

8 Experimentelle Psi-Forschung

Stefan Schmidt

8.1 Einleitung

Viele Menschen berichten von telepathischen Erfahrungen, präkognitiven Träumen, Vorahnungen der Zukunft, die sich bestätigen, oder haben den Eindruck, dass sie Kraft ihrer Gedanken Phänomene in der materiellen Welt bewirken können. In der experimentellen Parapsychologie wird empirisch untersucht, ob es solche Phänomene tatsächlich geben kann, und wenn ja, unter welchen Bedingungen diese auftreten. Dabei wird das Experiment, insbesondere das Laborexperiment, als der Königsweg zur Beantwortung dieser Frage verstanden und genießt eine entsprechende Reputation. Dies liegt daran, dass im Labor störende Einflussvariablen minimiert werden können und somit die geforderte interne Validität, die zur Führung eines solchen Beweises aus empirischer Perspektive unabdingbar ist, gewährleistet werden kann. Der Transfer dieser Phänomene ins Labor mit seinen streng kontrollierten Rahmenbedingungen geht jedoch oft auf Kosten der externen Validität (Generalisierbarkeit) bzw. der ökologischen Validität (Grad der Alltagsnähe) (s. Kap. 33). Letztendlich hat jeder wissenschaftliche Zugang seine Einschränkung und es gilt auch hier die Regel, dass ein umfassendes Bild zu einer bestimmten Fragestellung nur aus der Zusammenschau methodisch unterschiedlicher, sich ergänzender Herangehensweisen erreicht werden kann.

Versteht man das parapsychologische Experiment als eine methodisch kontrollierte Untersuchung von Phänomenen, die im Rahmen außergewöhnlicher Erfahrungen

berichtet werden, dann stellt sich die Frage, um welche Phänomene es sich handelt. Alltagssprachlich wird oft hinsichtlich **Hellsehen**, **Telepathie**, **Präkognition** oder **Psychokinese** unterschieden. Das Problem mit diesen Begrifflichkeiten ist, dass sie bestimmte Mechanismen implizieren, die nicht wissenschaftlich gesichert sind.

Definition

Präkognition: Von Präkognition wird immer dann gesprochen, wenn es darum geht, Informationen über zukünftige Ereignisse zu erlangen. Hier wird bewusst ein neutraler Begriff verwendet, um assoziativ vorbelastete Ausdrücke wie Prophezeiung o. Ä. zu vermeiden.

Telepathie: Telepathie geht von einem Sender/In-Empfänger/in-Modell aus. Im engeren Sinne kann nur dann von einem telepathischen Ereignis oder Phänomen gesprochen werden, wenn eine Person zu einer Information gelangt, die ausschließlich in den Gedanken oder Gedächtnisinhalten einer anderen Person vorhanden ist, und wenn keiner der bekannten Übertragungswege für diese Information benutzt wurde.

Hellsehen: Im Unterschied zur Telepathie kann beim Hellsehen (engl. clairvoyance) die Psi-vermittelte Information auch in der Umwelt verfügbar sein. Da in den meisten Situationen, in denen es um Psi-Kommunikation zwischen zwei Personen geht, nicht zu klären ist, ob es sich um Hellsehen oder Telepathie handelt, fasst man diese beiden Begriffe auch unter dem Oberbegriff *general extrasensory perception* (GESP; dt.: *Allgemeine Außersinnliche Wahrnehmung* bzw. AASW) zusammen.

Psychokinese: Von Psychokinese (PK) wird gesprochen, wenn Zusammenhänge zwischen mentalen (bzw. intentionalen) Anstrengungen und Veränderungen in der materiellen Umwelt ohne Zuhilfenahme jeglicher Handlungen gefunden werden können. PK wird unterteilt in *Makro-PK* und *Mikro-PK*. Makro-PK bezieht sich auf alle offensichtlichen Phänomene, z. B. das Verbiegen von Gabeln oder das Verrücken von Schränken. Mikro-PK bezieht sich auf die Beeinflussung von Prozessen, die nicht mit dem bloßen Auge erkannt werden können, sondern statistische Auswertungen nötig machen, z. B. die Beeinflussung eines fallenden Würfels oder eines Zufallsgenerators.

Auch ist es möglich, das eine Phänomen mit einem anderen zu erklären. Wenn eine Person wiederholt vorhersagen kann, welche Lottokugel als nächstes fällt, kann dies – Betrug und Zufall ausgeschlossen – ein Fall von Präkognition sein. Das Ereignis ließe sich aber auch als Psychokinese interpretieren, wenn man davon ausgeht, dass die Person die Lottomaschine durch entsprechende mentale Anstrengungen beeinflussen kann. Geht man weiterhin davon aus, dass die Person die genauen Randbedingungen des Lottoautomaten erkennen und somit die zu ziehende Zahl vorab bestimmen kann, dann würde es sich um einen Fall von Hellsehen handeln. Um diesem Dilemma zu entkommen, ist es üblich, die zu untersuchenden Phänomene mit dem griechischen Buchstaben *Psi* als Referenz zu einem noch unbekanntem Phänomenkomplex zu bezeichnen.

8.2 Geschichte der experimentellen Parapsychologie

Als Beginn der experimentellen Parapsychologie wird oft die Gründung der „Society for Psychical Research“ (SPR) in London im

Jahre 1882 bezeichnet. Ihr Ziel war es, mit wissenschaftlichen Methoden und ohne Vorurteile Phänomene zu untersuchen, die sich der (damaligen) wissenschaftlichen Erklärung entzogen. Ein erster Umbruch in der experimentellen Herangehensweise ergab sich in den späten 1920er-Jahren durch den Botaniker J. B. Rhine, der an der Duke University in North Carolina ein Forschungsprojekt begann, das für die nächsten drei Jahrzehnte die wissenschaftliche Untersuchung der Psi-Phänomene dominierte. Das sogenannte **Rhine'sche Paradigma** zeichnete sich durch die Vorgabe der Lösungsmöglichkeiten aus, unter denen die Versuchsperson ein „Ziel“ auswählen musste, das sogenannte *forced-choice*-Vorgehen. Weitere bestimmende Merkmale waren die zahlreichen Wiederholungen kurzer einfacher Versuche, die in einer neutralen Umgebung durchgeführt und dann statistisch ausgewertet wurden. Rhine verwendete vor allem die bekannten *Zener-Karten* mit den fünf Symbolen Stern, Kreuz, Kreis, Quadrat und Welle für seine Experimente zur *Außer-sinnlichen Wahrnehmung* (vgl. Abb. 8-1).

Für die *Psychokinese*-Forschung verwendete er zahlreiche Apparate, die zum Beispiel kontinuierlich automatisch würfelten und somit die Einflussmöglichkeiten der Probanden/innen und der Versuchsleiter/innen minimierten. Bestimmendes Merkmal dieser Ansätze war, dass „*Psi-Effekte*“ im statistischen Heraustreten aus einer großen Anzahl von Experimenten gesucht wurden. *Psi* wurde hier nicht mehr als beeindruckendes Einzelphänomen mit Evidenzcharakter verstanden, sondern als **das systematische Abweichen von einer unter Zufall zu erwartenden Verteilung**. Dieser Ansatz ermöglichte zwar einerseits den empirischen Zugang zu diesem Phänomenbereich, führte aber auch zu einer Reduktion der externen Validität, also des Alltagsbezugs. Viele Ver-

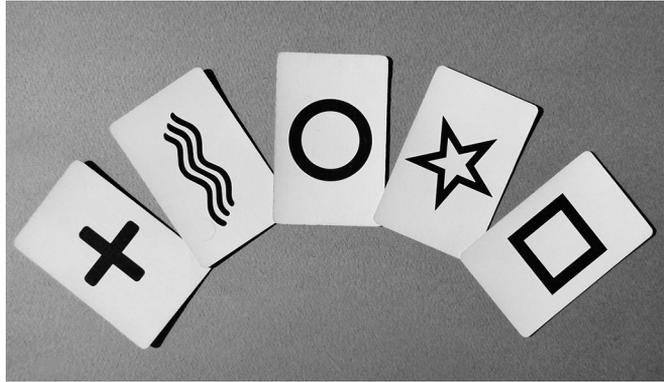


Abb. 8-1 Zener-Karten
(IGPP-Archiv).

suchspersonen empfanden die sich endlos wiederholenden Aufforderungen, die nächste Karte zu erraten, als langweilig und ermüdend. Ein Resümee der Studien zu Außer-sinnlicher Wahrnehmung aus dieser Phase findet sich bei Pratt und Rhine (1940). Hier werden die Ergebnisse von 139 Studien der Jahre 1882–1939 zusammengefasst.

Die 1960er-Jahre mit ihren vielen gesellschaftlichen und wissenschaftlichen Umbrüchen zogen auch einen Paradigmenwechsel in der parapsychologischen Forschung nach sich (Schmidt 2008). Insbesondere das Interesse an veränderten Bewusstseinszuständen und Transpersonaler Psychologie hatte großen Einfluss auf die experimentelle Herangehensweise an Psi-Phänomene. So ergab sich beispielsweise eine Hinwendung zur Introspektion als empirische Quelle für außersinnliche Informationen. Dem liegt die Idee zugrunde, dass außersinnliche Wahrnehmungen nicht von regulären sinnlichen Wahrnehmungen zu trennen sind. Da letztere jedoch das innere Erleben weitestgehend dominieren, gehen diese außersinnlich erlangten Informationen im Alltagsbewusstsein unter. Folglich wurde versucht, in **veränderten Bewusstseinszuständen** (z. B. Tiefenentspannung, Hypnose oder Meditation) Experimente

durchzuführen (Schmidt 2012a). Damit änderte sich aber auch das prinzipielle experimentelle Herangehen. Die einschränkende *forced-choice*-Methode wurde durch das sogenannte **free-response-Verfahren** ersetzt. Dabei soll die Versuchsperson ein zu ermittelndes, unbekanntes sogenanntes „Target“ (z. B. ein Ort oder ein Objekt) zeichnen oder verbal beschreiben. In einem zweiten Schritt beurteilen dann unabhängige Personen, die über das wahre Target nicht informiert sind, die Übereinstimmung des Berichts der Versuchsperson mit dem Target (im Vergleich zu einer Anzahl falscher Targets). Da diese Verfahren zeitaufwendig waren, wurde sehr viel Wert auf die Qualität des einzelnen Durchgangs gesetzt und, im scharfen Kontrast zum Rhine'schen Paradigma, nur sehr wenige Einzeldurchgänge ausgeführt.

Bedingt durch die Erfolge dieser Strategien und der Rezeption dieser Daten innerhalb einer breiteren Öffentlichkeit folgte ab den 1980er-Jahren eine weitere Phase, die sich durch die Konsolidierung einzelner Experimentalparadigmen auszeichnete. Bis dato waren zahlreiche unterschiedliche Experimente in verschiedenen Labors durchgeführt worden. Die kritische Betrachtung dieser Experimente durch Mitglieder von sogenannten Skeptiker-Organisationen, aber

vor allem auch durch Kollegen/innen innerhalb der Parapsychologie, führte zu zahlreichen methodischen Verbesserungen und Weiterentwicklungen. Es zeigte sich, dass es beinahe unmöglich ist, ein nahezu „wasserdichtes“ parapsychologisches Experiment aus dem Stand zu entwerfen und umzusetzen. Vielmehr bedarf es des wissenschaftlichen Diskurses mit den sogenannten Peers sowie mehrerer verbesserter Replikationen, um das anvisierte Ziel zu erreichen. Vor diesem Hintergrund bildeten sich in den letzten 30 Jahren einige gut definierte **Experimentalparadigmen** heraus, die dann in unterschiedlichen Labors durchgeführt wurden.

Eine zweite wichtige Komponente in dieser Phase war Ende der 1970er-Jahre die methodische Entwicklung der **Metaanalyse**. Bei dieser Methode werden die Ergebnisse mehrerer einzelner Studien auf der Basis der erzielten Effektstärke quantitativ zusammengefasst. Es kann also eine statistische Aussage hinsichtlich der Signifikanz und Größe eines Effektes über **mehrere unabhängige Studien mehrerer unabhängiger Autoren/innen** gemacht werden. Diese Methode wurde nun auch auf die parapsychologischen Experimentalparadigmen angewandt und erlaubte somit empirisch wesentlich besser gestützte Aussagen als auf der Basis von Einzelexperimenten (s. z. B. Radin 1997; Schmidt 2014).

8.3 Methodische Aspekte

Bei der Planung und Durchführung parapsychologischer Experimente hat die verwendete Forschungsmethodik einen sehr zentralen Stellenwert. Dies gilt besonders für die sogenannten **beweisorientierten Studien**, in denen versucht wird, den experimentellen Nachweis für die Existenz von

Psi zu führen. Durch den fortwährenden iterativen Prozess kritischer Debatten und methodischer Verbesserungen des Versuchsaufbaus sowie der Auswertungsstrategien ist heute ein sehr hoher Standard erreicht, der in einzelnen Fällen, z. B. bei der Qualität der Randomisation von Zufallsabfolgen, das Niveau der psychologischen Experimentalmethodik übersteigt. Da die Vielzahl der zu berücksichtigenden Aspekte in unterschiedlichen Experimenten hier kaum umfassend dargestellt werden kann (für eine detaillierte Ausführung s. Irwin u. Watt 2007), sollen exemplarisch vier Beispiele für eine spezifische Methodenentwicklung und -diskussion im Kontext der experimentellen Parapsychologie dargestellt werden.

Optional stopping: Unter *optional stopping* versteht man eine Einflussnahme auf die Ergebnisse eines Experimentes dadurch, dass die Zahl der Durchgänge in Abhängigkeit von den Resultaten gezielt gestoppt und das Experiment als beendet erklärt wird. Will man z. B. zeigen, dass eine Person in der Lage ist, beim Roulette das Auftreten der Farbe Rot signifikant zu erhöhen, so könnte man dafür theoretisch einen hin und wieder zufällig auftretenden Rotüberhang nutzen und das Experiment zu diesem Zeitpunkt der Beobachtung abbrechen. Zur Absicherung und Vorbeugung gegen diese Manipulationsmöglichkeit wurde – übrigens lange vor der Einführung von Studienregistern in der klinischen Forschung – in der Parapsychologie die Anzahl der Durchgänge **vor Beginn des Experimentes** in einem Protokoll festgelegt. Auch hat es sich eingebürgert, in der Regel keine Zwischenauswertung eines laufenden Experimentes durchzuführen.

Stacking effect: Beim *stacking effect* geht es um die Frage, ob eine bestimmte Randomi-

sationssequenz, also z. B. die (zufällige) Abfolge von Zener-Karten in einem Präkognitionsexperiment, für mehrere Versuchspersonen verwendet werden kann, oder ob es für jede Versuchsperson einer eigenen, neu generierten Sequenz bedarf. Der Hintergrund für diese Frage ist, dass menschliches Rateverhalten nie zufällig ist. So haben Menschen die Tendenz, eine Rateserie mit einem bestimmten Merkmal zu starten (z. B. eher ja als nein), oder sie neigen dazu, die real vorkommende Häufigkeit von Mehrfachereignissen (z. B. dreimal oder viermal Rot hintereinander) zu unterschätzen. Korreliert man nun die verwendete Zufallssequenz mit einer dieser Tendenzen, so könnte sich in Abwesenheit eines echten Psi-Effekts trotzdem ein künstlicher Trefferüberhang ergeben. Daher wird in parapsychologischen Experimenten großer Wert auf die **Qualität der Randomisation** gelegt. Zufallssequenzen werden idealerweise für jeden Durchgang neu generiert oder es wird ein limitiertes Set von balancierten Zufallssequenzen verwendet (Schmidt 2014). In Fällen, in denen sich diese Maßnahmen nicht ergreifen lassen, z. B. bei Tests mit großen Gruppen, sollten die Ergebnisse im Nachhinein statistisch mit der sogenannten *Greville-Formel* (Schmidt 2014) korrigiert werden.

Versuchsleiter-Effekt: Schon früh wurde in der Parapsychologie beobachtet, dass einige der Versuchsleiter/innen öfter positive Ergebnisse in Experimenten erzielten als andere. Man spricht hier von einem *Versuchsleiter-Effekt*. Dieser unterscheidet sich von dem in der Psychologie von Robert Rosenthal erforschten *Versuchsleiter-Erwartungseffekt*. Beim Versuchsleiter-Erwartungseffekt verfügt der oder die Versuchsleiter/in über lösungsrelevante Informationen für die Experimentalaufgabe, die sie der Versuchs-

person unbeabsichtigt und nonverbal signalisiert. Die in den parapsychologischen Experimenten beobachteten Versuchsleiter-Effekte lassen sich jedoch nicht vor diesem Hintergrund erklären. Schon in den 1950er-Jahren wurde die erste Studie durchgeführt, bei der zur Untersuchung dieser Beobachtung die **Variable Versuchsleiter** systematisch variiert wurde. Viele weitere zufällige Beobachtungen und auch mehrere Untersuchungen folgten seither (Schmidt 2014). Am bekanntesten sind die Experimente, bei denen die Versuchspersonen per Zufall entweder dem Psi-Phänomenen skeptisch gegenüberstehenden Richard Wiseman oder der für ihre erfolgreichen Experimentalserien bekannten Psi-Forscherin Marilyn Schlitz zugeteilt wurden. Hier zeigte sich in den ersten Versuchsserien ein Unterschied, der jedoch später nicht repliziert werden konnte (Schlitz et al. 2005). Insgesamt ist ein solcher Psi-Versuchsleitereffekt oder auch nicht klassischer Versuchsleitereffekt, wie er oft genannt wird, nicht ausreichend empirisch gesichert.

Decline effect: Unter dem *decline effect* versteht man die Beobachtung, dass die Größe der in Experimenten gefundenen Effektstärken mit der Wiederholung dieser Experimente abnimmt. Diese Beobachtung wurde schon früh in der Parapsychologie berichtet und dokumentiert. Man unterscheidet dabei zwischen einem **internen decline effect** innerhalb eines Experimentes oder einer Versuchsreihe und einem **externen decline effect**, der sich auf einen solchen Trend bei der mehrfachen Replikation eines ganzen Experimentes bezieht (Colborn 2007). Empirisch zeigt sich hier ein Missverhältnis von zahlreichen kursorischen Berichten und so gut wie keinen systematischen Untersuchungen oder Analysen zu diesem Thema. Interessant ist, dass dieses in der Parapsychologie

bereits seit vielen Jahren diskutierte Phänomen nun auch in der Biologie und vor allem in der Psychologie aufgegriffen wird. Auch hier gibt es bei vielen Experimenten ein Replikationsproblem dahingehend, dass sich die anfangs gefundenen Effektstärken später nicht mehr in dieser Größe replizieren lassen und es somit schwierig ist, ein bisher als solide angesehenes Experimentalparadigma zuverlässig zu implementieren (Yong 2012).

8.4 Einzelne Experimentalparadigmen

Nachfolgend werden die am weitesten verbreiteten parapsychologischen Experimentalparadigmen kurz vorgestellt und ihre Ergebnisse – basierend auf metaanalytischen Befunden – dokumentiert.

8.4.1 Ganzfeld

Beschreibung und alltagsweltliche Analogie

Das Ganzfeldparadigma wurde Mitte der 1970er-Jahre zeitgleich von Honorton und Braud in die Parapsychologie eingeführt (Schmidt 2014). Es entstand im Zuge des Paradigmenwechsels hin zu Experimenten mit veränderten Bewusstseinszuständen und *free-response*-Techniken. Ausgangspunkt war, dass man die „Übertragungsqualität“ in Telepathieexperimenten verbessern wollte, indem man den oder die „Empfänger/in“ sensorisch von lokalen und Alltagseinflüssen abkoppelt. Dies sollte durch den speziellen Bewusstseinszustand erreicht werden, der bei der sogenannten Ganzfeldstimulation entsteht. Dazu werden der „Empfänger“-Versuchsperson halbierte Tischtennisbälle auf die Augen gelegt und

diese mit rotem Licht bestrahlt. Gleichzeitig wird über Kopfhörer ein sogenanntes *weißes Rauschen* eingespielt (vgl. Abb. 8-2).

So beginnen die meisten Versuchspersonen nach einer Weile Bilder zu sehen und Geräusche zu hören, die nichts mit der unmittelbaren Außenwelt zu tun haben, sondern „intern“ generiert werden. Eine zweite Versuchsperson in einem anderen Raum – der oder die „Sender/in“ – versucht nun, der sich in der Ganzfeldstimulation befindenden Person Bilder zu „senden“. Dafür bekommt die Sender-Versuchsperson entweder ein Foto präsentiert oder sieht wiederholt einen kleinen Videoclip. Gleichzeitig hat die Person in der räumlich entfernten Ganzfeldstimulation die Möglichkeit, ihre Wahrnehmungen laut zu verbalisieren. Diese Berichte werden aufgenommen und der/die Sender/in kann diese Berichte zeitgleich



Abb. 8-2 Ganzfeldstimulation (IGPP).

hören und sie so eventuell in ihre Bemühungen, das Bild oder Video zu übermitteln, berücksichtigen. Die Experimente werden immer in einem sensorisch gut abgeschirmten Aufbau ausgeführt, der jegliche konventionelle Kommunikation unterbindet.

Die alltagsweltliche Analogie dieses Experimentes ist ganz offensichtlich das Phänomen Telepathie. Die beiden Versuchspersonen versuchen, ohne jegliche konventionellen Kommunikationsmöglichkeiten, mentale Inhalte auszutauschen.

Zur Auswertung des Experimentes müssen entweder die Person, die sich in der Ganzfeldstimulation befand („Empfänger/in“), oder aber unabhängige Gutachter/innen, denen die verbalen Protokolle vorgelegt werden, aus einer Serie von vier Bildern/Videos das im Experiment tatsächlich verwendete auswählen. Die Zufallswahrscheinlichkeit für einen Treffer beträgt also 25%. Bei Erfolg dieses Experimentes sollte die Trefferrate der Versuchspersonen signifikant über diesem Wert liegen.

Methodische Schwierigkeiten und Besonderheiten

In den 1980er-Jahren gab es eine öffentliche Kontroverse zwischen dem Experimentalparapsychologen Charles Honorton, der viele Ganzfeldexperimente selbst durchgeführt hatte, und dem Skeptiker Ray Hyman, der Honortons positiven Befunden kritisch gegenüberstand (Schmidt 2014). Hyman legte 1985 eine Metaanalyse vor, in der er auf einige methodische Mängel hinwies: Probleme mit der Randomisation, mit den statistischen Verfahren der Auswertung und mit der sensorischen Abschirmung. Honorton konterte mit seiner eigenen Metaanalyse, die diese methodische Kritik aufgriff. Auch der prominente Methodiker Robert Rosenthal schaltete sich in diese Debatte ein. Schließ-

lich verfassten Hyman und Honorton ein gemeinsames Communiqué, das methodische Standards und Verfahren für künftige Ganzfeldexperimente festlegte. In den folgenden Jahren entwickelte Honorton einen Aufbau und Ablauf für Ganzfeldexperimente, der weitestgehend automatisiert war und zur Unterscheidung unter dem Begriff „**Autoganzfeld**“ geführt wird (Palmer 2003). Dieser Standard hat heute noch weitestgehend Gültigkeit; allerdings kam es im Zuge der Weiterentwicklung der Computertechnik zu vielen weiteren Automatisierungen, besonders bei Auswahl, Präsentation und Bewertung der Bilder/Videos.

Metaanalysen

Das Ganzfeldexperiment ist mit der Methodik der **Metaanalyse** eng verknüpft. Es war das erste Experimentalparadigma, bei dem man sich erhofft hatte, die inkonsistenten Befunde der Primärstudien auf metaanalytischem Niveau auflösen und damit eine Entscheidung der Frage, ob es *Psi* gibt oder nicht, herbeiführen zu können. Im Verlauf einer zwanzigjährigen Diskussion zeigte sich, dass auch das Instrument der Metaanalyse subjektive Elemente enthält (Auswahl der Studien, Konzeption des statistischen Modells etc.) und somit die grundsätzliche Frage nicht endgültig lösen konnte.

Über die beiden ersten konkurrierenden Metaanalysen aus dem Jahre 1985 wurde bereits weiter oben berichtet. In der Folge führte Honorton mit seinem neuen Versuchsaufbau zehn weitere Studien mit 329 Einzelsitzungen durch. Die Metaanalyse dieser Experimente erschien 1994 im *Psychological Bulletin* (Bem u. Honorton 1994) und dokumentierte für die zehn Studien eine durchschnittliche Trefferrate von 32% ($p = 0.002$). Im Jahr 1999 erschien in derselben Fachzeitschrift eine Metaanalyse von

Milton und Wiseman, in der 30 neue Studien berücksichtigt wurden (Milton u. Wiseman 1999), die nach dem Erscheinen des Honorton-Hyman-Communiqués durchgeführt worden waren (die Autoganzfeldstudien von Honorton blieben dabei unberücksichtigt; vgl. Palmer 2003). Dieser neue Datensatz verfehlte die statistische Signifikanz ($p = 0.24$). Doch bereits zwei Jahre später erschien ebenfalls im *Psychological Bulletin* die nächste Metaanalyse (Storm u. Ertel 2001), in der die Autoren erstmals versuchten, *alle* verfügbaren Ganzfeldexperimente zu kombinieren. Insgesamt 79 Studien ergaben einen hochsignifikanten p-Wert von $7,8 \times 10^{-9}$ bei einer Effektstärke von $r = 0,138$ (entspricht $d = 0,28$). Nimmt man lediglich die 40 Studien nach Erscheinen des Communiqués, ergibt sich ein deutlich schwächeres, aber immer noch signifikantes Ergebnis von $p = 0.03$. Die letzte Metaanalyse schließlich stammt von Storm, Tressoldi und Di Risio (2010), die alle neueren Studien nach der Metaanalyse von Storm und Ertel aufnahmen. Für 29 Studien berechneten sie eine Effektstärke von $r = 0,142$ ($d = 0,29$) mit $p = 2,1 \times 10^{-8}$. In einer 2011 erschienenen Übersichtsarbeit über die Ganzfeld-Kontroverse findet sich eine einfache Zusammenfassung aller Studien seit dem Honorton-Hyman-Communiqué (Williams 2011).

Insgesamt zeigt sich über die beachtliche Studienanzahl also ein sehr komplexes Bild, das im Ganzen immer wieder in einen **signifikanten Effekt** mündet. Dabei ist die Effektstärke mit einem d -Wert von knapp unter 0,3 zwar klein – aber auch nicht vernachlässigbar. Es ist daher vorstellbar, dass solche Effekte auch unter Alltagsbedingungen vorkommen. Die Streuung der Effekte weist aber auch darauf hin, dass hier kaum von einem stabilen und in diesem Sinne zuverlässig kontrollierbaren Effekt ausgegangen werden kann.

8.4.2 Direkte mentale Interaktion

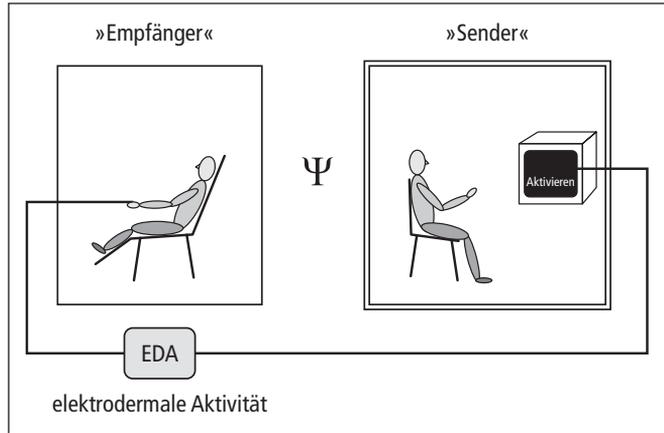
Beschreibung und alltagsweltliche Analogie

Unter *direkter mentaler Interaktion* (englisch DMILS: *direct mental interaction with living systems*) gruppiert man Experimente, bei denen eine Person durch mentale Anstrengungen versucht, mit einer zweiten, räumlich separierten Person in Kontakt zu treten. Das Besondere ist, dass diese Interaktion nicht über bewusste Inhalte erfolgt, sondern über **physiologische oder verhaltensbezogene Maße**. In dem am häufigsten durchgeführten Experiment, dem *EDA-DMILS-Experiment*, wird eine Versuchsperson des Versuchspaares gebeten, die zweite Person aus der Ferne abwechselnd in einem physiologischen Sinne zu aktivieren oder zu beruhigen. Dieses Ziel der Interaktion ist die sogenannte **Intention**. Die Person, die diese Intention betrifft, kennt in der Regel die Aufgabe der anderen Person, aber sie ist nicht über den Zeitplan der „Beruhigen“- , „Pause“- oder „Aktivieren“-Intervalle informiert. Als Maß der Aktivierung wird die elektrodermale Aktivität (EDA, auch Hautleitfähigkeit oder Hautwiderstand), ein Indikator sympathischer Erregung, von der Zielperson abgeleitet und dessen Fluktuation mit der Intention der anderen Versuchsperson verglichen (s. Abb. 8-3). Das erste diesbezügliche Experiment wurde von William Braud 1977 durchgeführt (Schmidt 2014).

Auf diesem experimentellen Ansatz basierend, haben sich drei spezifische Paradigmen herausgebildet, die häufig wiederholt wurden und somit einer metaanalytischen Betrachtung zugänglich sind. Dies sind:

- *EDA-DMILS*: Die Intention ist Aktivieren vs. Beruhigen. Die abhängige Variable ist die EDA der Zielperson.

Abb. 8-3 EDA-DMILS-Experiment: Die Person auf der rechten Seite sieht in ihrem Bildschirm die elektrodermale Aktivität der linken Person als einen Indikator autonomer Erregung. Ihre Aufgabe ist es, gemäß der Instruktion auf dem Bildschirm die linke Person aus der Entfernung zu aktivieren oder zu beruhigen.



- *Remote Staring*: Die Intention besteht darin, die andere Person über einen Monitor, der ein mit einer Videokamera aufgenommenes Bild live überträgt, entweder anzuschauen oder nicht anzuschauen. Die abhängige Variable ist ebenfalls die EDA (Schmidt 2008).
- *Attention Focusing Facilitation Experiment (AFFE)*: Hier wird die Zielperson gebeten, sich auf eine Kerze zu konzentrieren und jedes Mal einen Knopf zu drücken, wenn sie merkt, dass sie diesen Aufmerksamkeitsfokus verloren hat. Die Intention der zweiten Person besteht darin, die Zielperson zu unterstützen (oder nicht), indem sie sich entweder auch auf eine gleich aussehende Kerze in ihrem Experimentalraum konzentriert und sich mental mit der Zielperson verbindet (oder nicht). Die abhängige Variable ist die Anzahl der Knopfdrücker, verglichen in den Unterstützungs- und Nichtunterstützungsphasen (Schmidt 2012b).

Hinsichtlich der alltagsweltlichen Analogie werden die Studien zur direkten mentalen Interaktion oft als Grundlagenexperiment für die Frage herangezogen, ob es **Fernhei-**

lung geben kann und ob man aus der Ferne mit körperlichen, physiologischen oder gesundheitlichen Aspekten interagieren kann. Beim Remote-Staring-Experiment tritt hingegen die Frage in den Vordergrund, ob man merkt, dass man aus der Ferne angeschaut wird (zu dieser Frage s.a. Abschn. 8.4.3, wobei „merken“ hier nicht in einem bewussten kognitiven Sinne gemeint ist, da ja lediglich untersucht wird, ob sich **bestimmte physiologische Parameter** verändern. Prinzipiell kann die Idee der direkten mentalen Interaktion mit der Frage in Zusammenhang gebracht werden, ob eine wohlwollende (oder feindliche) mentale Haltung oder Absicht einen Zusammenhang mit tatsächlichen Ereignissen in der physikalischen Welt haben. Bringt es beispielsweise etwas, jemandem die Daumen zu drücken, einem kranken Menschen liebevolle Gedanken zu senden oder für das Wohl eines anderen zu beten?

Methodische Schwierigkeiten

Bei den EDA-DMILS- und den Remote-Staring-Experimenten zeigte sich bei einer Reanalyse Ende der 1990er-Jahre (Schmidt

2014), dass in vielen Studien inadäquate EDA-Methoden verwendet wurden. Dies kann während des Experiments zu einem kontinuierlichen Trend in den Daten führen. Solche Verschiebungen sind aber auch aus anderen Gründen denkbar (z. B. Ermüdung der Versuchsperson). Der Schlüssel im Umgang mit potenziellen Trends in den physiologischen Daten ist die angemessene Balancierung der Randomisationssequenz. In einer Metaanalyse ergab sich eine signifikante negative Korrelation für die Qualität der EDA-Messung mit den gefundenen Effektstärken von $r = -0,35$ (Schmidt 2014).

Metaanalysen

William Braud und Marilyn Schlitz verfassten zwei Übersichtsarbeiten über die große Serie von DMILS- und Remote-Staring-Experimenten, die sie an der *Mind Science Foundation* in San Antonio durchführten. Hierbei handelt es sich jedoch um simple Aufzählungen von Effektstärken ohne prinzipielle methodische Überlegungen, die bei jeder Metaanalyse zu berücksichtigen sind (Schmidt 2014). Zwei vollständige Metaanalysen sämtlicher Studien zu EDA-DMILS und Remote Staring bis Ende 2000 wurden von uns selbst durchgeführt. In diesen Arbeiten wird auch detailliert die metho-

dische Qualität der Primärstudien berücksichtigt (Schmidt 2014; Schmidt et al. 2004). Für das AFFE-Paradigma wurde eine ähnliche Metaanalyse im Jahr 2012 veröffentlicht (Schmidt 2012b). Insgesamt ergaben alle drei Metaanalysen ein nahezu identisches Bild einer **kleinen Effektstärke** von etwas mehr als $d = 0,1$, die aber bedingt durch die große Studienzahl (insgesamt 62 Primärstudien mit nahezu 2000 Einzelversuchen) signifikant ist (vgl. Tab. 8-1). Die Gesamtheit dieser Forschungsarbeiten dokumentiert somit sehr solide eine **Anomalie**, die mit den untersuchten Hypothesen in Einklang steht. Die gefundene Effektstärke ist dabei jedoch so gering, dass alle bisher durchgeführten Studien hinsichtlich ihrer Versuchsdurchgänge als zu klein betrachtet werden müssen, und dass von diesem Effekt auch nahezu keine klinische Relevanz zu erwarten ist. Interessant ist, dass trotz der unterschiedlichen Versuchsaufbauten und trotz der unterschiedlichen Präsentationen der jeweiligen Experimente für die Versuchspersonen die Effektstärken in den drei Paradigmen nahezu identisch sind. Dies kann als starker Hinweis darauf gedeutet werden, dass die Gemeinsamkeit der drei Experimente, nämlich die Operationalisierung einer *Intention* hinsichtlich der Beeinflussung einer zweiten Versuchsperson, hier tatsächlich die Schlüsselvariable ist.

Tab. 8-1 Ergebnisse der drei Metaanalysen zur direkten mentalen Interaktion. k bezieht sich auf die Anzahl der Studien, n auf die Anzahl der Versuchsdurchgänge, d gibt die gemittelte und gewichtete Effektstärke an, p die zugehörige Signifikanz.

Experiment	k	n	d	p
DMILS	36	1015	0,106	0,001
Remote Staring	15	379	0,128	0,013
AFFE	11	576	0,114	0,029
Insgesamt	62	1970		

8.4.3 Blickwahrnehmung

Beschreibung und alltagsweltliche Analogie

Das Gefühl, Blicke im Rücken zu spüren oder von hinten angeschaut zu werden, ist fast jedem Menschen vertraut. Zur Frage, ob es sich hierbei um eine außersinnliche oder doch sinnlich vermittelte Wahrnehmung handelt, wurden immer wieder Experimente durchgeführt, von denen das erste gut dokumentierte bereits 1898 an der Stanford University stattfand (Schmidt 2008). Im Jahr 1994 fasste *Rupert Sheldrake* die unklare empirische Lage zu diesem Phänomen zusammen und führte in den folgenden Jahren zahlreiche Experimente durch. Dabei handelte es sich meist um **Gruppenexperimente mit Schulklassen**. Der experimentelle Aufbau ist denkbar einfach. Eine Person wendet der anderen den Rücken zu und diese starrt für kurze Zeit (10–20 s) entweder auf den Rücken der ersten Person oder anderswohin im Raum. Danach wird die angeschaute Person befragt, ob sie gerade angeschaut wurde oder nicht, dann folgt der nächste Durchgang. Meist finden 20 Durchgänge statt und oft werden danach die Rollen gewechselt.

Methodische Schwierigkeiten

Sheldrake führte seine Experimente unter einfachsten Bedingungen in Schulklassen durch. Diese Arbeiten zeigen viele methodische Schwächen, die von Kritikern nach und nach benannt und von Sheldrake auch bereitwillig aufgegriffen und verbessert wurden. So gibt es natürlich das Problem der sensorischen Hinweisreize. Das angeschaute Kind hört zum Beispiel, ob sich das andere abwendet oder nicht, oder sieht dies in einem spiegelnden Fenster. Über das direkte **Feedback** können die Schüler so die Bedin-

gungen schnell erkennen. Weiterhin ergaben sich Probleme mit den Zufallssequenzen, abgesehen von der Frage, ob die Schüler/innen schummeln. Zuletzt wurden die Experimente über größere Entfernungen durchgeführt: Die anschauenden Kinder waren im Klassenzimmer und die anzuschauenden im Schulhof vor den Fenstern des Klassenzimmers. Die Kinder im Hof trugen Augenbinden und Ohrstöpsel, direktes Feedback wurde nicht gegeben, die Zufallssequenzen wurden balanciert und optimiert (Sheldrake 2005).

Metaanalysen

Diese Experimente sind den oben beschriebenen *Remote-Staring*-Experimenten sehr ähnlich (für eine Systematisierung s. Schmidt 2008). Eine adäquate Metaanalyse der hier beschriebenen Blickwahrnehmungsexperimente liegt bislang nicht vor, abgesehen von den Zusammenfassungen von Rupert Sheldrake (2005) zu seinen eigenen Studien. Radin (2005) publizierte eine metaanalytische Zusammenfassung der vorliegenden Experimente, aber seine Arbeit berücksichtigt keine systematische Literaturrecherche und Qualitätsmerkmale der doch sehr stark variierenden Studien. Insgesamt sind 60 Experimente mit 33 357 Einzeldurchgängen enthalten, von denen allerdings 92% von Rupert Sheldrake beigetragen wurden. Die Metaanalyse ergab eine mittlere Trefferrate von knapp über 54% (Zufallserwartung 50%) und eine Effektstärke von $r = 0,11$ ($d = 0,23$; $p = 10^{-28}$). Allerdings erwies sich die Stichprobe der Einzelexperimente statistisch als sehr heterogen. Ein weiteres Problem mit diesem Datensatz ist, dass diese Art von Experimenten durch ihre Einfachheit und dadurch, dass sie oft als Projekte von Studierenden oder Laien durchgeführt werden,

vermutlich einen starken *Publikationsbias* aufweisen. Damit ist gemeint, dass es eventuell noch zahlreiche weitere Studien (z.B. Bachelorarbeiten, Haus- oder Seminararbeiten, private Experimente) gibt, von denen ein Großteil keine Effekte gezeigt haben und die deswegen nicht publiziert wurden. Das Gesamtbild wird dadurch verzerrt und die Ergebnisse der Metaanalyse müssen insofern zurückhaltend interpretiert werden.

8.4.4 Mikro-Psychokinese

Beschreibung und alltagsweltliche Analogie

Unter Psychokinese (PK) versteht man die **mentale Einwirkung** auf die materielle Welt. Kann ein Mensch ausschließlich mit der Kraft seiner Gedanken Veränderungen seiner materiellen Umwelt bewirken? Die Geschichte der systematischen PK-Experimente beginnt ebenfalls mit *J.B. Rhine* und konzentriert sich auf die Frage, ob es Menschen gelingen kann, durch intentionale Anstrengung einen Würfel auf eine bestimmte Zahl fallen zu lassen. Zur methodischen Absicherung ließ Rhine sogar eine spezielle Würfelmachine anfertigen. In einer 1991 durchgeführten Metaanalyse wurden 73 Studien mit insgesamt 2,6 Millionen Würfeln von 2569 Personen zusammengefasst (Radin u. Ferrari 1991). Es ergab sich eine sehr kleine, aber hochsignifikante Effektstärke von $r = 0,012$ ($d = 0,024$, $p < 10^{-70}$).

Im Jahre 1970 baute der deutsch-amerikanische Physiker *Helmut Schmidt* den ersten **Zufallsgenerator**, der auf einem radioaktiven Zerfallsprozess beruht. Dieser quantenmechanisch beschreibbare Zerfall ist lediglich mit einer bestimmten Wahrscheinlichkeit vorherzusagen, aber niemals

der genaue Zeitpunkt. Somit erzeugt der Schmidt'sche Zufallsgenerator ein **definitiv unvorhersagbares Ereignis**. Man spricht bei diesem Aufbau von einem *Random Number Generator* (RNG) und ordnet Experimente mit solchen Zufallsgeneratoren der Mikro-PK-Ebene zu. In RNG-Experimenten werden Versuchspersonen gebeten, den Output des Zufallsgenerators intentional in die eine oder andere Richtung zu verändern. Dazu werden die nicht erkennbaren Zufallsereignisse im Versuchsaufbau in ein erkennbares Signal umgewandelt. Dieses sogenannte *Feedback* kann zum Beispiel aus unterschiedlichen Tonhöhen oder aus wandernden Lichtpunkten bestehen. Mit diesem Paradigma wurde in den folgenden 30 Jahren sehr umfassende Forschung durchgeführt. Wichtige Arbeiten dazu stammen unter anderem von dem deutschen Parapsychologen Walter von Lucadou (z.B. 1995). Eine weitere zentrale Rolle spielte das Projekt *Princeton Engineering Anomalies Research* um Roger Nelson und Robert Jahn (Jahn et al. 1997). Diese fanden in insgesamt 522 einzelnen Experimenten immer wieder sehr kleine Abweichungen der Zufallssequenzen in die intendierte Richtung (in der Größenordnung von 16–26 vom Erwartungswert abweichenden Ereignissen je 100000). Ende der 1990er-Jahre bildete sich ein großes Konsortium aus mehreren Arbeitsgruppen, mit dem Ziel, diese Experimente mehrfach in verschiedenen Labors zu wiederholen und die Befunde zu replizieren. Die Replikation misslang. Zwar zeigten sich in exploratorischen Analysen einige nicht zu erwartende Unregelmäßigkeiten in den Zufallsdaten, aber hinsichtlich der hypothetisierten Zielgröße, der Veränderung der Anzahl vordefinierter Zufallsereignisse, ergab sich keine Abweichung von der Zufallserwartung (Jahn et al. 2000).

Metaanalysen

Hinsichtlich metaanalytischer Zugänge wurde das Feld lange von einer frühen, methodisch jedoch nicht ausgereiften Metaanalyse von Radin und Nelson (1989) geprägt, die insgesamt positive Effekte des RNG-Paradigmas berichtet. Im Jahre 2006 wurde dann von Bösch, Steinkamp und Bolter im *Psychological Bulletin* (2006) die bisher sicherlich aufwendigste Metaanalyse der Parapsychologie publiziert. Die Metaanalyse umfasst 380 Studien, deren Größe und Umfang sehr stark variieren. Insgesamt ergab sich eine sehr kleine, aber signifikante Effektstärke von $\pi = 0,500035$ ($p = 0,01$, beim Effektstärkemaß π steht $\pi = 0,5$ für keinen Effekt), allerdings war die Verteilung der Effektstärke so inhomogen, dass diese Zusammenfassung eigentlich nicht gerechnet werden sollte. Aufgrund der Verteilung der Effektstärken und einer Simulationsrechnung argumentieren die Autoren, dass vermutlich viele Studien mit nichtsignifikantem Ausgang nicht publiziert worden sind und dass die Berücksichtigung dieser Studien den ohnehin sehr kleinen Effekt vollends verschwinden lassen würde. Diese einseitige Interpretation wurde allerdings von Kritikern infrage gestellt (Radin et al. 2006), die argumentierten, die große Inhomogenität ließe sich auch zum Beispiel durch die Annahme erklären, dass die Effektstärke in diesem Paradigma nicht unabhängig von der Stichprobengröße sei. Es ist mittlerweile zu beobachten, dass die klassischen RNG-Experimente mit Mittelwertsverschiebung als Zielgröße zunehmend anderen Studiendesigns, z.B. der Veränderung der Varianz der Verteilung, weichen.

Es sei hier noch auf das von Roger Nelson initiierte *Global Consciousness Project* verwiesen. Die Verantwortlichen dieses Projekts unterhalten an ca. 65 Orten in der Welt

kontinuierlich laufende RNGs. Die Daten dieser Zufallsgeneratoren können von jeder Person über das Internet abgerufen und ausgewertet werden. Die Idee ist, dass besondere Ereignisse, die eine weltweite Öffentlichkeit haben oder hervorrufen, sich im Sinne eines globalen Bewusstseins in den Daten widerspiegeln. In der Tat finden die Autoren immer wieder überraschende Abweichungen hinsichtlich prospektiver und retrospektiver Analysen (s. <http://noosphere.princeton.edu/>).

8.4.5 Präkognition

Beschreibung und alltagsweltliche Analogie

Seit jeher ist es der Wunsch der Menschen, Informationen über die Zukunft zu erhalten, und in vielen Kulturen sind die unterschiedlichsten Praktiken dokumentiert, dieses Ziel zu verwirklichen. Parapsychologische Experimente, die die Präkognitionshypothese testen, sind oft einfach zu realisieren. Meist muss die Versuchsperson bestimmen, welches Ereignis aus einem vorgewählten Set eintreten wird (*forced-choice*). Es lassen sich aber auch *free-response*-Experimente durchführen (Schmidt 2014). Da die Angabe der Versuchsperson zeitlich vor der Ermittlung des „Targets“ festgestellt wurde, sind hier die zu beachtenden Sicherheitsvorkehrungen gering. Wichtig ist lediglich die adäquate Randomisierung, sodass sich keine alternativen Möglichkeiten ergeben, das Ziel besser als unter Zufall erwartet zu bestimmen. Wie aber bereits angedeutet, lässt sich aus solchen Experimenten nicht zwingend die Existenz von Präkognition schlussfolgern, da auch alternative Hypothesen ohne zeitliches Paradox denkbar sind (z. B. Psychokinese).

Metaanalysen

Es liegen zwei Metaanalysen vor. Honorton und Ferrari (1989) fassten 309 Studien aus den Jahren 1935–1987 und damit die Daten von insgesamt 50 000 Versuchspersonen und mehr als 2 Mio. Ratedurchgängen zusammen. Sie berechneten eine Effektstärke von $r = 0,02$ ($d = 0,04$, $p = 6 \times 10^{-25}$). Storm und Kollegen veröffentlichten 2012 eine Metaanalyse der Studien der Jahre 1987–2010 mit strengeren Einschlusskriterien (Storm et al. 2012). Diese enthält ausschließlich *forced-choice*-Experimente und schließt Fallstudien und Versuche mit Tieren aus. Die mittlere Effektstärke von 25 Studien beträgt hier $r = 0,01$ ($d = 0,02$). Übereinstimmend lässt sich sagen, dass es sich bei einer soliden Datenbasis um **sehr kleine Effektstärken** handelt. In Einzelfällen sind diese wohl kaum relevant, aber bei häufig wiederholten Vorgängen (z. B. im Casino) könnten sich diese Effekte auch alltagsweltlich bemerkbar machen. Dies gilt immer unter der Annahme, dass der Effekt gleichbleibend und überall auftaucht und es sich nicht um singuläre Einzelereignisse handelt.

8.4.6 Presentiment

Beschreibung und alltagsweltliche Analogie

Sogenannte Presentiment-Studien (von engl. „*presentiment*“ = Vorahnung) stellen die Frage nach einer sehr kurzfristigen und nicht bewusst erlebten Präkognition. Die Idee ist, dass durch eine präkognitive Wahrnehmung gefährlicher Situationen wenige Sekunden vor deren Eintreffen evolutionsbiologisch ein klarer Vorteil entstehen könnte (s. Kap. 7). Entsprechend ist auch das Experimentalparadigma aufgebaut, das von Radin erstmals 1997 berichtet wurde (Radin

1997). In diesem Experiment wurden bei den Versuchspersonen die EDA als Indikator autonomer Erregung abgeleitet. Den Versuchspersonen wurden wie in einem klassischen psychophysiologischen Experiment entweder neutrale oder erregende bzw. aversive Stimuli präsentiert. Es zeigte sich nach der Stimuluspräsentation die typische zu erwartende Erregungskurve in Abhängigkeit von der Stimulusqualität. Radin wertete allerdings die EDA-Daten unmittelbar vor der Stimuluspräsentation aus und konnte zeigen, dass sich hier bereits signifikante Unterschiede ergaben. Vor der Präsentation eines erregenden Stimulus zeigte sich eine leichte Erhöhung des autonomen Erregungsniveaus, die sich bei neutralen Stimuli so nicht fand.

Methodische Schwierigkeiten und Metaanalyse

Das Presentiment-Experiment birgt einige methodische Schwierigkeiten. So ist zum Beispiel der Zeitpunkt, zu dem der Effekt vermutet und damit gesucht wird, für ein konfirmatorisches Experiment zentral. Soll dieser nun 1, 3 oder 5 Sekunden vor der Stimuluspräsentation liegen? Das andere Problem, das in der Literatur bereits recht umfangreich thematisiert wurde, ist der Ausschluss von Erwartungseffekten durch eine adäquate Randomisierung. Hat die Versuchsperson z. B. aufgrund der Zufallsabfolge bereits 5 neutrale Stimuli gesehen, wird sie vermutlich stark davon ausgehen, dass der nächste ein affektiv aufgeladener sein könnte, und mit einer entsprechenden Erregung vor der Stimuluspräsentation reagieren, die dann als Presentiment fehlinterpretiert werden könnte. Ein solches Artefakt kann allerdings durch geeignete Randomisations- und Analysemethoden verhindert werden.

Hinsichtlich der Ergebnisse zeigten sich seit 1997 einige signifikante, aber auch nicht signifikante Befunde. Das Paradigma und die abhängige psychophysiologische Variable wurden mehrfach variiert, z. T. wurden auch bereits bestehende Datensätze aus ganz normalen psychophysiologischen Experimenten auf die Existenz einer stimulusabhängigen Variation vor der Stimuluspräsentation untersucht. Mossbridge et al. (2012) führten eine Metaanalyse mit insgesamt 26 Experimenten durch und fanden eine Effektstärke von $r = 0,21$ ($d = 0,43$), die signifikant war ($p = 2 \times 10^{-12}$) und auf einem homogenen Datensatz beruhte. Die Analyse des sogenannten *Funnelplots* ergab, dass vermutlich nur wenige der unpublizierten Studien (Publikationsbias) nicht gefunden wurden und diese die Gesamteffektstärke nur unwesentlich verringern würden. Somit kann hier von einem sehr soliden Gesamtbefund ausgegangen werden.

Die Experimente von Daryl Bem

Eine spezielle Gruppe von Experimenten, die an die Presentimentforschung angelehnt ist, wurde von Bem (2011) veröffentlicht. Bem nutzte zahlreiche Standardexperimente der kognitiven Sozialpsychologie und änderte diese in Präkognitionsexperimente, indem er den Ablauf des Versuches umkehrte. Ein Beispiel: In affektiven Priming-Experimenten wird Versuchspersonen z. B. für einen sehr kurzen Zeitraum ein positives oder ein negatives Wort präsentiert (*prime*), dann taucht ein Bild auf, das so schnell wie möglich als angenehm oder unangenehm bewertet werden soll. Dies gelingt besser, wenn das Bild mit dem *prime* kongruent ist. In Bems *Retroactive Priming Experiment* muss nun zuerst das Bild eingeschätzt werden und erst danach wird das Wort von einem Zufallsgenerator ausgewählt und prä-

sentiert. Bem führte hier beide Versionen durch; das Standardpriming ergab eine Effektstärke von $d = 0,45$, das retroaktive, das durch *Psi* zu erklären wäre, ergab $d = 0,25$ ($p = 0,006$). Insgesamt führte Bem neun solcher transformierten Experimente mit mehr als 1 000 Versuchspersonen durch, von denen acht signifikante Resultate erbrachten; die Gesamteffektstärke lag bei $d = 0,22$. Die Publikation löste ein großes Echo aus; zurzeit werden die Experimente in vielen Labors wiederholt.

8.4.7 Traum-Telepathie

Beschreibung und alltagsweltliche Analogie

In den großen parapsychologischen Fallsammlungen zeigt sich ein umfangreicher Anteil an Berichten über *telepathisches* oder *präkognitives Träumen*. Viele Menschen, die Rat oder Hilfe wegen außergewöhnlicher Erfahrungen bei Beratungsstellen suchen, berichten über sehr detaillierte Träume, die sich später bewahrheiteten. Die Schauspielerin Christine Mylius beispielsweise hinterlegte in den 1950er-Jahren Hunderte solcher Traumberichte beim Institut für Grenzgebiete der Psychologie und Psychohygiene e. V. (IGPP) in Freiburg, von denen sich etliche später bewahrheiteten (vgl. Kap. 10). Diese vielfältigen Berichte waren Anlass, das Phänomen auch experimentell im Labor zu untersuchen. Am Maimonides Medical Center in Brooklyn wurden von 1966–1972 unter der Leitung von Stanley Krippner, Montague Ullman und Charles Honorton knapp 400 solcher Einzelexperimente mit unterschiedlichen Versuchsdesigns durchgeführt. Beim Standardvorgehen wurden einer schlafenden Versuchsperson das EOG (= Augenbewegung) und das EEG abgenommen. Der Versuchsleiter konnte anhand

dieser Daten erkennen, wann sich die Versuchsperson in einem REM-Zustand befand, und signalisierte dies an eine zweite Versuchsperson (Sender/in), die daraufhin ein Bild aus einem verschlossenen Umschlag nahm und versuchte, dies der schlafenden Person zu „senden“. Am Ende der REM-Phase weckte der Versuchsleiter die träumende Versuchsperson und bat sie um ihren Traumbericht. Anschließend bekamen unabhängige Gutachter/innen diesen Traumbericht zusammen mit acht Bildern, von denen nur eines das im Experiment verwendete darstellte. Die Gutachter/innen mussten die Bilder gemäß ihrer Passung zum Traumbericht in eine Rangreihenfolge bringen. War das wirklich verwendete Bild unter den ersten vier, wurde es als Treffer gewertet (Zufallserwartung 50 %).

Metaanalysen

Aus den **Maimonides-Experimenten** wurden mehrere Zusammenfassungen gerechnet (für Details s. Schmidt 2014), die aber alle nicht die formalen Kriterien einer methodisch korrekten Metaanalyse erfüllen. Es ergab sich insgesamt in 379 Durchgängen eine Trefferrate von 61 % ($p < 0,001$). Sherwood und Roe (2003) rechneten eine Metaanalyse für die Maimonides-Studien sowie für die in der Nachfolge von anderen Labors durchgeführten Experimente. Sie fanden eine Effektstärke für die Maimonides-Studien von $r = 0,33$ ($d = 0,70$) und für die Post-Maimonides-Experimente von $r = 0,14$ ($d = 0,28$, $n = 604$ Einzelexperimente aus 21 Studien) (vgl. Kap. 10). Die beiden Datensätze unterscheiden sich signifikant voneinander, die berichtete Datenlage ist aber eindeutig positiv hinsichtlich der untersuchten Phänomene.

8.4.8 Remote Viewing

Eine interessante Serie von *free-response*-Studien stellen die sogenannten *Remote Viewing*-Experimente dar, die seit Anfang der 1970er-Jahre durchgeführt wurden. Bei diesen Experimenten versuchen Personen Informationen über einen entfernten Ort zu erlangen. Sie fertigen dafür Zeichnungen an und liefern verbale Beschreibungen (ein Beispiel zeigt Abb. 8-4).

Diese Dokumente werden dann ebenfalls von unabhängigen Gutachter/innen hinsichtlich der Passung eingestuft. In den einzelnen Experimenten gibt es große Variationen. Mal begibt sich einer der Versuchsleiter/innen zu dem jeweiligen Zielort, mal wird der Versuchsperson aber auch nur eine Koordinate gegeben oder sie wird gebeten, den Ort einer vermissten Person oder eines vermissten Flugzeuges genauer zu bestimmen oder zu beschreiben. Diese Experimente wurden bereits ab Mitte der 1970er-Jahre von der amerikanischen Regierung gefördert und daraufhin im Modus der Geheimhaltung weitergeführt. 1995 wurden große Teile des Materials freigegeben. Eine formale Metaanalyse wurde nie durchgeführt und ist vermutlich auch aufgrund der Heterogenität und der mangelhaften Publikation der einzelnen Studien/Versuche nur schwer möglich. Das von der amerikanischen Regierung unterstützte Programm wurde 1996 von Jessica Utts evaluiert. Sie errechnete für insgesamt 1215 Durchgänge Effektstärken von $d = 0,21$ bzw. $0,23$. Auffallend war dabei, dass nur eine sehr kleine Population von Versuchspersonen (ca. 1 % aller Untersuchten) wiederholt zu guten Ergebnissen gelangte. Im Unterschied zu vielen anderen Paradigmen scheint es hier begabte und nichtbegabte Versuchspersonen zu geben. Der Bericht von Utts geht nicht davon aus, dass diese Fähigkeit trainierbar ist. In einer

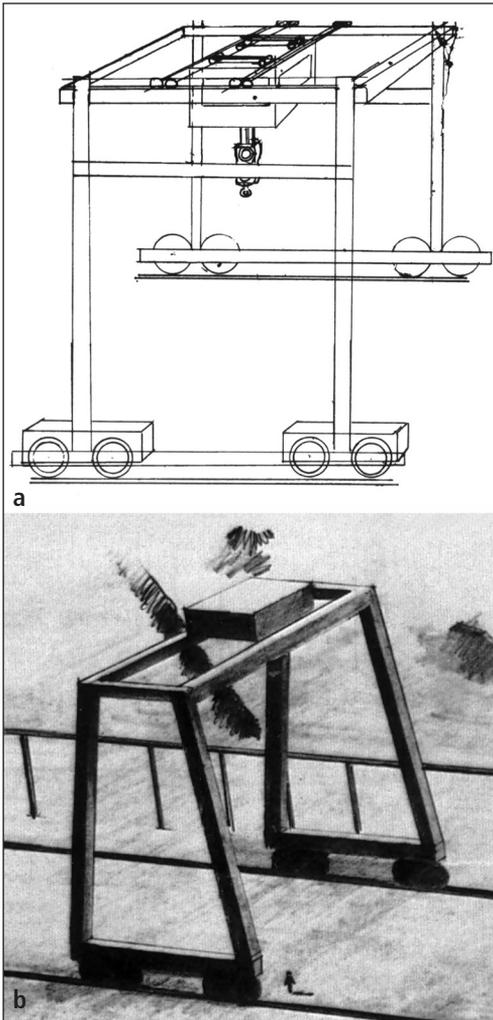


Abb. 8-4 a Zeichnung eines Kranes, die von einem Remote Viewer 1974 angefertigt wurde. Aufgabe war es herauszufinden, was in einem russischen Atomlabor stattfindet. Der Remote Viewer erhielt dafür lediglich die Koordinaten der Anlage; b vom CIA angefertigte Zeichnung, die auf einer Satellitenaufnahme des Geländes beruht (aus: „Remote Viewing at Stanford Research Institute in the 1970s: A Memoir“ von Russell Targ, *Journal of Scientific Exploration*, Vol. 10, S. 77–88, 1996, veröffentlicht durch die Society for Scientific Exploration; www.scientificexploration.org).

kritischen Evaluation des Berichtes von Hyman (1996) werden diese Befunde jedoch als zu verfrüht eingeschätzt und weitere Replikationen gefordert.

8.5 Zusammenfassung

Fasst man die qualitativ angemessenen Metaanalysen zu den oben berichteten Paradigmen zusammen, so ergeben sich für die hier berichteten Experimente durchweg signifikante und solide **Effektstärken** im Bereich von $d = 0,11$ (DMILS) bis zu $d = 0,43$ (Presentiment), die Werte für das Ganzfeldparadigma ($d = 0,28$ – $0,29$) und für die Studien zu Bem ($d = 0,29$) liegen dazwischen. Bei den Bereichen Remote Viewing und Blickwahrnehmung steht eine adäquate Metaanalyse noch aus. Zwei Paradigmen fallen aus dem Rahmen. Zum einen die Präkognitionsexperimente mit sehr kleinen Effektstärken (von $d = 0,02$ – $0,04$) und zum anderen Mikro-PK, wo die Existenz des Effektes auf der Basis der vorliegenden Daten nicht als gesichert angesehen werden kann.

Aus forschungsmethodischer Sicht weisen die hier berichteten Effektstärken auf zwei wichtige Fakten hin. Zum einen sind so gut wie alle der bisher durchgeführten Experimente zu klein angelegt, um Effektstärken dieser Größenordnung zuverlässig zu zeigen. Das oft berichtete Replikationsproblem der Parapsychologie könnte sich eventuell durch diese sogenannte **mangelnde statistische Power** erklären lassen. Zum anderen zeigt sich aber auch mehrfach, dass das Problem der nicht berichteten (negativen) Studien – als **Publikationsbias** – hier ein wichtiger Faktor ist, der für jeden einzelnen Experimentalbereich kritisch untersucht und bewertet werden muss.

Hinsichtlich der Gesamtbewertung dieser Befunde ergeben sich zwei Perspektiven.

Da ist zunächst die Ebene, dass die hier untersuchten Effekte im Widerspruch zu unserem momentanen wissenschaftlichen Weltbild stehen. Aus dieser Perspektive spielt die Größe der Effektstärke keine Rolle, denn auch ein sehr kleiner Effekt, der hochsignifikant ist, steht im direkten Widerspruch zu dem, was unter der momentan gültigen Theorie zu erwarten ist. Hier kann man auf Basis der Befunde klar und deutlich auf die Existenz einer (oder mehrerer) **Anomalien** hinweisen, die Anlass zu weiteren Untersuchungen und vor allem zu theoretischen Modellierungen geben (vgl. z. B. Kap. 7). Die zweite Ebene bezieht sich auf die **Alltagsrelevanz** der Effekte. Prinzipiell muss bei Effekten, die kleiner sind als $d = 0,25-0,30$, davon ausgegangen werden, dass sie im Alltag nicht bedeutsam sind. Diese Einschätzung beruht aber immer auf der Idee der Mittelung und des gleichmäßigen Auftretens der statistischen Effekte. Tritt der Effekt aber zum Beispiel aufgrund bestimmter uns unbekannter Bedingungen nur jedes zehnte oder zwanzigste Mal auf, dann wäre er in diesem Fall zehn oder zwanzigmal so groß und damit deutlich alltagsrelevant. Diese Überlegung deckt sich mit der Sichtweise, dass Telepathie oder Interaktion über die Ferne in unserem gewöhnlichen Alltagsleben eigentlich keine Rolle spielen, aber dann vereinzelt doch Berichte über ungewöhnliche, schwer zu erklärende Wahrnehmungen auftauchen. Was die Bedingungen für ein solches Auftreten sind, hat die experimentelle Forschung der letzten 130 Jahre trotz vielfacher und intensiver Bemühungen nicht herausfinden können.

Zur vertiefenden Lektüre

- Holt N, Simmonds-Moore C, Luke D, French C. *Anomalistic Psychology*. Houndmills, Basingstoke: Palgrave Macmillan 2012.
- Irwin HJ, Watt C. *An introduction to parapsychology*. 5th ed. Jefferson, NC; London: McFarland 2007; darin Kap. 4: Experimental research on extrasensory perception.
- Radin DI. *The Conscious Universe. The Scientific Truth of Psychic Phenomena*. New York: Harper-Edge 1997.
- Schmidt S. *Experimentelle Parapsychologie. Eine Einführung*. Würzburg: Ergon 2014.

Literatur

- Bem DJ. Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology* 2011; 100(3): 407–25.
- Bem DJ, Honorton C. Does PSI exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin* 1994; 115: 4–18.
- Bösch H, Steinkamp F, Boller E. Examining Psychokinesis: The interaction of human intention with random number generators – a meta-analysis. *Psychological Bulletin* 2006; 132: 497–523.
- Colborn M. The decline effect in spontaneous and experimental psychical research. *Journal of the Society for Psychical Research* 2007; 71(1): 1–21.
- Honorton C, Ferrari DC. „Future Telling“: A meta-analysis of forced-choice precognition experiments, 1935–1987. *Journal of Parapsychology* 1989; 53: 281–308.
- Hyman R. Evaluation of a program on anomalous mental phenomena. *Journal of Scientific Exploration* 1996; 10: 31–58.
- Jahn RG, Dunne BJ, Nelson RD, Dobyys YH, Bradish GJ. Correlations of random binary sequences with pre-stated operator intention: a review of a 12-year program. *Journal of Scientific Exploration* 1997; 11(3): 345–67.

- Jahn RG, Dunne BJ, Bradish GJ, Dobyns YH, Lettieri A, Nelson RD et al. Mind/Machine Interaction Consortium: PortREG replication experiments. *Journal of Scientific Exploration* 2000; 14: 499–555.
- Lucadou W v. Psyche und Chaos. Theorien der Parapsychologie. Frankfurt: Insel 1995.
- Milton J, Wiseman R. Does psi exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin* 1999; 125: 387–91.
- Mossbridge J, Tressoldi PE, Utts J. Predictive physiological anticipation preceding seemingly unpredictable stimuli: a meta-analysis. *Frontiers in Perception Science* 2012, doi: 10.3389/fpsyg.2012.00390. www.frontiersin.org/Perception_Science/10.3389/fpsyg.2012.00390/full (20 August 2013).
- Palmer J. ESP in the Ganzfeld: Analysis of a Debate. *Journal of Consciousness Studies* 2003; 10(6–7): 51–68.
- Pratt JG, Rhine JB. Extra-sensory perception after sixty years. In: Holt H (ed). *Extra-Sensory Perception After Sixty Years: A Critical Appraisal of the Research in Extra-Sensory Perception*. New York: Henry Holt and Company 1940.
- Radin D. Unconscious perception of future emotions: an experiment in presentiment. *Journal of Scientific Exploration* 1997; 11(2): 163–80.
- Radin D. The sense of being stared at: a preliminary meta-analysis. *Journal of Consciousness Studies* 2005; 12(6): 95–100.
- Radin D, Ferrari DC. Effects of consciousness on the fall of dice: a meta-analysis. *Journal of Scientific Exploration* 1991; 5: 61–84.
- Radin D, Nelson RD. Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics* 1989; 19: 1499–514.
- Radin D, Nelson R, Dobyns Y, Houtkooper J. Reexamining psychokinesis: comment on Bösch, Steinkamp, and Boller (2006). *Psychological Bulletin* 2006; 132 (4): 529–32.
- Schlitz MJ, Wiseman R, Radin DI, Watt CA. Of two minds: skeptic-proponent collaboration within parapsychology. *The Parapsychological Association, Inc. 48th Annual Convention, August 11th–15th, 2005*. Durham, NC: The Parapsychological Association 2005; 171–77.
- Schmidt S. Bohrende Blicke? Stechende Blicke? Das Phänomen der „Blickwahrnehmung“ wissenschaftlich untersucht. *Zeitschrift für Anomalistik* 2008; 8: 32–54.
- Schmidt, S. Die Fliege des Aristoteles. Bemerkungen zur Anomalistik und eine Forschungsübersicht zum Zusammenhang zwischen Meditation und Psi. *Zeitschrift für Anomalistik* 2012a; 12: 158–78.
- Schmidt S. Can we help just by good intentions? A meta-analysis of experiments on distant intention effects. *Journal of Alternative and Complementary Medicine* 2012b; 18(6): 529–33.
- Schmidt S, Schneider R, Utts JM, Walach H. Distant intentionality and the feeling of being stared at – two meta-analyses. *British Journal of Psychology* 2004; 95: 235–47.
- Sheldrake R. The sense of being stared at part 1: Is it real or illusory? *Journal of Consciousness Studies* 2005; 12(6): 10–31.
- Sherwood SJ, Roe CA. A review of dream ESP studies conducted since the Maimonides dream ESP studies. *Journal of Consciousness Studies* 2003; 10(6–7): 85–109.
- Storm L, Ertel S. Does psi exist? Comments on Milton and Wiseman's (1999) meta-analysis of ganzfeld research. *Psychological Bulletin* 2001; 127: 424–33.
- Storm L, Tressoldi PE, Di Risio L. Meta-analysis of ESP studies, 1987–2010: assessing the success of the forced-choice design in parapsychology. *Journal of Parapsychology* 2012; 76 (2): 243–73.
- Storm L, Tressoldi PE, Di Risio L. Meta-analysis of free-response studies, 1992–2008: assessing the noise reduction model in parapsychology. *Psychological Bulletin* 2010; 136(4): 471–85.
- Williams BJ. Revisiting the ganzfeld ESP debate: A basic review and assessment. *Journal of Scientific Exploration* 2011; 25(4): 639–61.
- Yong E. Replication studies: bad copy. *Nature* 2012; 485 (7398): 298–300. doi:10.1038/485298a.