hier etwas ausführlicher zu zitieren, welche Schlüsse Libet und seine Kolleg*innen selbst aus den Versuchsdaten zogen. Die Entscheidung für eine Bewegung stellten sie nämlich keineswegs als nur durch unbewusste Gehirnprozesse determiniert dar:

„Wenn man jedoch unsere Schlussfolgerung akzeptiert, dass spontane, freiwillige Handlungen unbewusst initiiert werden können, bleiben mindestens zwei Möglichkeiten, unter denen eine bewusste Kontrolle wirksam sein könnte. (1) Es könnte ein bewusstes ‚Veto‘ geben, das die Ausführung sogar der hier untersuchten ‚spontanen‘ selbstinitiierten Handlung abbricht. Dies ist möglich, weil die berichtete bewusste Absicht, auch wenn sie deutlich später als das Einsetzen des [Bereitschaftspotenzials] auftrat, eine beträchtliche Zeit (etwa 150 bis 200 Millisekunden) vor dem Beginn der Bewegung auftrat, wie sie durch das [Elektromyogramm] signalisiert wird. Sogar in unseren gegenwärtigen Experimenten haben Versuchspersonen berichtet, dass einige bewusste Handlungsimpulse ‚abgebrochen‘ oder gehemmt wurden, bevor eine tatsächliche Bewegung stattfand [...]. (2) Bei freiwilligen Handlungen, die nicht ‚spontan‘ und schnell ausgeführt werden, das heißt, bei denen eine bewusste Überlegung (ob man handeln soll oder welche Handlungsalternative man wählen soll) der Handlung vorausgeht, würde die Möglichkeit einer bewussten Initiierung und Kontrolle durch die vorliegenden Beweise nicht ausgeschlossen werden.“

(Libet, Gleason, Wright & Pearl, 1983, S. 641; dt. Übers.)

Die Fachleute hatten hier also mindestens zwei offene Möglichkeiten für die bewusste Kontrolle unserer Entscheidungen formuliert: Erstens hatte es schon damals experimentelle Befunde gegeben, dass das (angeblich) unbewusste Bereitschaftspotenzial im Gehirn auch dann auftrat, wenn die Bewegung *nicht* ausgeführt wurde; dies bestätigten ebenfalls spätere Wiederholungen dieses und ähnlicher Versuche (Schultze-Kraft et al., 2016; Trevena & Miller, 2010). Dann konnte das im Gehirn gemessene Bereitschaftspotenzial schon prinzipiell nicht die (alleinige) Ursache der Bewegung sein. Was Libet und Kolleg*innen immer wieder als bewusstes „Veto“ beschrieben, ließen andere Forscher, die die Willensfreiheit bestritten, schlicht unter den Tisch fallen (Haynes & Eckoldt, 2021; Wegner, 2002). Dies kritisierte schon Libet als Auslassungsfehler (Libet, 2004, S. 144).

Zweitens wurden mit dem Versuchsaufbau und insbesondere wegen der spontanen Natur der Bewegungen hier keine richtigen Willensentschlüsse untersucht. Mit anderen Worten: Was in dem Versuch erforscht wurde, lässt sich gar nicht auf das übertragen, worum es gemeinhin bei der Willensfreiheit geht, schon gar nicht in einem tieferen moralischen, rechtlichen oder philosophischen Sinn. Die Übertragung solcher reduzierter Ergebnisse auf die Willensfreiheitsdiskussion ist vielmehr eine Begriffsverwirrung. Diese harte Kritik am Vorgehen mancher medial stark ver-

treterer Fachleute aus der Hirnforschung wird durch anerkannte psychologische Handlungsmodelle gestützt, für die Abwägen, Planen und Bewerten von Handlungen von zentraler Bedeutung sind (zum Beispiel Heckhausen & Gollwitzer, 1987). Die fehlende Berücksichtigung, teils sogar ausdrückliche Ausgrenzung dieser Vorgänge in den neurowissenschaftlichen Experimenten schränkt deren Aussagekraft erheblich ein. Auch wenn man Gehirnaktivität heute genauer untersuchen kann als früher, sollte man darum bei der psychologischen Modellierung oder der ökologischen Validität der untersuchten Vorgänge keine Abstriche machen. Wie wir gleich sehen werden, wurden diese Defizite in den neueren Experimenten nicht ausgeräumt, sondern teils sogar noch vergrößert. Das ist, wohlgemerkt, keine Kritik an Benjamin Libet, der ja nie vorgab, mit seinen Versuchen die Willensfreiheit zu widerlegen.

Der renommierte Bewusstseinsforscher Anil K. Seth von der University of Sussex räumte die Existenz des Vetos später zwar ein, doch nur, um gleich zu relativieren: „Jedes bewusste ‚Veto‘ hat jedoch wahrscheinlich auch identifizierbare neuronale Vorläufer [...]“ (Seth, 2018, S. 2–3; dt. Übers.). In seinem etwas später erschienenen Buch über Bewusstsein schrieb er ausführlicher:

„Eine gängige Interpretation von Libets Experiment lautet, dass es ‚den freien Willen widerlegt‘. In der Tat ist dies eine schlechte Nachricht für die gespenstische Willensfreiheit (nicht, dass es noch mehr schlechte Nachrichten bräuchte), denn es scheint die Möglichkeit auszuschließen, dass das Erleben des Willens die freiwillige Handlung verursacht hat. Libet selbst war durch diese Implikation so beunruhigt, dass er in einem heute wie ein verzweifelter Rettungsversuch anmutenden Versuch die Idee in den Raum stellte, dass zwischen dem Moment des Drangs und der daraus resultierenden Handlung genügend Zeit für den spukhaften freien Willen bleibt, um einzugreifen und die Handlung zu verhindern. Wenn es keinen echten (das heißt gespenstischen) freien Willen gibt, dachte Libet, dann gibt es vielleicht immer noch ein ‚freies Nichtwollen‘. Das ist ein netter Trick, aber er funktioniert natürlich nicht. Die bewusste Hemmung ist ebenso wenig ein kleines Wunder wie die ursprüngliche bewusste Absicht.“

(Seth, 2021, S. 221–222; dt. Übers.)

Hier wird nach meinem Verständnis allerdings gleichzeitig das Thema gewechselt und ein Strohmännchen-Argument aufgebaut: Dass *nach* der unbewussten Einleitung der Bewegung aber bis zu 150 bis 200 Millisekunden *vor* ihrer Ausführung ein anderer kognitiver Prozess eingreifen kann, ist nicht nur eine von Libet und Kollegen getätigte empirische Hypothese (Libet, 2004; Libet et al., 1983), sondern wurde tatsächlich experimentell bestätigt (Schultze-Kraft et al., 2016). Mit einem „geisterhaften“ freien Willen hat das nichts zu tun. Außerdem verschiebt Seths Darstellung die Frage, ob eine Bewegung durch bewusste oder unbewusste Prozesse gesteuert wird

und was deren zeitlicher Ablauf ist, zur Spekulation darüber, ob Bewegungen *überhaupt* durch Gehirnprozesse verursacht werden.

Benjamin Libet und seine Kolleg*innen merkten schlicht an, dass es noch keine vollständige kausale Erklärung für den Vorgang gibt – und die fehlenden Variablen in der Gleichung bewusste oder unbewusste Prozesse sein können (siehe auch Dominik, Mele, Schurger & Maoz, 2023; Nestor, 2019). Wir sahen in diesem Kapitel schon, warum der Unterschied unbewusst/bewusst für das Recht so wichtig ist. Gleich wird deutlicher werden, warum ich diesem Thema so viel Aufmerksamkeit widme. Denn abgesehen von der leider oftmals mangelhaften Darstellung der Originaldaten in solchen Diskussionen geht es hier um den Kern des Zusammenhangs von Hirnforschung, Psychologie und unseren normativen Praktiken.

Bei einer Variante des Libet-Experiments im Hirnscanner behaupteten die Forscher*innen sogar, die Entscheidung, ob Versuchspersonen links oder rechts auf einen Knopf drücken würden, bis zu zehn Sekunden vor ihrer Bewusstwerdung vorhersagen zu können (Soon, Brass, Heinze & Haynes, 2008). Die viel zitierte Studie – zurzeit kommt sie auf über 2600 Zitationen auf Google Scholar – wurde sowohl von dieser Forschungsgruppe als auch in den Medien wieder gleich als Herausforderung für das Rechtssystem dargestellt (zum Beispiel Welberg, 2008). Auch sei das metaphysische Willensfreiheitsproblem nun gelöst, da wir in Wirklichkeit nur – Carl Vogt aus dem 19. Jahrhundert lässt grüßen – einem unpersönlichen Kausalgesetz folgten, das unsere Entscheidungen unbewusst festlege (Smith, 2011). Zum Thema unseres Buchs hieß es ganz konkret:

„Obwohl es schwer vorstellbar ist, dass unsere Entscheidungen unbewusst getroffen werden, haben diese Erkenntnisse wichtige Folgen. Können Menschen für ihre Handlungen zur Rechenschaft gezogen werden, wenn sie sich ihrer Entscheidungen erst bewusst werden, nachdem sie sie getroffen haben? Sie entscheiden.“

(Welberg, 2008, S. 411; dt. Übers.)

Davon abgesehen, dass auch in dieser Studie wieder spontane Bewegungen und keine Willensentschlüsse untersucht wurden; dass zwei Drittel (!) der Teilnehmer*innen vom Versuch ausgeschlossen werden mussten, teils sogar nach der Datenerhebung, weil sie nicht den Anforderungen der Versuchsleitung entsprachen; und dass die Vorhersage nur leicht über dem Zufallsniveau lag, ist mein zentraler Punkt: Bei all diesen Versuchen wie denen von Libet und Kolleg*innen und den Folgeuntersuchungen handelte es sich um *bewusste Kontrollprozesse*. Darauf deuteten sogar die Hirnregionen, die in der letztgenannten Studie mit den Entscheidungen in Zusammenhang standen, nämlich Teile des Frontalhirns und des Parietallappens (Soon et al., 2008). Das sind Regionen *par excellence*, wenn es um bewusstes Erleben und

Kontrollfunktionen geht. Unbewusste Einflüsse werden demgegenüber häufig mit subkortikalen Strukturen in Zusammenhang gebracht.

Wie man es auch dreht und wendet: Der vom Gehirn fremdgesteuerte Mensch scheint eher ein Konstrukt mancher Forscher*innen als ein Ergebnis der Hirnforschung zu sein (Schleim, 2023c). Wie wir sahen, baut das Recht nicht auf einer abstrakten Vorstellung vom freien Willen auf, sondern auf Wissen um Richtig und Falsch im Zusammenhang mit bewusster Kontrolle. In den hier genannten Experimenten wurde den Versuchspersonen zudem all das verboten, was Willensent-

Infobox

Warum das Libet-Experiment Willensfreiheit nicht widerlegt

Selbst 40 Jahre nach ihrer Veröffentlichung herrscht große Unklarheit darüber, wie Libets Experimente zu interpretieren sind. Die stärksten Argumente *gegen* die angebliche Widerlegung der Willensfreiheit sind:

- Bei diesen (und ähnlichen) Experimenten ging es um spontane Bewegungen, nicht um echte Willensentschlüsse; insbesondere fehlten Planung, Überlegung und ein sinnvoller Kontext der Entscheidungen.
- Diese (und ähnliche) Versuchsaufbauten erforderten gerade die *bewusste Kontrolle* durch die Versuchspersonen; die Teilnehmer*innen hätten auch jederzeit abbrechen können.
- Im gemeldeten Moment der bewussten Wahrnehmung kann die Ausführung der Bewegung oft noch gestoppt werden; dies ist nicht nur als „Veto“ theoretisch beschrieben, sondern auch mehrfach empirisch bestätigt worden.
- In Fortsetzung des letzten Punktes: Da das (angeblich) unbewusste Bereitschaftspotenzial im Gehirn selbst dann auftritt, wenn es gar nicht zur Bewegung kommt, *kann* es auch nicht deren (hinreichende) Ursache sein.
- In den Originalveröffentlichungen von Libet und Kolleg*innen wurde der freie Wille nicht einmal erwähnt und Benjamin Libet hat dessen Möglichkeit nie bestritten.
- Neurowissenschaftler*innen haben kürzlich gezeigt, dass das Bereitschaftspotenzial bei den einzelnen Entscheidungen gar nicht messbar ist und es möglicherweise nur ein Artefakt der statistischen Analyse ist (Schurger, Pak & Roskies, 2021; aber siehe Schmidt, Jo, Wittmann & Hinterberger, 2016). Aufgrund des schlechten Signal-Rausch-Verhältnisses der EEG ist das aber schwer zu entscheiden.
- In den Medien besteht ebenso wie in wissenschaftlichen Fachzeitschriften ein größeres Interesse an unerwarteten, neuen und sensationellen Funden, da diese mehr Aufmerksamkeit erhalten (Caulfield & Condit, 2012). Wahrscheinlich hat das die Diskussion verzerrt (vgl. Racine, Nguyen, Saigle & Dubljevic, 2017).

schlüsse überhaupt erst ausmacht: zum Beispiel Abwägen, Planen und Handeln aus Gründen. Wenn man Menschen dazu zwingt, eine Art Zufallsgenerator zu simulieren, dann kann das Ergebnis wohl kaum mehr als das repräsentieren.

Ein grundlegendes Problem hat zudem mit dem bereits erwähnten Selbstwiderspruch von Vertreter*innen des Materialismus und Naturalismus zu tun: Woher wissen die Fachleute bei diesen Experimenten, welche Hirnprozesse bewusst und welche unbewusst sein sollen? Wie wir sahen, sind wir von einer abgeschlossenen neurowissenschaftlichen Theorie des Bewusstseins noch weit entfernt. Stattdessen vertraut man ausschließlich auf die Versuchspersonen, die das Bewusstsein auf einen bestimmten Punkt festlegen müssen. Alles, was vorher stattfand, wird dann schlicht als unbewusst definiert (Soon et al., 2008; Libet et al., 1983). Das bedeutet aber: In ihrer Widerlegung der Willensfreiheit setzen die Forscher*innen einerseits voraus, dass die Versuchspersonen diesen Zeitpunkt – der in Wirklichkeit vielmehr ein ausgedehnter Prozess ist – punktgenau festlegen können; andererseits behaupten die Fachleute hinterher, dass wir Menschen uns prinzipiell über den Ablauf unseres Bewusstseins irren.

Das passt alles nicht wirklich zusammen und widerlegt mit Sicherheit nicht unsere normative Praxis. Und wie es zu Recht heißt: Außergewöhnliche Behauptungen erfordern außergewöhnliche Beweise. Wie wir sahen, sind die generell gegen jede Möglichkeit von Willensfreiheit gerichteten Argumente nur außergewöhnlich spekulativ. Weil sich die Missverständnisse über diese Versuche seit Jahrzehnten trotzdem hartnäckig halten, sind die Haupteinwände hier noch einmal in einer Infobox zusammengefasst. Somit führt uns diese Diskussion schrittweise zu dem letzten grundlegenden Punkt, wie die Ursachen unseres Verhaltens mit Recht und Moral und insbesondere dem Begriff der Verantwortung zusammenhängen.

3.1.5 Determination und Verantwortung

Wir sahen oben, dass Carl Vogt im 19. Jahrhundert Willensfreiheit für unmöglich und das Strafrecht für grundlegend falsch hielt, weil all unsere Entscheidungen und Handlungen *naturgesetzlich-kausal determiniert* seien. In Reaktion auf Experimente wie das von Libet und Kolleg*innen lautete die Argumentation im 20. und 21. Jahrhundert meist anders: Weil unsere Entscheidungen und Handlungen durch *unbewusste* Prozesse determiniert seien, hätten wir keinen freien Willen und basiere das Strafrecht auf einem falschen Menschenbild. Das Bewusstsein laufe den Gehirnentscheidungen gewissermaßen hinterher, so wie es auch Thomas Huxley mit dem Epiphanomenalismus formulierte.

Vogt – aber auch Anil Seth mit seinem Argument gegen den „geisterhaften“ freien Willen – richtete sich damit gegen eine bestimmte philosophische Position, nämlich die libertarianische (Roskies, 2006; Schleim, 2023c). Für diese sind unsere

Entscheidungen und Handlungen nicht vollständig naturgesetzlich-kausal determiniert. Stattdessen oder darüber hinaus wird ein Wille angenommen, der nach eigenen Gesetzmäßigkeiten entscheidet. Diesen Standpunkt kann man vielleicht philosophisch vertreten, mit allen Problemen, die die Annahme einer solchen metaphysischen Entität aufwirft (zum Beispiel Clarke, 2003; Franklin, 2018; Kane, 2009). Er ist aber keine Grundvoraussetzung des Strafrechts.

Das Argument gegen die *bewusste* Kontrolle unseres Verhaltens richtet sich zwar tatsächlich gegen eine Bedingung von Verantwortlichkeit in Recht und Moral. Wie ich zeigte, handelte es sich bei den genannten Experimenten aber gerade um bewusste Kontrollaufgaben (vgl. Wittmann, Montemayor & Dorato, 2025). Dass die Versuchsanweisungen die Teilnehmer*innen zu spontanem Verhalten anhielten und tatsächlich auch diejenigen ausgeschlossen wurden, die nicht spontan genug reagierten, ändert daran nichts: Die Versuchsperson hatten vorher bewusst in die Prozedur eingewilligt, behielten während des Experiments die bewusste Kontrolle ihres Verhaltens – etwa der Handbewegung oder des Knopfdrucks – und konnten jederzeit abbrechen. Hätte man ihnen glaubhaft versichert, dass sie mit den Knopfdrücken zum Beispiel Kindern lebensgefährliche Stromstöße erteilen, hätten sie dafür auch zur Verantwortung gezogen werden können.

Stellen wir uns ein praktisches Beispiel vor: Nehmen wir an, es gäbe eine genetische Prädisposition für das zurzeit beliebte Intervallfasten. Auch wenn genetische Einflüsse in der empirischen Realität viel kleiner sind, gehen wir für das Beispiel von einer genetischen Determination von nahezu 100 Prozent aus. Nun sei Kim so eine Person, deren Essgewohnheiten gewissermaßen in die Gene geschrieben sind und die jetzt für mehrere Stunden pro Tag aufs Essen verzichten will. Idealerweise hat Kim sich vorher über die Vor- und Nachteile dieser Diät informiert (vgl. Schleim, 2024). Doch das ist keine zwingende Voraussetzung für das Beispiel.

Um diese Diät umzusetzen, wird Kim vor allem die Tageszeit im Auge behalten und bestimmten Impulsen widerstehen müssen, während der Fastenperiode etwas zu essen. In dieser Hinsicht ist die Aufgabe gar nicht so anders als die (angeblichen) Willensfreiheitsexperimente, in denen Versuchspersonen auf die Zeit achten und abwechselnd auf einen Drang warten oder spontane Bewegungen ausführen sollten. Und wie der Bewegungsdrang beginnt auch das Hungergefühl oft unbewusst.

Der springende Punkt hierbei ist: Selbst wenn die Entscheidung für das Intervallfasten genetisch determiniert ist, setzt ihre Ausführung Wissen und bewusste Kontrolle voraus, dass also Kim im genannten Sinne eine minimal rationale Person ist. Wie wir sahen, sind genau das die Elemente moralischer und rechtlicher Verantwortlichkeit – und diese stehen völlig im Einklang mit den bekannten experimentellen Daten. Würde Kim stattdessen schlafwandelnd zum Kühlschrank laufen und das Fasten brechen, würden die Voraussetzungen für Verantwortlichkeit nicht vorliegen. Ähnlich würde es sich verhalten, wenn eine Person vor Hunger fast stürbe oder sich

wegen einer Impulskontrollstörung nicht beherrschen könnte. Dann würden wir ihr das Verhalten wahrscheinlich nicht oder zumindest weniger zurechnen. Auch das entspricht der üblichen Logik von Recht und Moral. Dass bei alledem neuronale Vorgänge stattfinden, steht dazu nicht im Widerspruch.

Dementsprechend setzt der seit dem 19. Jahrhundert wiederholt versuchte neurobiologische Angriff auf das Fundament des Strafrechts entweder eine Verwechslung eines bestimmten philosophischen Willensfreiheitsbegriffs mit der vorhandenen normativen Praxis voraus oder eine fragwürdige Interpretation experimenteller Befunde. Doch es ist nicht nur das Recht, das dieses Menschenbild von uns als *in der Regel* wissentlich und kontrolliert handelnden Personen verwendet: Wir sahen in Kapitel 2, dass Psycholog*innen Prozesse wie kognitive Kontrolle oder Impulsivität experimentell messen. Es handelt sich also um Eigenschaften, die wir mal mehr und mal weniger haben und die bei manchen von uns stärker ausgeprägt sind als bei anderen. Wer das Strafrecht in der genannten Weise unterminieren will, müsste auch zeigen, dass psychologische Forschung bei solchen Messungen Gespenstern nachjagt.

3.1.6 Beispiele aus dem Recht

Im US-amerikanischen Strafrecht kommt es auf den Kausalzusammenhang zwischen einem bestimmten psychologischen Vorgang, *mens rea* (wörtlich: schuldiger Geist) und einer Tathandlung beziehungsweise Unterlassung an, dem *actus reus* (wörtlich: schuldige Tat). Wenn beispielsweise John dem neuen Partner seiner Ex-Freundin aus Rache auflauert und ihn aus dem Hinterhalt erschießt, handelt es sich um ein Morddelikt. Falls ein Zufall ergäbe, dass am selben Ort aber einen Moment früher eine andere Person den Mann erschießt, weil sie eine Psychose hat und ihn für Satan persönlich hält, könnte John – trotz gleicher Bewusstseinsvorgänge und gleichen Verhaltens – höchstens für *versuchten* Mord verurteilt werden. Einen schon Toten kann man schließlich nicht ermorden. In diesem Fall fehlt es also am nötigen Kausalzusammenhang für das Morddelikt: Nicht John, sondern jemand anders tötete den Mann. Auch wenn das eher Lehrbuchbeispiele sind, veranschaulichen sie die „Denkweise“ des Strafrechts (vgl. Dressler, 2015).

Dieses Beispiel zeigt, dass das Vorliegen oder die Art einer Straftat nicht nur von Eigenschaften der Person, sondern auch der Konfiguration der Welt abhängt; Recht und Moral finden demnach nicht nur in unserem Kopf beziehungsweise Gehirn statt. Stellen wir uns 999 Personen vor, die mit dem Auto in Eile auf dem Weg nach Hause die zulässige Höchstgeschwindigkeit überschreiten. Bei der tausendsten Person aber springt ein Kind unerwartet auf die Fahrbahn und wird vom Auto tödlich verletzt. Auch wenn alle tausend Personen subjektiv dasselbe tun, begeht nur die

tausendste eine fahrlässige Tötung. In der Philosophie ist das als das „Problem des moralischen Zufalls“ bekannt (zum Beispiel Nagel, 1979).

Dieser Pechvogel könnte sich allenfalls vor der Verantwortung für den Unfall drücken, wenn er beispielsweise aufgrund einer schweren Bewusstseinsstörung die Kontrolle über das Fahrzeug verloren hatte, etwa wegen eines epileptischen Anfalls. Aber auch dann geht es nicht um die Frage, ob die Tat verursacht oder unverursacht war oder ob ein metaphysischer freier Wille in das Geschehen eingriff oder nicht – sondern um die *Art* der den Unfall bedingenden Ursache. Oder um es mit einer klassisch gewordenen Gerichtsentscheidung des britischen House of Lords zu einem Tötungsdelikt an einer Anhalterin zu sagen:

„Keine Handlung ist strafbar, wenn sie unwillkürlich erfolgt: Eine unwillkürliche Handlung in diesem Sinne – manche sprechen heute lieber von ‚Automatismus‘ – ist eine Handlung, die von den Muskeln ohne jegliche Kontrolle durch den Verstand ausgeführt wird, wie ein Krampf, eine Reflexhandlung oder eine Zuckung, oder eine Handlung, die von einer Person ausgeführt wird, die sich nicht bewusst ist, was sie tut, wie eine Handlung, die bei einer Gehirnerschütterung oder beim Schlafwandeln ausgeführt wird.“¹⁵

Wiederum geht es um die Frage, ob jemand die Kontrolle über das hatte, was er oder sie tat. Der Strafrechtler Joshua Dressler von der Ohio State University erklärte dazu in seinem einflussreichen Lehrbuch:

„Wenn also der Arm von D infolge eines epileptischen Anfalls V trifft, haben wir das Gefühl, dass der Körper von D, nicht aber die Person D, den Schlag verursacht hat. Im strafrechtlichen Kontext ist die Bewegung von Ds Arm begrifflich dasselbe wie ein Ast, der sich im Wind biegt und V trifft. Wenn D jedoch ‚will‘, dass sich ihr Arm bewegt, haben wir das Gefühl, dass D, und nicht nur ihr Arm, für die Verletzung von V verantwortlich ist. Ihr ‚handelndes Selbst‘ ist involviert. Ein persönliches, menschliches Handeln ist an der Verursachung des körperlichen Kontakts beteiligt.“

(Dressler, 2015; dt. Übers.)

Ob die Tat der Person zugerechnet wird, hängt also davon ab, ob sie darüber Kontrolle hat. Schon in den 1990ern identifizierte der Strafrechtler Stephen J. Morse von der University of Pennsylvania es als den „psycholegalen Fehler“, Verursachung für sich genommen als einen Entschuldigungsgrund aufzufassen (Morse, 1994). Nach dem Aufkommen des Neurorechts führte er aus:

15 Bratty v. Attorney-General, 1963, A. C. 386, United Kingdom House of Lords, dt. Übers.; online: <https://www.bailii.org/uk/cases/UKHL/1961/3.html> (Zugriff am 04.11.2025).

„Für die Beurteilung der Verantwortlichkeit spielt es keine Rolle, ob die Ursache für das fragliche Verhalten biologisch, psychologisch, soziologisch oder eine Kombination aus diesen drei Faktoren ist. Die Angabe einer genetischen oder neurophysiologischen Ursache leistet nicht mehr als die Angabe einer umweltbedingten Ursache. Die Frage ist immer, ob das betreffende rechtliche Kriterium für die Nichtverantwortlichkeit erfüllt ist, wie auch immer dieser Zustand verursacht worden sein mag.“

(Morse, 2007, S. 217; dt. Übers.)

Bei der Beurteilung der strafrechtlichen Verantwortlichkeit geht es demnach, wie wir bereits feststellten, um die vorliegenden psychologischen Vorgänge:

„Nach geltendem, positivem Recht ist ein Handelnder prima facie strafrechtlich verantwortlich, wenn er vorsätzlich und mit dem entsprechenden mentalen Zustand (der mens rea) handelt, der nach der Definition des Straftatbestands erforderlich ist, wie etwa Vorsatz, Wissen, Leichtfertigkeit oder Fahrlässigkeit. Das Strafrecht definiert eine Handlung in der Regel als eine vorsätzliche Körperbewegung, die von einem Handelnden ausgeführt wird, dessen Bewusstsein einigermaßen intakt ist.“

(Morse, 2007, S. 210; dt. Übers.)

Wir haben bereits ganz am Anfang dieses Buches psychische Vorgänge als verkörpert und eingebettet beschrieben. Von Fachleuten aus sowohl den Rechtswissenschaften als auch den empirischen Disziplinen wurde gleichermaßen immer wieder darauf verwiesen, dass unser Verhalten (auch) biologisch ist (zum Beispiel Greenberg & Bailey, 1994; Hyman, 2021; Turkheimer, 1998). Dazu stehen weder unsere Psychologie noch die strafrechtlichen Kriterien im Widerspruch. Darum *kann* das Aufdecken der neuronalen Grundlagen unserer psychologischen Vorgänge diese normative Praxis prinzipiell nicht widerlegen.

Eine Sensation wäre es vielmehr, würde die Hirnforschung eine Handlung als *unverursacht* identifizieren. Doch anstatt seine materialistische oder naturalistische Überzeugung zu ändern, würde man dann eher einen Fehler im Versuchsaufbau vermuten oder davon ausgehen, noch länger suchen zu müssen. Wir scheinen uns auch nicht dem eliminativen Materialismus anzunähern; im Gegenteil decken Neurowissenschaftler*innen auf, welche Strukturen beispielsweise der kognitiven Kontrolle unterliegen, etwa im dorsolateralen präfrontalen, vorderen cingulären und parietalen Kortex (zum Beispiel Breukelaar et al., 2017). Mit anderen Worten: Unser psychologisches Wissen wird neurobiologisch unterfüttert, im Einzelfall vielleicht verfeinert, bis auf Weiteres aber nicht grundlegend ersetzt.

Insofern stellte sich die Willensfreiheitsdiskussion wiederholt als falscher Alarm heraus, sowohl im 19. als auch im 20. und 21. Jahrhundert. Dass Benjamin Libet und seine Kollegen von Anfang an Möglichkeiten für die bewusste Kontrolle sahen, hätte

man als Warnzeichen ernst nehmen können (Libet, 2004; Libet et al., 1983). Aber dann wäre die Schlussfolgerung viel weniger sensationell gewesen: Ein Experiment zeigt, dass menschliche Entscheidungen auch im Laborversuch der bewussten Kontrolle unterliegen. Welches Medium hätte sich dafür interessiert?

Die Tatsache, dass Medienberichte widersprüchlich über solche Experimente berichteten, doch der sensationellen Interpretation bis in unsere jüngere Zeit Vorrang gaben (Racine et al., 2017), hätte uns ebenfalls zur Vorsicht mahnen können. Wir erinnern uns daran, dass der Anatom Josef Hyrtl die mediale Verzerrung für sensationelle Übertreibungen in dieser Debatte schon im 19. Jahrhundert wahrnahm (Hyrtl, 1864/1897; vgl. Hlade, 2025). Und wenn man sich auf lange Dauer nicht über die Bedeutung eines Experiments für die Willensfreiheit einigen kann, dann ist es wahrscheinlich gar kein aussagekräftiges Experiment zu diesem Thema.

Für den folgenden Abschnitt, in dem wir uns mit den normativen Folgen von Gehirnverletzungen beschäftigen, ist vor allem diese Feststellung von Bedeutung: Recht und Moral gehen für unsere Praxis der Verantwortlichkeit davon aus, dass wir in aller Regel Richtig und Falsch voneinander unterscheiden und nach dieser Einsicht handeln können. Wir sind vielleicht nicht immer, doch oft genug minimal rational handelnde Personen. Das schließt übrigens emotionale, irrationale oder unbewusste Einflüsse auf unser Entscheiden und Verhalten nicht prinzipiell aus, wie mehrere Beispiele gezeigt haben.

Diese Praxis ist zudem nützlich, kohärent und sinnvoll. Das äußert sich nicht zuletzt in der Geschichte und dem nachhaltigen Bestehen verschiedener Rechtssysteme weltweit, trotz der Behauptungen mancher Wissenschaftler und Philosophen, dass sie auf einem Irrtum beruhen. Dass es strittige Grenzfälle gibt, beschäftigt nicht nur Gerichte, sondern auch die juristische und moralphilosophische Forschung. Solche Grenzfälle unterminieren aber nicht den Sinn und Zweck der verwendeten Kriterien insgesamt, wie auch Zweifelsfälle zwischen den Kategorien „gesund“ und „krank“ nicht den generellen Sinn dieser Unterscheidung widerlegen.

3.2 Berühmte (un)verantwortliche Gehirne

Gegen Ende der 1960er-Jahre schrieben drei nordamerikanische Ärzte einen Brief mit dem Titel „Die Rolle von Gehirnerkrankungen bei Unruhen und Gewalt in Städten“, der im *Journal of the American Medical Association* abgedruckt wurde (Mark, Sweet & Ervin, 1967; dt. Übers.). Der Letztautor, Frank R. Ervin (1926–2015), würde einige Jahre später das Buch *Violence and the Brain* mitveröffentlichen und in seinem Nachruf als „weltweit führend im Erfassen der Bedeutung von Gehirnerkrankheiten für Verhaltensstörungen aller Arten“ und Experte in „mehreren Beratungsgremien

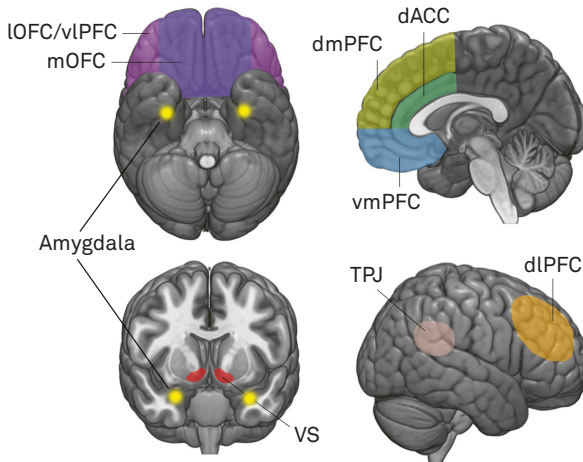


Abbildung 3-2: Hier sind verschiedene Gehirnregionen markiert, die in diesem Abschnitt von besonderer Bedeutung sind. Die Kleinbuchstaben geben grob eine Richtung an (d = dorsal, zum Rücken; l = lateral, zur Seite; m = medial, zur Mitte; v = ventral, zum Bauch). Die Abkürzungen in Großbuchstaben stehen für: ACC = anteriorer cingulärer Kortex; PFC = präfrontaler Kortex; OFC = orbitofrontaler Kortex; TPJ = temporoparietale Kreuzung; VS = ventrales Striatum. Quelle: Messimeris, Levy und Le Bouc (2023). Lizenz: CC BY 4.0

für den Präsidenten zu den Problemen der Gewalt, Kriminalität und Delinquenz“ gelobt werden.¹⁶

Der Gedanke der drei Ärzte war: Zwar leiden viele Menschen unter Armut, Arbeitslosigkeit, dem Leben in Slums und schlechter Bildung, doch nicht alle werden gewalttätig. Könnte der Unterschied zwischen Täter*innen und Nicht-Täter*innen nicht an Differenzen zwischen den Menschen liegen, beispielsweise in ihrem Gehirn? Sie verwiesen auf Untersuchungen einer damaligen „Neuro-Research Foundation“, wonach Verbrecher*innen häufiger Gehirnabweichungen hätten (Mark et al., 1967; siehe auch Pustilnik, 2009). Dabei ging es insbesondere um die Mandelkerne (Amygdalae). Diese sind Teil des limbischen Systems (von lat. *limbus*, Saum) und befinden sich an der Innenseite der Temporallappen (Abbildung 3-2). Auch heute noch erhalten diese kleinen Hirnstrukturen wegen ihrer Rolle bei der Emotionsverarbeitung große Aufmerksamkeit im Neurorecht (zum Beispiel Tonnaer, van Zutphen, Raine & Cima, 2023).

Der Vorschlag der drei Ärzte fiel in die Zeit der amerikanischen Bürgerrechtsbewegung, in der Schwarze gegen Diskriminierung kämpften. Laut Jonathan M. Metz, Psychiater und Professor für Soziologie und Medizin an der Vanderbilt University in Nashville, veränderte sich damals auch die Diagnose Schizophrenie in eine „schwarze Krankheit“ (Metz, 2009). Mit anderen Worten: Die Aggression, mit der sich Unterdrückte für gleiche Rechte einsetzten, wurde von verschiedenen Seiten neurobiologisch oder psychiatrisch pathologisiert. Unter der Überschrift „Angriffs-

¹⁶ <https://www.mcgill.ca/psychiatry/channels/news/obituary-dr-frank-r-ervin-1926-2015-249027> (Zugriff am 04.11.2025).