

Fortgesetzte Diskussionen zu früheren Beiträgen

Kommentar zum Aufsatz von Martina Belz und Thomas Berger „Psychisches Wohlbefinden, Außergewöhnliche Erfahrungen und Emotionsregulation“

In: Zeitschrift für Anomalistik 8 (2008), 118-134

WOLFGANG AMBACH¹

Gemischte Ratsuchende – gemeinsame Pläne

Ratsuchende, die von eigenen außergewöhnlichen Erfahrungen (AgE) berichten, landen, wenn das Gegenüber der gängigen psychiatrischen und psychotherapeutischen Lehre folgt, fast zwangsläufig in einer Art Niemandsland: Die unerklärten Phänomene, die berichtet werden, gibt es nach gängiger Lehre nicht. Gleichzeitig genügen allerdings die vorgebrachten Schilderungen vielfach den Kriterien, die sich in diagnostischen Manualen für psychische Störungen finden. Hieraus ergibt sich bevorzugt eine Einordnung des Geschilderten, die entweder bagatellisiert / trivialisiert („... wird er/sie sich halt eingebildet haben ...“) oder pathologisiert („... typisches Symptom von paranoider Störung: ... KRANK ...“). Die entsprechende Hilfestellung, die das System in Folge anbietet, wird dem/der Ratsuchenden dann häufig nicht gerecht und geht ins Leere. Auch bilden die Personen, die AgE schildern, nach dieser Logik keine wissenschaftlich untersuchenswerte Einheit, sondern gehen in verschiedenen psychopathologischen Kategorien auf.

Zentrales Anliegen einer auf AgE spezialisierten Beratungsarbeit ist eine psychotherapeutische Hilfestellung, die die geschilderten Erfahrungen (*auch*) jenseits der gängigen diagnostischen Kategorien und jenseits der Frage der „Tatsächlichkeit“ des Erlebten betrachtet. Der hieraus resultierende Ansatz beschäftigt sich vorrangig mit der individuellen Bedeutung und Entwicklung der geschilderten Erfahrungen und der individuellen Lebenssituation und -geschichte der Ratsuchenden. Die Beratungsabteilung des IGPP ist hier wegweisend. Eines der zentralen Anliegen der hieran angeknüpften Forschung zu AgE ist die Beschreibung der AgE-Klientel als eigenständige, abgrenzbare und nicht zwangsläufig

1 Dr. med. Wolfgang Ambach leitet die Forschungsgruppe Klinische und Physiologische Psychologie, Institut für Grenzgebiete der Psychologie und Psychohygiene (IGPP), Freiburg i.Br.

fig mit klinischen Kategorien assoziierte Personengruppe. Eine solche möglichst detaillierte Beschreibung sollte auf verschiedenen differentialpsychologischen Dimensionen und mit verschiedensten Methoden vollzogen werden, um der wissenschaftlichen Beschäftigung mit dieser Personengruppe überhaupt einen Platz im wissenschaftlichen Austausch in einer breiteren Community zu verschaffen.

Genau hierzu leistet die Analysemethode, die von Belz und Berger präsentiert wird, einen hervorragenden Beitrag. Die Annahme, dass hinter der Heterogenität der Erscheinungsformen von AgE eine psychische Tiefenstruktur besteht, die zwischen den Personen weit weniger heterogen ist, liegt diesem Ansatz zugrunde. Sie motivierte die wissenschaftliche Untersuchung mehrerer Personen, die zunächst ihre Ratsuche am IGPP als einzige Gemeinsamkeit hatten, und die Beschreibung eines aus den Schilderungen erschlossenen gemeinsamen psychodynamischen Musters. Die ausführlich beschriebene Vorgehensweise, einschließlich zahlreicher plastischer Beispiele im Ergebnisteil, ist gut nachvollziehbar und transparent. Es wird ohne weiteres deutlich, wie aus den aufgezeichneten Gesprächen die einzelnen Reaktionen und Pläne abstrahiert werden und wie hieraus in der Gruppenanalyse ein prototypisches Muster der Emotionsregulation erschlossen wird. Die hinzugefügte methodenkritische Analyse der Inter-Rater-Übereinstimmung belegt für ein qualitatives Analyseverfahren gute Werte. Die Ergebnisse sind in gutem Einklang mit den von Tölle (2003) gefundenen prototypischen Planstrukturen bei AgE-Ratsuchenden, werden von Belz und Berger aber stärker im Hinblick auf die Emotionsregulation differenziert. Wie in der zitierten Vorstudie zeigt sich im Endeffekt auch hier ein psychodynamisch beschreibbares, differentialpsychologisch bedeutsames Merkmal, das in der Gruppe der Ratsuchenden hervortritt und damit – was anhand einer Vergleichsgruppe noch zu verifizieren wäre – jenseits von psychopathologischen Kriterien zu deren Charakterisierung beiträgt.

Im Gegensatz zu der klar beschriebenen Vorgehensweise lassen die Autoren allerdings einige Unklarheit über die in die Studie eingeschlossenen Teilnehmer und die von ihnen berichteten AgE. Auch wenn für die Analyse keine Differenzierung der Teilnehmer nach der Art der beschriebenen Erfahrungen zu erfolgen hatte, genügt der bloße Hinweis auf eine bestehende Heterogenität der Erfahrungen dem Leser nicht. Die Frage, worum es in den einzelnen Schilderungen überhaupt geht, bleibt zunächst ganz offen und wird auch später in vereinzelt Beispielsituationen nur gestreift. Am Ende erfährt der Leser, dass die Beschreibungen der AgE durch die Ratsuchenden „allesamt relativ diffus“ geblieben sind, und im Schlusssatz, dass sie „vor allem Intuition und Ahnungen“ berichtet haben. Wenn man berücksichtigt, dass offensichtlich der Akt des Ratsuchens das wesentliche Einschlusskriterium war, dann erfährt der Leser letztlich kaum, welche Personen hier unter-

sucht wurden. Dies ist gerade insofern bedauerlich, als der dargestellte Beratungs- und Forschungsansatz hochgradig individuell ausgelegt ist. Am Ende bilden die Ergebnisse der Plananalyse das hauptsächliche Beschreibungsmerkmal für die Teilnehmer; es wird schwierig sein, extrapolierende Schlussfolgerungen für die dahinterstehende breitere Gruppe von Personen mit AgE-Schilderungen anzustellen.

Dieser Umstand erscheint noch problematischer, wenn man liest, dass bei *allen* Teilnehmern eine „sehr schwierige und teils traumatisierende Lebensgeschichte“ erschlossen wurde. Freilich lassen sich angesichts der gebotenen Kürze der Darstellung kaum anamnestische Angaben unterbringen; gerade hier werden aber Vorgeschichten angedeutet, die erhebliche Relevanz für die gefundenen Ergebnisse haben dürften (was die Autoren ja auch diskutieren).

Letztlich schränken die zu knappen Angaben über Teilnehmer, Vorgeschichte und berichtete AgE die Aussagekraft der Studienergebnisse, insbesondere ihre Extrapolierbarkeit auf einen größeren Personenkreis, ein.

Das diskutierte Modell der mehr oder weniger gelingenden Integration diskrepanter Wahrnehmungen erscheint sehr interessant und plausibel. Die prinzipiell bidirektional denkbare Interaktion zwischen dem Auftreten von AgE und dem individuellen Umgang mit Emotionen weiter zu klären, wäre von großem Interesse. Freilich wäre es auch für die Einbindung der vorliegenden Studienergebnisse in einen solchen breiteren Kontext wünschenswert, den zeitlichen Verlauf der individuellen AgE der Studienteilnehmer betrachten zu können, was hier nicht möglich ist. Insbesondere die beschriebene Entwicklung der „Versklavung“ des psychischen Gesamtsystems, ein im Bestätigungsfall sicher recht implikationsreiches Modell, wäre vermutlich nur anhand minutiöser Analysen einzelner Anamnesen nachvollziehbar, die den hier gesetzten Rahmen allerdings sprengen würden.

Die angeführten Kommentar- und Kritikpunkte sollen nicht verdecken, dass der von Belz und Berger eingeschlagene Ansatz als sehr vielversprechend betrachtet wird. Insbesondere lässt er erwarten, dass trotz der Heterogenität der AgE-Schilderungen weiterreichende differentialpsychologische Aussagen über eine Personengruppe getroffen werden können, die durch a) eigene außergewöhnliche Erfahrungen und b) ein überdurchschnittlich paranormales Überzeugungssystem definiert wird. Die grundlegende These, dass hinter den heterogenen AgE-Schilderungen eine weit weniger heterogene Tiefenstruktur besteht, dass es also differentialpsychologische Gemeinsamkeiten zwischen Personen gibt, die unterschiedliche AgE berichten, bedarf weiterhin der Untermauerung. Es ist sehr wertvoll, dass an einer solchen, nicht primär psychopathologisch orientierten Beschreibung von Persönlichkeitscharakteristi-

ka multiperspektivisch gearbeitet wird. Dem hier aufgezeigten psychodynamischen Ansatz dürfte dabei auch weiterhin eine bedeutende Rolle zukommen, vor allem dann, wenn es gelingt, anhand einer Vielzahl von Einzelfällen, in denen die individuelle Entwicklung der geschilderten Erfahrungen nachgezeichnet werden kann, sowohl spezifische disponierende Personenmerkmale als auch konstellative Entstehungsbedingungen von AgE zu definieren. Das Verständnis der berichteten Phänomene wird hiervon ebenso profitieren wie die Grenz-ziehung zwischen berichtetem Phänomen und Symptom.

Literatur

Tölle, P. (2003). *Typische Planstrukturen von Menschen mit außergewöhnlichen Erfahrungen*. Unveröffentl. Diplomarbeit. Freiburg i.Br.: Universität Freiburg.

Autorenantwort:

MARTINA BELZ & THOMAS BERGER²

Auf dem Weg zu vertieften Analysen

Wir stimmen Ambach uneingeschränkt zu, wenn er die knappen Angaben über die sechs untersuchten Ratsuchenden, u.a. die fehlenden Angaben zu deren Vorgeschichte, kritisiert. In unserem Artikel (Belz & Berger, 2008) wurde aus Platzgründen auf eine ausführliche Darstellung der Stichprobe zugunsten der Beschreibung der Methodik verzichtet.

Den hier zur Verfügung gestellten Platz nutzen wir gerne für einen Nachtrag und gehen im Folgenden auf zentrale Merkmale der Stichprobe ein. Wie in unserem Beitrag beschrieben, handelt es sich bei den untersuchten Personen um sechs zufällig ausgewählte Ratsuchende der Abteilung Information und Beratung für Menschen mit außergewöhnlichen Erfahrungen (AgE) des Instituts für Grenzgebiete der Psychologie und Psychohygiene e.V. (IGPP) in Freiburg, die zwischen 2002 und 2004 das Beratungsangebot in Anspruch genom-

2 Abteilung Klinische Psychologie und Psychotherapie der Universität Bern, Gesellschaftsstrasse 49, CH-3012 Bern, E-Mail: martina.belz@psy.unibe.ch.

men haben. Die Ratsuchenden sind zwischen 24 und 50 Jahre alt, drei leben allein, eine in Trennung, zwei in festen Partnerschaften. Vier Personen haben eigene Kinder. Der Bildungsstand ist relativ hoch, vier der Ratsuchenden haben ein Studium absolviert. An lebensgeschichtlich relevanten Ereignissen berichten vier der sechs Ratsuchenden von schwerwiegenden, teils lebensbedrohlichen Erkrankungen bzw. Unfällen in Kindheit und Jugend sowie gravierenden Verlusterlebnissen (Autounfälle, Beinaheertrinken, Tod eines Kindes und Fehlgeburt, früher Tod der Eltern, gleichzeitiger Tod von mehreren nahestehenden Bezugspersonen). Alle Ratsuchenden der Stichprobe befinden sich zum Zeitpunkt der Beratung und Untersuchung in einer beruflichen und persönlichen Umbruch- und Neuorientierungsphase, wobei die berichteten AgE eine zentrale Rolle einnehmen (werden z.B. als Zeichen für die Richtung der beruflichen Neuorientierung eingeordnet). Mit diesen soziodemographischen Merkmalen repräsentieren die ausgewählten Fälle die in grösseren Stichproben bereits gefundenen typischen soziodemographischen Merkmale von Ratsuchenden des IGPP recht gut (vgl. Belz-Merk & Fach, 2005).

Bevor wir Genaueres zur Phänomenologie der AgE der Ratsuchenden der Stichprobe berichten, stellen wir zur besseren Einordnung im Folgenden den aktuellen Stand der Auswertung der Häufigkeiten dar, mit denen Klientinnen und Klienten spezifische Phänomene schildern. Seit 1996 werden am IGPP im Rahmen der routinemässigen Dokumentation der Beratungsfälle vor allem auch genaue Beschreibungen der Phänomenologie der AgE, zur Soziodemographie und psychischen Auffälligkeit der Ratsuchenden erhoben. Die bis jetzt mittels einer Hauptkomponentenanalyse ausgewerteten N=1508 Fälle ergeben sechs verschiedene AgE-Formenkreise: 1. *Spuk und Erscheinungen* (externale Anomalien, wie das Verschwinden oder Auftauchen von Gegenständen, akustische Phänomene, insbesondere Mimikrygeräusche [z.B. Klopfen, Schritte, Stimmen] ohne identifizierbare Quelle, visuelle Eindrücke [Lichterscheinungen, Gestalten usw.]), 2. *Internale Präsenz und Beeinflussung* (Somatische Phänomene [Energieströme, Schmerzen], für die es keine objektive bzw. medizinische Grundlage zu geben scheint, das Hören von Stimmen „im Kopf“), 3. *Automatismen und Mediumismus* (psychophysische Dissoziation bei Automatischem Schreiben, Channeling etc.), 4. *Externale Präsenz und Alptrüben* („Spüren einer Anwesenheit“ in der physikalischen Außenwelt lokalisiert, gelegentlich zusätzlich noch mit einer psychophysischen Dissoziation, nämlich der Unfähigkeit, den Körper zu bewegen [Schlafparalyse], verknüpft), 5. *Außersinnliche Wahrnehmung „ASW“* (sinnvolles, aber angeblich nicht kausal vermitteltes Zusammentreffen z.B. von internalen Wahrnehmungen mit externalen Ereignissen oder Wahrnehmungen anderer Personen wie bei Telepathie, Hellsehen oder Präkognition) und 6. *Sinnvolle Fügungen* (sinnvolle Zusammenhänge zwischen Ereignissen

ohne befriedigende kausale Erklärung) (Belz-Merk & Fach, 2005; Belz 2009; Belz & Fach, im Druck).

Mittels Clusteranalyse lassen sich dann Gruppen von Ratsuchenden mit typischen Kombinationen von Phänomenen und AgE-Formenkreisen identifizieren. Sechs der gefundenen Cluster zeichnen sich dadurch aus, dass die zugeordneten Personen einen einzelnen AgE-Formenkreis gemeinsam haben. Das ist bei ca. 50% der Ratsuchenden der Fall. Im Einzelnen sind das der *Spuktyp*, *ASW-Typ*, *Externaler Präsenztyp*, *Internaler Präsenztyp*, *Fügungstyp* und *Medialer Typ*. Darüber hinaus gibt es drei Mischtypen, die durch das Auftreten von jeweils zwei Formenkreisen charakterisiert sind: *Internaler ASW-Typ* (Internale Präsenz und ASW), *Spuk-ASW-Typ* (Spuk und ASW) und *Internaler Spuktyp* (Internale Präsenz und Spuk). Vier der neun Cluster (Fügings-, Medialer, Externaler Präsenz- und Internaler ASW-Typ) enthalten zwar kleinere Subgruppen von Klienten und Klientinnen mit zusätzlichen Phänomenen anderer Formenkreise, insgesamt bietet die Aufteilung jedoch eine gute Heuristik.

Was die Phänomenologie der AgE in der Stichprobe unserer qualitativen Studie angeht, repräsentieren zwei Ratsuchende einen reinen Typus mit Phänomenen, die zum Formenkreis der sinnvollen Fügungen gehören. Die übrigen vier berichten relativ heterogene AgE, die sich aber vor allen Dingen den Bereichen der Außersinnlichen Wahrnehmung und der internalen Präsenz und Beeinflussung (insbes. somatische Phänomene, innere Bilder und Stimmen) zuordnen lassen, teilweise gemischt mit Spuk und Erscheinungen sowie einzelnen dissoziativen Phänomenen (wie z.B. Automatismen, Außerkörperliche Erfahrungen und Depersonalisation).

Aufgrund der kleinen Stichprobe und der Heterogenität der Erfahrungen, die zudem bei den Ratsuchenden überwiegend in gemischter Form auftraten, schien es wenig sinnvoll Untergruppen zu bilden. Auch in den beiden Diplomarbeiten von Toelle (2003) und Spitz (2005) lassen sich für die verschiedenen Phänomenbereiche keine unterschiedlichen motivationalen Strukturen finden, so dass die gewählte Darstellung gerechtfertigt schien.

Bisherige quantitative Analysen (Belz-Merk & Fach, 2005; Fach & Atmanspacher, 2006) des umfassenden Datenpools aus der Basisdokumentation, in der vor allem Oberflächenmerkmale erfasst werden (Soziodemographie, Phänomenologie der AgE, Rating der klinischen Auffälligkeit), zeigen zwar Unterschiede zwischen den verschiedenen Typen. So sprechen die Ergebnisse für eine besonders hohe Belastung und psychische Auffälligkeit bei Klientinnen und Klienten mit internalen Phänomenen (Internaler Präsenz- und Internaler Spuktyp) und Menschen mit mediumistischen Phänomenen. Die wenigsten Anzeichen für das Vorliegen einer psychischen Störung zeigen der Spuk-ASW- und der Spuktyp. Bei

letzterem finden sich auch die wenigsten Vorerfahrungen mit der Psychiatrie. Unterschiede in der Häufigkeit, mit der verschiedene Erfahrungen berichtet werden, finden sich auch hinsichtlich Alter, Geschlecht, beruflicher Situation, sozialer Situation und bezüglich verschiedener Belastungen (sozial, psychisch, körperlich).

Unser explorativ-qualitativer Ansatz geht jedoch ein wenig in die Breite (untersucht wurden nicht ein/e, sondern sechs IGPP-Klientinnen bzw. -Klienten) und mit Hilfe der Plananalyse relativ stark in die Tiefe. Der Kommentar von Ambach ist für uns gerade deshalb hilfreich und prüfenswert, weil er – wie vielleicht eher zu erwarten gewesen wäre – nicht mehr Breite, sondern noch mehr Tiefe im Einzelfall fordert, zum Beispiel durch eine minutiöse Analyse einzelner Vorgeschichten bzw. Anamnesen, um die Entwicklung der beschriebenen “Versklavung” des psychischen Systems besser zu verstehen. Dieser mögliche Weg des weiteren Vorgehens stellt eine interessante Alternative (oder besser: Ergänzung) zum Weg dar, den wir bereits eingeschlagen haben, nämlich die Erhöhung der Breite bzw. Vergrößerung der Stichprobe unter Beibehaltung des qualitativen Ansatzes. Die Vergrößerung der Stichprobe soll letztlich eine bessere Generalisierbarkeit unserer Hypothesen und Ergebnisse erlauben, und hier wird die genaue Beschreibung der Stichprobe und möglicherweise sogar die Bildung von Subgruppen, zum Beispiel aufgrund bestimmter und eben möglicherweise tiefer elaborierterer Vorgeschichten, von noch größerer Bedeutung sein.

Literatur

- Belz, M. (2009). *Außergewöhnliche Erfahrungen* (Fortschritte der Psychotherapie, Band 35). Göttingen: Hogrefe.
- Belz, M., & Berger, T. (2008). Psychisches Wohlbefinden, außergewöhnliche Erfahrungen und Emotionsregulation. *Zeitschrift für Anomalistik*, 8, 118-134.
- Belz, M., & Fach, W. (im Druck). Reflections on counseling and therapy for individuals reporting exceptional experiences. In Kramer, W.H., Bauer, E., & Hövelmann, G.H. (Eds.), *Perspectives of Clinical Parapsychology*. Utrecht: Stichting Het Johan Borgman Fond.
- Belz-Merk, M., & Fach, W. (2005). Beratung und Hilfe für Menschen mit außergewöhnlichen Erfahrungen. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 55, 256-265.
- Fach, W., & Atmanspacher, H. (2006). Außergewöhnliche Erfahrungen und Psychodynamik. In Vaitl, D. (Ed.), *Tätigkeitsbericht 2004 – 2005* (S. 54-55). Freiburg i.Br.: Institut für Grenzgebiete der Psychologie und Psychohygiene e.V.

- Spitz, H. (2005). *Emotionsregulation bei außergewöhnlichen Erfahrungen. Eine Fallstudie über Rat-suchende mit außergewöhnlichen Erfahrungen*. Unveröffentl. Diplomarbeit. Freiburg i.Br.: Uni-versität Freiburg.
- Tölle, P. (2003). *Typische Planstrukturen von Menschen mit außergewöhnlichen Erfahrungen*. Un-veröffentlichte Diplomarbeit. Freiburg i.Br.: Universität Freiburg.

**Kommentar zum Artikel von Thilo Hinterberger, Ursula Mochty, Stefan Schmidt, Lisa-Milena Erat und Harald Walach
„EEG-Korrelationen zwischen räumlich weit entfernten Paaren“
In: *Zeitschrift für Anomalistik* 8 (2008), 55-75**

WOLFGANG AMBACH³

EEG-Correlations: Schon wieder neue Effekte!

Der Beitrag von Hinterberger *et al.* beschreibt als Kernpunkt das engagierte und spektakuläre Unterfangen, die EEGs zweier über 750 km weit voneinander getrennter Versuchspersonen eines Paares auf überzufällige Zusammenhänge hin zu untersuchen. Als Leiter einer der einschlägigen Vorstudien (mit weit geringerer räumlicher Distanz der Versuchspersonen) möchte ich den Autoren für die Durchführung der sehr aufwändigen Studien zunächst meinen ehrlichen Respekt aussprechen. In den nachfolgenden Anmerkungen werde ich den Aufsatz nochmals Revue passieren lassen und komme insbesondere nicht umhin, einige aus meiner Sicht problematische Seiten der angewandten Methodik anzusprechen.

Zur ersten der drei Studien – Die „Replikation“

Vorstudien zum Thema der anomalen Kommunikation haben mehrfach konventionell nicht erklärbare Befunde erbracht. Diese gilt es zunächst zu replizieren, bevor es gerechtfertigt

3 Dr. med. Wolfgang Ambach leitet die Forschungsgruppe Klinische und Physiologische Psychologie, Institut für Grenzgebiete der Psychologie und Psychohygiene (IGPP), Freiburg i.Br.

und notwendig erscheint, die Annahmen über das Wesen der Natur in Teilen zu verändern. Dem werden die Autoren genau gerecht, indem sie in ihrer ersten Studie die Replikation zweier konkreter, einschlägiger Vorstudien anstreben.

In den „prinzipiellen Vorüberlegungen“ zur Methodik wird dann aber ein generelles Dilemma zwischen Replikation und Innovation deutlich. Schon allein aus theoretischer Sicht kann es ja nie eine „identische“ Replikation einer Studie geben, selbst wenn man dies noch so sehr anstrebt: „Man steigt nie zweimal in den gleichen Fluss“. Hierüber hinausgehende Veränderungen werden deutlich, als die Autoren ihren Methodikteil damit einleiten, dass „einige Modifikationen der bisher verwendeten Vorgehensweisen“ wesentlich erschienen. Die Frage ist also, nach welcher Veränderung der Methode eine Studie noch sinnvoll als Replikation einer bestimmten Vorläuferstudie (im Sinne einer Bestätigung der vorigen Ergebnisse mit all ihren Implikationen) angesehen werden kann. Hier geht es konkret um die Replikation zweier Studien von Wackermann *et al.* (2003, 2004).

Die Autoren geben für die beabsichtigte Replikation, die sie mit der ersten der drei Studien anstrebten, als methodische Details an: vollständig neuer Hardwareaufbau, anderer Ort, andere Versuchspersonenpaare, „gleiche“ Stimulation (das Schachbrettmuster ist leider nicht genau spezifiziert; die Interstimulus-Intervalle wurden „zufällig gewählt“ ohne nähere Angabe, vermutlich also anders als in den Vorstudien, was eventuell von Bedeutung sein kann), „ähnliches“ Analyseverfahren. Im weiteren werden zusätzlich folgende Veränderungen gegenüber den Vorstudien deutlich: 7-Kanal-EEG, geringfügig andere EEG-Filtereinstellung, 50% der Versuchspersonenpaare in Beziehung stehend. Was die Auswertungsprozedur angeht, wird zwar angegeben, sie ähnele der in früheren Studien beschriebenen, aber beim weiteren Lesen wird klar, dass hier – bezogen auf Wackermann *et al.* (2003, 2004) – fast kein Stein auf dem anderen belassen wurde. Die Details dazu gehen unmittelbar aus dem Aufsatztext hervor; insbesondere die EKP-Auswertung und die Auswertung spektraler Komponenten des EEG wurden komplett neu eingeführt und treten offensichtlich an die Stelle der EEG-Power-Analyse.

Was dieses erste Anliegen der Replikation angeht, mag in diesem speziellen Fall allerdings offenbleiben, ob die Autoren dem Anspruch einer „weitgehend identischen“ Replikation nun Genüge geleistet haben oder nicht. Eine zwischenzeitlich veröffentlichte kritische Auseinandersetzung mit der Analyseverfahren, die in den genannten Bezugsstudien, um deren Replikation es gehen sollte, verwendet wurde, stellt die dort beschriebenen Ergebnisse und die für signifikant befundenen Effekte ernsthaft in Frage (Ambach, 2008), so dass eine Replikation der Studien zwar weiterhin erlaubt, aber immerhin wissenschaftlich entbehrlich erscheint.

Zur Methodik der zweiten und dritten Studie

In den beiden weiteren Studien kamen eine Reihe zusätzlicher erheblicher Modifikationen der Methodik hinzu: große räumliche Distanz, DCF-Synchronisation, in enger Beziehung stehende Versuchspersonenpaare, Einstimmungsprozedur, Duplikation des Versuchs mit vertauschten Rollen, zusätzliche Stimuli (neutrale und affektive IAPS-Bilder; zusätzliche SEB-Bilder in einer Studie); Betrachtungszeit der Bilder in einer Studie variabel und mit zusätzlicher Reaktionsanforderung an die Versuchspersonen (Bildrating per Tastendruck) verbunden; Downsampling des EEG. Hier wird man sicher nicht mehr von einer Replikation konkreter früherer Studien sprechen; ein solcher Anspruch wird auch nicht formuliert. Sämtliche der genannten Modifikationen haben natürlich gute Gründe. Manche der methodischen Veränderungen erscheinen unter pragmatischen Gesichtspunkten ausgesprochen plausibel, vorteilhaft, vielversprechend (beispielsweise die elegante DCF-Synchronisation); manche anderen verändern oder erweitern den Suchfokus (so etwa die Analyse der spektralen Anteile des EEG); weitere entspringen einem etwas anders gelagerten Theoriegebäude und haben das Ziel, diesem gerecht zu werden (etwa der Einbezug emotionaler Stimuli).

Der Umstand aber, dass mit diesen Veränderungen letztlich ein völlig neues Studiendesign und ein neuer Analyseplan entstanden sind, hat auch eine nachteilige Implikation: Die Studien sind dadurch kaum geeignet, direkt mit Vorläuferstudien vergleichbare Ergebnisse hervorzubringen und die Hypothesen, die sich aus den vorangegangenen Studien ergeben haben, zu überprüfen. Vielmehr muss man den jetzt gewonnenen Resultaten wiederum explorativen Charakter zuschreiben. In der Tat, wie in unzähligen Grenzgebietsstudien zuvor, werden als Ergebnis neue Effekte (Unterschiede in der Alpha-Band-Aktivierung der nicht-stimulierten Person) beschrieben, auf die hin frühere Studien gar nicht ausgewertet worden waren.

Zur großen räumlichen Distanz: Wieviele Kilometer sind genug?

Unerklärliche Zusammenhänge über 750 Kilometer hinweg zu untersuchen, kann mancher für absurd halten, andere vielleicht für gewagt oder beeindruckend, für erstaunlich oder unnötig; in jedem Fall ist das Vorhaben nicht „gewöhnlich“, und gerade das verleiht den beiden „Long-Distance“-Studien ihre Einzigartigkeit und ihren Abenteuercharakter. Die räumliche Distanz ist wissenschaftlich damit begründet, dass man konventionelle, also bekannte, Mechanismen der Informationsübertragung zwischen beiden untersuchten Personen unbedingt ausschließen will. Dieses Anliegen haben alle „Anomalous-Communi-

cation“-Studien seit jeher gemeinsam, und vielen der früheren Projekte sind Schwächen gerade in diesem Punkt der physikalischen Abschirmung kritisch entgegengehalten worden. Kein Wunder also, dass die Experimentatoren sich dies nicht nachsagen lassen wollen und durch eine imposante Distanz vorbeugen.

Ein Problem dabei ist: Wieviel Distanz ist nötig, und wieviel ist ausreichend, um „konventionelle“ Informationsübertragung auszuschließen? Wenn es um die fünf Sinne, insbesondere um akustische oder direkte optische Übertragung geht, scheint dies noch halbwegs gut beantwortbar zu sein. Dazu wären übrigens nicht unbedingt Kilometer-Distanzen erforderlich; 750 km Abstand wären hiernach nicht besser als die Trennung der Versuchspersonen durch mehrere Mauern oder vielleicht durch zwei Straßenzüge. Schwieriger wird es, wenn man dem Menschen weitere, unbewiesene oder unbekante, aber im Bereich der konventionellen Physik liegende Kommunikationswege zuschreiben würde. Solche werden aber im Text nicht konkret benannt, auch nicht in Mutmaßungen. Welchen Abstand bräuchte man beispielsweise, wenn beide Partner etwa „magnetisch“ oder „wie über Ultrakurzwelle“ kommunizieren könnten, oder „wie Elektronenstrahlen“ oder gar „auf ähnlichem Weg, wie das DCF-Signal der Funkuhren“? Kurzum, es wäre wünschenswert, dass die Art der potentiellen, konventionellen Kommunikationswege genauer benannt wird, gegen die in diesen beiden Studien besser als in früheren abgeschirmt wurde. Sonst wäre die Verwendung einer so großen Distanz einfach nur unter den wissenschaftlichen Kuriositäten zu sammeln.

Zum Randomisationstest

Die Art der Stichprobenziehung, die zu erheblicher Ergebnisverzerrung führen kann (vgl. Ambach, 2008) ist leider nicht genau beschrieben. Es ist unter anderem entscheidend, aus welchem Pool die Stichproben gezogen wurden und ob mit oder ohne Zurücklegen gezogen wurde (z.B. Edgington, 1987).

Zum multiplen Testen und zur Datenreduktion

Die verwendete „Mittelung über alle EEG-Kanäle“ ist ein wenig übliches und vermutlich auch wenig sensitives Verfahren, freilich aber nicht falsch. Die Autoren haben es gewählt, um dem Problem der wechselseitigen Abhängigkeit der EEG-Kanäle voneinander zu begegnen. Als sensitiveres Verfahren rege ich an, die Kanäle zunächst getrennt voneinander auszuwerten und in einem zweiten Schritt durch eine aufwändigere Prozedur das genaue Ausmaß dieser wechselseitigen Abhängigkeit abzuschätzen und diesem durch ent-

sprechende Korrektur Rechnung zu tragen. Dies ist etwa durch eine Monte-Carlo-Simulation mit multipler Rekombination der kanalweisen Ergebnisse möglich (beschrieben in Ambach, 2008). (Ferner erhielt ich mündlich die Anregung, das Problem der Abhängigkeit der Kanäle durch eine Faktorenanalyse mit wesentlich geringerem Rechenaufwand zu lösen; hierzu kann ich allerdings nichts berichten.)

Die Autoren führen an mehreren Stellen an, dass die wegen des multiplen Testens erforderliche und genau angemessene Korrektur schwer zu bestimmen sei. Dies bezieht sich nicht nur auf die EEG-Kanäle, sondern auch auf weitere unternommene Tests, insbesondere Tests auf Basis höher aggregierter Daten, deren Ergebnisse nicht unabhängig voneinander sind. Hier wird es wirklich kompliziert, insbesondere auch rechenaufwändig – aber es ist möglich, auch die statistische Abhängigkeit von bereits höher aggregierten Ergebnissen (z.B. Ergebnisse der Probandenpaare unter zwei variierten Versuchsbedingungen) durch eine Monte-Carlo-Simulation abzuschätzen. (Nebenbei bemerkt, lassen sich mit dieser Methode auch potentielle Einflüsse der zeitlichen Abfolge von Stimuli verschiedener Klassen in ihrer Bedeutung abschätzen. Dies hat sich in unveröffentlichten Analysen allerdings als marginal herausgestellt.)

Zum beschriebenen Alpha-Band-Effekt: Sind drei Studien besser als eine?

Wie oben begründet, schreibe ich den referierten Studien eher hypothesengenerierenden als hypothesenbestätigenden Charakter zu. Dies schließt auch den hier erstmals beschriebenen Effekt der Unterschiede in der Alpha-Band-Aktivität mit ein. Die Autoren führen an, sie hätten diesen Effekt in allen drei Studien gefunden. Dies macht die Beobachtung sicher bedeutender als wenn er nur ein oder zwei Mal aufgetreten wäre; für eine Hypothesenbestätigung würde ich allerdings dennoch erwarten, dass eine neue Studie, die zeitlich nach der Formulierung dieser Hypothese initiiert wurde, diesen Effekt repliziert.

Der vielzitierte texanische Scharfschütze zieht seinen Colt und schießt auf das große Scheunentor. Er sucht dann das Einschussloch, malt mit Kreide eine Zielscheibe darum herum, und ruft „Volltreffer!“. Der Wissenschaftler grenzt sich hiervon ab, indem er das Postulat einhält, nicht dieselben Daten zum Generieren und zum Bestätigen einer Hypothese heranzuziehen. Dies impliziert eine Abfolge „Studie A → Hypothese formulieren → Studie B zur Prüfung genau dieser Hypothese“. Wenn, wie hier, die Abfolge „Studie A → Studie B → Studie C → Hypothese formulieren und bestätigen“ lautet, genügt dies dem genannten Postulat nicht; der Effekt ist ja gerade durch sein mehrfaches Auftreten augenfällig, selektiert und dann hypothesengenerierend geworden.

Schlussbemerkung: „Mentales multiples Testen“

Die letzten Sätze leiten zu einer kritischen abschließenden Betrachtung über, die über die Auseinandersetzung mit dem Aufsatz von Hinterberger et al. hinausgeht: Warum werden in diesem Forschungsbereich, der sich mit der Existenz von Unerklärlichem beschäftigt, immer wieder neue Effekte gefunden, auch für signifikant befunden, sind dann aber in Folgestudien nicht replizierbar?

Freilich ist, vor allem im Falle geringer Effektstärken, nicht in jedem Fall eine gelingende Replikation zu erwarten; die Existenz eines bestimmten unerklärten Phänomens ist im übrigen ohnehin auch durch eine mangelnde Reproduzierbarkeit nicht widerlegbar. Darüber hinaus gibt es längst Theorien, die behaupten, dass gerade das fluktuierende, nicht stabil reproduzierbare Auftreten unerklärlicher Dinge und Zusammenhänge deren wahrhaftige Existenz bestätigt. Die Frage, welche Bedeutung solchen Theorien zukommt, soll hier außen vor bleiben.

Mein Anliegen ist es vielmehr, zunächst den möglichen Einfluss von einfacher und konventionell verstehbaren ergebnisverzerrenden Mustern, die sich im Verlauf des experimentellen Arbeitens ergeben, so weit wie möglich auszuloten. Solche Muster, die einen systematischen Einfluss auf die berichteten Ergebnisse haben, zu identifizieren erscheint mir wertvoller als eine frühzeitige Erweiterung der theoretischen Modelle. Dies folgt der wissenschaftlichen Devise, den bestehenden theoretischen Rahmen so lange beizubehalten, bis er aufgrund einer Widerlegung der Anpassung bedarf.⁴

Ein gutes Quantum Beachtung verdient in diesem Zusammenhang die Abfolge des Generierens und Prüfens von Hypothesen, speziell unter dem Aspekt des mehrfachen Testens.

Wie die Autoren sehr dezidiert feststellen, ist der Umstand des multiplen Testens von Hypothesen bei der Signifikanzbewertung der Ergebnisse zu berücksichtigen. Das heißt, das Signifikanzniveau ist danach zu korrigieren, (1) wie viele Hypothesen getestet wurden und (2) in welchem Maße die Einzelergebnisse statistisch voneinander abhängig sind. Was die angemessene Korrektur darstellt, ist oft nur mit Schwierigkeiten, mit erheblichem Aufwand und mit einiger Ungenauigkeit zu ermitteln. Der offene Umgang mit diesem Problem im Aufsatz von Hinterberger et al. entspricht guter wissenschaftlicher Praxis und lässt auf ein gutes Maß der in diesem Bereich besonders wertvollen Selbstkritik schließen.

4 In diesem Sinne halte ich auch die angestellten Überlegungen zu quantentheoretischen Erklärungen des neu gefundenen Zusammenhangs zumindest zum jetzigen Zeitpunkt für entbehrlich.

Um Missverständnissen gleich vorzubeugen: Ich will im Folgenden keineswegs den Autoren des Aufsatzes irgendein unwissenschaftliches Vorgehen anlasten; im Gegenteil bin ich von einer sehr sorgfältigen und korrekten Versuchsdurchführung überzeugt. Vielmehr möchte ich hier ein Phänomen aufzeigen, das bei der Auswertung experimenteller Untersuchungen fast immer auftritt, aber vermutlich zu wenig Beachtung findet.

Bei den vielfältigen Anstrengungen zur Korrektur multiplen Testens wird immer davon ausgegangen, dass man genau weiß, wie viele und welche Hypothesen getestet wurden; in der Regel sind es die, die der Autor im Text zu beschreiben gedenkt. Ich füge dem meine kritische Überzeugung hinzu, dass es darüber hinaus, wenn die zu prüfenden Hypothesen nicht vor der Studiendurchführung beim Notar liegen, ein „mentales multiples Testen“ gibt, das sich vorwiegend im Geist des Experimentators abspielt. Wie viele und welche weiteren Hypothesen, von denen wir nie erfahren, wären denkbar gewesen? Wie viele denkbare Effekte („Effekte“ immer verstanden im statistischen Sinn) hatten vielleicht die prinzipielle Chance, auf ihre Signifikanz hin getestet zu werden, wenn sie augenfälliger gewesen wären? Wie viele und welche davon wurden vielleicht vom Experimentator kurz, intuitiv und ohne weitere Dokumentation im Hinblick auf die Aussicht einer möglichen Bestätigung bei näherer Analyse betrachtet und dann schnell wieder vergessen? Muster und Effekte in den Daten, die intuitiv auffallen, werden mit einer höheren Wahrscheinlichkeit später zu einer zu testenden Hypothese erhoben als solche, die prima vista schon zufällig wirken. Mental getestet, wenn auch zum Teil nur flüchtig, wurde aber eine größere Anzahl als später (erinnert und) benannt wird. Es wird deutlich, dass hier eine Grenzziehung zwischen den getesteten Hypothesen und den nur potenziell denkbaren, aber nicht getesteten Hypothesen gar nicht mehr möglich ist; an eine angemessene nachträgliche Korrektur des „mentalenen multiplen Testens“ ist gar nicht erst zu denken.

Meine Hypothese ist nun, dass das Phänomen des „mentalenen multiplen Testens“ zu wenig Beachtung findet und auch von vielen Experimentatoren vernachlässigt wird. In der Folge wird die Zahl der (zum Teil implizit und flüchtig) getesteten Hypothesen unterschätzt. Es werden dann auch unter Zufallsbedingungen mehr Effekte gefunden als es dem explizit formulierten Signifikanzniveau entspricht. Von den dann für signifikant befundenen Effekten ist übrigens zu erwarten, dass sie (1) nicht stabil reproduzierbar sind und (2) von wechselnder Natur (betroffener Messparameter, Lokalisation, Richtung, zeitliche Zusammenhänge usw.) sind.

Während die sich aufdrängende Konsequenz, die aus dieser Überlegung zum „mentalenen multiplen Testen“ letztlich zu ziehen wäre (nämlich die Hinterlegung von Hypothesen vor Studienbeginn), in keiner Weise neu ist, könnten die hier formulierten Gedanken vielleicht

in der wissenschaftlichen Praxis dazu anregen, auch den statistisch signifikanten Ergebnissen noch etwas länger als bisher einen vorläufigen Charakter zuzuschreiben. Dies scheint gerade in der Grenzgebietenforschung angemessen angesichts der Tragweite, die ein stabil replizierbares Ergebnis hier hätte.

Literatur

- Ambach, W. (2008). Correlations between the EEGs of two spatially separated subjects – a replication study. *European Journal of Parapsychology*, 23, 131-146.
- Edgington, E.S. (1987). *Randomization Tests*. New York: Marcel Dekker.
- Wackermann, J., Naranjo, J.R., & Pütz, P (2004). Event-related correlations between brain electrical activities of separated human subjects. *The Parapsychological Association 47th Annual Convention. Proceedings of Presented Papers* (S. 465-468). New York, August 5-8, 2004.
- Wackermann, J., Seiter, C., Keibel, H., & Walach, H. (2003). Correlations between brain electrical activities of separated human subjects. *Neuroscience Letters*, 336, 60-64.