

## Fortgesetzte Diskussionen zu früheren Beiträgen

**Kommentar zur Replik von Rudolf Henke**

**„Banale und spektakuläre UFO-Sichtungsfälle nicht vermengen“**

In: *Zeitschrift für Anomalistik* 4 (2004), 251-255

HUGH DEASY<sup>1</sup>

### **UFO-Fälle mit hoher Strangeness werden auch von Profis beobachtet**

Es war gut, im Jahresband 2004 der *Zeitschrift für Anomalistik* auch nicht-„skeptische“ Kommentare über UFOs und Kornkreise zu lesen, denn nach dem abschätzigen Tenor von Ulf Harendarskis Eröffnungsworten zu diesem Themenbereich in der allerersten Ausgabe der Zeitschrift (Harendarski 2001) war zu befürchten, dass nur ein ähnlicher derartiger postmoderner „Skeptizismus“ die Seiten der jungen Publikation zieren würde. Daher war es ermutigend, in Band 4 neben meinem eigenen Beitrag auch Kommentare von Andreas Müller, Jan Schwochow und Wolfgang Schindler zu lesen, die nicht der Linie der GWUP und ihrer Anhänger folgen. Doch gleichzeitig wurde der Geist von Harendarskis Artikel wieder ins Leben gerufen durch die Replik von Rudolf Henke (2004) auf meinen Kommentar (Deasy 2004) zu der Umfrage über UFO-Sichtungserfahrungen (Wunder & Henke 2003). Henke greift nur allzu gerne die Idee auf, dass alle UFO-Berichte einfach nur eine Form der Literatur seien, obwohl Harendarski diese Überlegung tatsächlich nur auf Entführungsberichte bezogen haben wollte. Daher ist es an dieser Stelle sinnvoll, ein für alle Mal festzustellen, dass der Wert dieser Berichte mit völliger Sicherheit nicht nur im Reich der Fiktion anzusiedeln ist.

Die Idee, UFO-Berichte als Fiktion zu abzutun, erinnert an das Vorgehen jener Psychologen, Neurowissenschaftler und Philosophen, die in dem Versuch, jede Vorstellung eines *deus ex machina* gründlich auszurotten, auch dem Konzept des subjektiven Bewusstseins jegliche Realität absprechen. Eines ihrer Argumente ist die Idee, dass der „Konfabulator“ in der linken Gehirnhälfte nach irgendwelchen kognitiven Fehlern stets nachträglich die Realität neu konstruiert, sodass sie mit der eigenen Erwartung einer bequemen, aber leider fehlgeleiteten Realität übereinstimmt. Dieselben Psychologen schrecken auch vor dem Homunculus mit seinen Implikationen für die Seele oder den unendlichen Regress zurück. Doch interessanterweise

---

<sup>1</sup> Dr. Hugh Deasy ist Physiker und im Aufgabenfeld der Satellitenflugdynamik bei der ESOC in Darmstadt tätig. E-Mail: hugh.deasy@esa.int.

gibt es auf das Homunculus-Argument keine Antwort – siehe etwa Sutherlands (1997) Widerlegung von Metzingers „Homophobia“ –, was impliziert, dass wir uns entweder in einer dualistischen oder einer idealistischen Realität wiederfinden. Die Überkompensation dieser Psychologen, die sich auf „Konfabulatoren“ berufen müssen, während sie die unvermeidlichen Homunculus-Konsequenzen des Bindungsproblems<sup>2</sup> verdrängen, erinnert an die kognitiven Defizite, die Henke zur Erklärung von UFO-Berichten heranzieht.

Nach diesen Bemerkungen zu den philosophischen Implikationen möchte ich nun auf die konkreteren Behauptungen Henkes eingehen. Erstens sei es seltsam, dass unter den Meldern von UFO-Sichtungen mit hohem Strangeness-Grad praktisch keine Astronomen sind. Dabei wird impliziert, dass nur Menschen ohne astronomische Kenntnisse diese Sichtungen mit hoher Strangeness berichten. Allerdings ist das Fehlen von Astronomen nicht besonders überraschend, denn der Anteil von Astronomen in der Bevölkerung beträgt nur etwa 1 zu 40000, sodass es extrem unwahrscheinlich wäre, wenn ein Astronom an einer Sichtung mit hoher Strangeness beteiligt wäre. Nach meiner eigenen Erfahrung als Berufsastronom in Irland gab es in ganz Irland niemals mehr als etwa 100 Astronomen bei einer Gesamtbevölkerung von 4 Millionen Einwohnern. Interessanterweise begann sich noch ein weiterer dieser irischen Astronomen für UFOs zu interessieren: Eamonn Ansbro, der jetzt das Kingsland Observatory leitet und dort nach extraterrestrischen Sonden sucht, um die Astronautentheorie von Dutton zu testen (Ansbro & Overhauser 2001). Die Astronomen-Welt ist klein: Ansbro arbeitet nun mit der italienischen EMBLA/Hessdalen-Gruppe zusammen, die von Teodorani geleitet wird. Man beachte auch, dass während der EMBLA-Missionen Teodorani und einige weitere italienische Astronomen Zeugen der Hessdalen-Lichter wurden – und mindestens eine dieser Beobachtungen hatte einen hohen Strangeness-Grad. Allein dies genügt schon, die Behauptung Henkes zu widerlegen, dass kein Astronom an Sichtungen mit hohem Strangeness-Grades beteiligt war.

Ein weiterer Beleg dafür, wie klein und eng verbunden die internationale astronomische Gemeinschaft ist: Auf einem Treffen der Irish Astronomical Science Group traf ich Bernhard Haisch, dessen Webseite <http://www.ufoskeptic.org> eine gute Einführung für ernsthaft interessierte Astronomen oder andere Wissenschaftler in den Bereich der UFO-Forschung darstellt. Ein weiterer Zufall: UFO-Papst Illobrand von Ludwiger war an der Bamberger Sternwarte an einem Projekt zur Untersuchung von Veränderlichen Sternen beteiligt – und ich lernte mein erstes Deutsch durch das Lesen von Artikeln über Beobachtungen zu diesen Sternen an dieser Sternwarte.

Diese hohe Quote von mir bekannten Astronomen, die an UFO-Studien beteiligt sind, zeigt eher das Gegenteil von Henkes Bild eines Mangels an bedeutungsvollen Verbindungen zwischen UFO-Forschern und Astronomen. Weitere Astronomen, die die UFO-Frage stu-

---

2 Unter dem Bindungsproblem versteht man die Frage, wie das Gehirn eine Vielzahl einzelner Sinnesindrücke zu einer einheitlichen Wahrnehmung integriert.

dierten und zu dem Schluss kamen, dass an den Berichten etwas dran sein könnte, waren u.a. Peter Sturrock und Allen Hynek. Daher sind es gewöhnlich die astronomisch oder akademisch schwachen „Skeptiker“ und nicht die echten Astronomen, die die Wichtigkeit der Beschäftigung mit dem Kern des UFO-Phänomens nicht erkennen.

Zudem interessiert sich eine überproportionale Zahl von Menschen aus dem Luft- und Raumfahrtbereich für UFOs, angefangen von Militärfluglotsen und Piloten bis hin zu Astronauten wie Gordy Cooper (der kürzlich verstorben ist) und Ed Mitchell. Während ein Pseudo-Skeptiker wohl darauf hinweisen würde, dass Menschen in diesen Berufen auch am ehesten angesprochen werden, wenn es darum geht, solche Beobachtungen zu begutachten (oder zu entlarven), ist es doch offensichtlich, dass diese Berufsgruppe mehr als andere über Himmelsphänomene aller Art Bescheid weiß und deshalb jedes UFO, das sie nicht erklären können, viel wahrscheinlicher unser gegenwärtiges Wissen über physikalische Grenzen sowie über Antriebssysteme in Frage stellt.

Nun zu der Sichtung des Astronomen und Pluto-Entdeckers Clyde Tombaugh. Nach Swords (1999) berichtete Tombaugh:

I saw the object about eleven o'clock one night in August, 1949 from the backyard of my home in Las Graces, New Mexico. I happened to be looking at zenith, admiring the beautiful transparent sky of stars, when suddenly I spied a geometrical group of faint bluish-green rectangles of light similar to the "Lubbock lights". My wife and her mother were sitting in the yard with me and they saw them also. The group moved south-southeasterly, the individual rectangles became foreshortened, their space of formation smaller, (at first about one degree across) and the intensity duller, fading from view at about 35 degrees above the horizon. Total time of visibility was about three seconds. I was too flabbergasted to count the number of rectangles of light, or to note some other features I wondered about later. There was no sound. I have done thousands of hours of night sky watching, but never saw a sight so strange as this. The rectangles of light were of low luminosity; had there been a full moon in the sky, I am sure they would not have been visible.

Henke meint immer noch, dass die Sichtung von Tombaugh und seiner Familie wohl ein meteorischer Feuerball gewesen sein muss, der die meisten Astronomen überrascht, weil sie im Durchschnitt noch nie zuvor einen gesehen haben. Doch angesichts der obigen Beschreibung ist das offensichtlich Unsinn – kein Meteor besteht aus regelmäßig angeordneten Rechtecken. Aber diese Art von Theorie ist typisch für „Skeptiker“. Sie stellen eine wilde Behauptung in den Raum und versuchen durch den Überraschungseffekt alle möglichen Einwände zu ersticken. Doch die Erklärung, die hier geboten wird, ist kaum besser als die Vermutung, dass mehrere Polizeiwägen in einer wilden Verfolgungsjagd über Staatsgrenzen hinweg die Venus verfolgten. Dies passt einfach nicht zu den Fakten.

Henkes Aussagen über Schwierigkeiten der Wahrnehmung sind ebenso ungültig – Clyde Tombaugh war einer der besten, oder sogar der beste, beobachtende Astronom der USA im 20. Jahrhundert. Seine Beschreibung hier ist ein Beispiel für seine sorgfältige Beobachtungsgabe. Einfach nur zu behaupten, dass Tombaugh unter Wahrnehmungstäuschungen litt,

macht dies noch lange nicht wahr. Ganz im Gegenteil. Ich kannte einen Beobachter persönlich, der dieselben Qualitäten wie Tombaugh hatte – solche Beobachter haben die Geduld Hiobs und die ruhige, sicherere Genauigkeit der Beobachtung, die wir normalen Sterblichen gar nicht nachvollziehen können. Mein Kollege konnte die schwächsten Sterne sehen und beeindruckte mich einmal dadurch, dass er mir Kondore über den Anden zeigte, obwohl sie nur unglaublich winzige Punkte darstellten.

Wiederum erinnert die Taktik von „Skeptikern“ wie Henke an diejenige von bestimmten Neurowissenschaftlern oder Psychologen, die anscheinend immer nur an kognitiven Defiziten, Behinderungen und Pathologien interessiert sind. Sie scheinen nicht daran interessiert zu sein, wie ein perfekt funktionierendes Gehirn das Wunder der perzeptuellen Bindung erreicht. In ähnlicher Weise wollen diese „Skeptiker“ auch nichts von der Großartigkeit der wahrhaft guten Beobachter wissen.

Als weiteren Beleg dafür, dass Henkes Spiel des verschwindenden UFO-Anteils und andere Statistiken irreführend sind, möchte ich folgendes anführen: Zunächst einmal sind UFOs nicht wie Zufallszahlengeneratoren (RNGs) – bei jenen ist es nötig, die Statistik von Millionen von Durchgängen zu untersuchen. Doch im Falle der UFOs ist eine einzige Ausnahme genug, um die Nullhypothese zu widerlegen. Daher möchte ich einige neue, recht gut dokumentierte Fälle mit mehreren Zeugen zitieren, um die negative Darstellung Henkes auszugleichen.

Erstens hat sich mit dem Ende des kalten Krieges auch die Haltung südamerikanischer Regierungen zur Offenheit bei UFO-Berichten gewandelt. So berichtet etwa Corrales (o.J.):

Evidence of a thaw in the icy wall of silence surrounding the Chilean military's position on the subject of UFOs became evident in the spring of 1997, when air traffic controllers at the Chacalluta International Airport, on the outskirts of the city of Arica, reported seeing three UFOs hovering over the Pacific Ocean within plain view from the tower. Uncharacteristically, the unknown vehicles allowed the onlookers to take a good, long look -- they remained in position for over two hours, finally speeding away toward the Andes.

A news story appearing in Arica's La Tercera on April 1, 1997 featured an interview with the airport's director, in which he explained that the UFOs had not been visible to his own radar system, and that the assistance of other airports in Chile and nearby Perú had been requested. The latter were also unable to pick the objects up on their systems.

But the shocker came on April 2, 1997, when Chile's General Directorate of Civil Aeronautics (DGAC, in Spanish), confirmed that the trio of saucers seen over the water off Arica had been confirmed by the Chilean Air Force, which had recorded the objects moving at speeds in excess of 12,000 kilometres an hour. A spokesman for the directorate went on to add that the UFO phenomenon was decidedly real and was neither "meteoric nor climactic" -- a bold admission we have yet to hear other countries make.

Auch <http://ufologie.net/htm/offichili.htm> berichtet in ähnlicher Form von diesem Fall.

Zweitens: Die Haltung Frankreichs zu UFOs war niemals ganz so abgeschlossen wie die anderer Nationen. GEPAN/SEPRA, die Abteilung der französischen Raumfahrtagentur CNES, deren Aufgabe die Untersuchung von „OVNIs“ ist, konnte, da sie offiziellen Zugang zu den französischen Polizei- und Militärdokumenten hatte, bemerkenswerte Fortschritte in einigen Fällen mit relativ hohem Strangeness-Grad erzielen. So berichtet die *Times* (Sage 2003):

Consider, for instance, a case reported in 1994, when the crew of an Air France flight from Nice to London saw a dark, 300m (1,000ft) long object over the Paris region. The object disappeared before the aircraft had got near it, and the flight continued without difficulty. A few days later Velasco travelled from his office in Toulouse to the military aviation control centre outside Paris, where he was given a read-out of the radar information from the day in question. It revealed that an unknown object had indeed flown over the French capital.

Hierzu auch Sturrock et al. (1998) im Abschnitt „Radar Evidence“:

One of these cases is particularly interesting. This case occurred on January 28, 1994, about 70 kilometers southeast of Paris, at a height of 11,700 meters, under excellent meteorological conditions. An object was first noticed by a steward who happened to be in the cockpit, and his observation was then confirmed by the copilot. The captain then saw the object. It was above the thick layer of altocumulus clouds at 10,500 meters. The captain described the object as resembling a gigantic disk (diameter about 1000 meters, thickness about 100 meters) with slightly fuzzy edges. The witnesses suddenly lost sight of the object when the edges appeared to go out of focus and the object disappeared.

Corresponding radar information was obtained from the military air traffic control (ATC). The object was positively detected by radar for a period of 50 seconds. The apparent speed of the object was measured first as 110 knots, then as 84 knots, and subsequently as zero. The altitude of the object was not recorded by radar. The radar was also tracking a nearby commercial aircraft and appeared to be in good working order. There appears to be good correspondence between the radar measurements and the visual observations.

Auch die britische Flugbehörde hat Informationen über UFO-Sichtungen von Piloten veröffentlicht. Beispielsweise hat ein Bericht aus dem Jahr 1982 (CAA o.J.) einen relativ hohen Strangeness-Grad.

Und schließlich haben die Sichtungen nicht aufgehört. Im Jahr 2003 filmte ein BBC-Kameramann UFOs, die wahrscheinlich Beispiele für die „Hanbury Orbs“ waren (BBC 2003).

Daher kann ich die Punkte, die ich in meinem ursprünglichen Kommentar angeführt habe, nochmals bestätigen. Die „Suppe“ der UFO-Studien ist keineswegs die dünne Brühe, die Henke servieren möchte, sondern ein reicher und nahrhafter Eintopf, der großzügig mit Fällen von hohem Strangeness-Grad gewürzt ist.

### Literatur

- Ansbro, E.; Overhauser, C. (2001): SETV: Opportunity for European initiative in the search for extraterrestrial intelligence. Conference Proceedings on "First European Workshop on Exo/Astrobiology", 21-23 May 2001, ESRIIN, Frascati (Rome), Italy, ESA SP-496, 285-288.
- BBC (2003): "UFO" lights caught on film. BBC, 23. Juli 2003, [http://news.bbc.co.uk/2/hi/uk\\_news/england/hereford/worcs/3090649.stm](http://news.bbc.co.uk/2/hi/uk_news/england/hereford/worcs/3090649.stm).
- CAA (o.J.): Fallbericht der Civil Aviation Authority (CAA), [http://www.narcap.org/international/orgs/case\\_files/UK\\_CAA78\\_84.DOC](http://www.narcap.org/international/orgs/case_files/UK_CAA78_84.DOC).
- Corrales, Scott (o.J.): The South American Military and the UFO Phenomenon. <http://www.ufoevidence.org/documents/doc1749.htm>.
- Deasy, H. (2004): UFO-Sichter sind ganz normale Leute. *Zeitschrift für Anomalistik* 4, 248-250.
- Harendarski, U. (2001): Mord und Entführung: was man alles tun kann, um Literatur zu erkennen. *Zeitschrift für Anomalistik* 1, 6-19.
- Henke, R. (2004): Banale und spektakuläre UFO-Sichtungsfälle nicht vermengen. *Zeitschrift für Anomalistik* 4, 251-255.
- Sage, A. (2003): Salut, Earthlings. *Times of London*, February 5, 2003.
- Sturrock, P. et al. (1998): Physical evidence related to UFO reports. *Journal of Scientific Exploration* 12, 179-229.
- Sutherland, K. (1997): Homophobia in the Cartesian theatre. <http://www.imprint.co.uk/online/homophob.html>.
- Swords, M.D. (1999): Clyde Tombaugh, Mars, and UFOs. *Journal of Scientific Exploration* 13, 685-694.
- Teodorani, M. (2001): Instrumented Search for Exogenous Robotic Probes on Earth. Conference Proceedings on "First European Workshop on Exo/Astrobiology", 21-23 May 2001, ESRIIN, Frascati (Rome), Italy, ESA SP-496, 379-381.
- Teodorani, M. (2004): A Long-Term Scientific Survey of the Hessdalen Phenomenon. *Journal of Scientific Exploration* 18, 217-251.
- Wunder, E.; Henke, R. (2003): Menschen mit UFO-Sichtungserfahrungen. Eine Umfrage unter 447 Besuchern von Volkshochschul-Vorträgen zum UFO-Phänomen. *Zeitschrift für Anomalistik* 3, 112-133.

**Kommentare zum Aufsatz von Andreas Sommer**  
**„Im Niemandsland der Bewusstseinsforschung: Survival Research“**  
**In: *Zeitschrift für Anomalistik* 5 (2005), 178-237**

PETER MULACZ<sup>3</sup>

### **Hundertfünfzig Jahre Stagnation**

Der Aufsatz gibt einen sehr umfassenden Überblick über die „Survival“-Forschung, wobei sowohl die ältere wie auch neuere internationale Literatur Berücksichtigung findet. Dennoch vermisst man so manches Grundsätzliche bzw. lässt sich so manches auch ganz anders betrachten.

Von einem „bottom-up approach“ ausgehend, ist die grundlegende Frage, „wer“ oder „was“ möglicherweise „überlebt“ – letztlich die Frage nach dem Leib-Seele-Problem, die philosophische Anthropologie. Noch habe ich die Stimme des „Altmeisters“ Hans Bender gleichsam im Ohr, der nicht müde geworden war, immer wieder zu betonen, dass wir empirisch das Psychische nur in physischer Bindung kennen. Dabei ist die empirische Forschung nur die eine Seite (welcher ich auch gerne den Primat zuschreibe), während die andere Seite die philosophische Durchdringung des Problems ist (Beloff 1989; Beloff 1990). Hinzuweisen wäre in diesem Kontext auch auf die klassische „Psi“-Theorie von Thouless und Wiesner, welche in einem großen Wurf eine Vereinheitlichung des normalen und des paranormalen Seelenlebens auf einer dualistischen Basis anstrebt. Solange diese Grundfragen nicht geklärt sind, bleibt jede Diskussion der „Survival“-Problematik notwendigerweise konditional („falls – dann“) (Mulacz 1976). Diese grundlegende Frage wird freilich ganz allgemein in der Literatur recht stiefmütterlich behandelt, auch in den jüngst erschienenen Werken. So geht David Fontana (2005) in seinem opus magnum von fast 500 Seiten nur recht beiläufig darauf ein, und Michael Grosso (2006) scheint diese Problematik überhaupt nicht wahrzunehmen.

Der „top-down approach“ – der mehr als hundert Jahre währende Disput zwischen Animisten und Spiritisten (oder, modisch gesagt, zwischen „Super-psi“, also einer Extrapolation von psi, und „Survival“) stellt sich somit ähnlich einem Koloss auf tönernen Füßen dar. Dennoch ist dieser Disput wichtig und man muss dem Autor dankbar sein, dass er ihn kenntnisreich und ausführlich abhandelt.

Es muss angemerkt werden, dass es konkret nicht um „Survival“ in einem irgendwie gedachten Jenseits geht, sondern um die (angebliche) Kommunikation zwischen „Jenseitigen“

---

3 Prof. Peter Mulacz ist Vizepräsident der Österreichischen Gesellschaft für Parapsychologie und Grenzbereiche der Wissenschaften. E-Mail: peter.mulacz@parapsychologie.info.

und Diesseitigen. Man mag sich über diese Unterscheidung wundern und sie als müßig empfinden, denn nur durch eine solche gedachte Kommunikation, wenn überhaupt, erhalten wir Kenntnis von „Survival“ – ansonsten wären wir in der sprichwörtlichen Situation, wo der Baum im Wald fällt und niemand hat es wahrgenommen. Dennoch scheint mir diese Unterscheidung sinnvoll, weil sie Raum gibt für die traditionellen Jenseitsvorstellungen der Offenbarungsreligionen, welche freilich in unserer westlichen Gesellschaft eine immer geringere Rolle spielen. Zwei der „vier letzten Dinge“ der traditionellen katholischen Eschatologie sind bekanntlich „Himmel“ und „Hölle“, beides Zustände, die für die Gläubigen wohl eine Art von „Survival“ darstellen (genau genommen noch mehr, nämlich „immortality“), dennoch aber keine Interaktion mit den Lebenden erlauben. Eine solche gäbe es allenfalls aus dem Zwischenzustand des Purgatoriums heraus, was zur Diskussion der spezifisch katholischen Version spiritistischer Spontanphänomene führen könnte, zu den Fegefeuer- und Arme Seelen-Erscheinungen, die der Autor allerdings nicht erwähnt.

Auf das psychophysische Problem zurückkommend, muss anerkannt werden, dass Andreas Sommer die Wichtigkeit der philosophischen Anthropologie keineswegs übersieht. Er weist auf „tiefgehende subjektive Erfahrungen leibfreien ich-bewussten Daseins“ hin – AKE<sup>4</sup> und Todesnäheerfahrungen. Aber er verbucht diese Indizien für einen substantiellen Seelenbegriff auf der Seite von „Survival“, also eine argumentative Bei-Ordnung zu den beiden miteinander in Konflikt stehenden Hypothesen, während es sich in Wirklichkeit um eine andere hierarchische Ebene handelt: nur auf der Basis der Annahme eines substantiellen Seelenbegriffs ist die ganze Diskussion überhaupt möglich. Dabei ergibt sich aus einem substantiellen Seelenbegriff keineswegs zwingend die „Survival“-Position, mit der Akzeptanz eines substantiellen Seelenbegriffs ist lediglich die *conditio sine qua non* für die Aufnahme der Diskussion der beiden Alternativhypothesen gegeben, aber keine Vorentscheidung für eine der beiden gewonnen. Die argumentative Tragfähigkeit der beiden Konkurrenzhypothese muss unabhängig geprüft werden.

Was ferner wünschenswert gewesen wäre, ist eine Darstellung der logischen Aporie, in welcher die Jagd nach einem Identitätsbeweis enden muss, d.h., die Tatsache, dass ein solcher Beweis *grundsätzlich* nicht zu liefern ist, wie ich das schon vor 30 Jahren ausgeführt habe (Mulacz 1976). Der Grundgedanke dessen, was ich als „parapsychologischen Relativismus“ bezeichnet habe, ist einfach: nachdem Psi-Phänomene unter Lebenden nie ausgeschlossen werden können, gibt es keine Möglichkeit, auszuschließen, dass gerade jene Information, die zur Verifikation einer angeblich aus dem „Jenseits“ stammenden Information dient, in Wirklichkeit vielleicht die Quelle war, aus welcher (von einem Medium etc.) geschöpft worden ist. Das ist natürlich keineswegs eine neue Erkenntnis, vielmehr hat das schon der Berliner For-

---

4 Gerade der auch von Andreas Sommer zitierte, von Sabom beschriebene Pam Reynolds-Fall stellt sich, wie Woerlee gezeigt hat, keineswegs eindeutig dar und kann daher nicht als entscheidender Durchbruch gewertet werden.

scher Fritz Grunewald (1885-1925) erkannt, der es so ausgedrückt hat: „Ein regelrechtes parapsychologisches Relativitätsprinzip tut sich hier auf“ (Grunewald 1925), ein Gedanke, den mit andern Worten auch der Tübinger Philosoph Traugott Konstantin Oesterreich (1880-1949) ausgesagt hat, indem er von der grundsätzlichen Unbeweisbarkeit der spiritistischen Hypothese (freilich auch ihrer Unwiderlegbarkeit) spricht („Es gibt überhaupt kein Beweisverfahren, das uns zwingen könnte, hinter irgendeinem medialen Produkt einen anderen Geist als Urheber anzunehmen als den Geist des Mediums selber.“) – sämtlich Quellen, die Andreas Sommer in seiner sonst so verdienstvollen Darstellung leider nicht berücksichtigt hat. In einer etwas anderen Sichtweise stellt auch Neal Grossman (2005) die Äquivalenz der Hypothesen „Survival“ und „Super-psi“ dar, insofern, als jeweils die eine Hypothese die andere (mit)enthält.

Erfreulich ist, dass Andreas Sommer den interessanten und aufgrund der Ungunst der bio- und bibliographischen Umstände viel zu wenig rezipierten Ansatz Mattiesens darstellt – dabei räumt er ganz richtig ein, dass „wir alle diese formalen Aspekte des Trancegeschehens auch als unbewusstes ‚Theaterspiel‘ aus dem Traumleben von Trancemedien beschreiben“ können. Allerdings wird dieser Gedanke dann nicht weiter verfolgt, sodass es sich bloß um eine Salvierungsklausel zu handeln scheint. Nicht nur die Trancemedien, wir alle sind in unseren Träumen „Dramatiker“, die eine Vielzahl von Personen auf einer inneren Bühne auftreten lassen. Darüber hinaus sind viele der „Kommunikatoren“ höchst fragmentarisch bleibende Personifikationen (Mulacz 1976), von deren Persönlichkeit und Schicksal wir nicht viel erfahren.

Dabei diskutiert Sommer auch einen Gedanken, den man öfters in der Literatur findet, nicht nur bei Mattiesen, indem er schreibt: „Darüber hinaus werden u.a. spezifische Manierismen, Ausdrucksweisen, humorige Eigenheiten, Temperament, Dialekte oder auch Fremdsprachenkenntnisse, die Handschrift und evtl. künstlerische Fertigkeiten eines bestimmten Verstorbenen beigebracht, also nicht nur biographische, sondern auch psychologische Informationen.“ Diese vermitteln eine „psychologische ‚Dreidimensionalität‘ der sich angeblich manifestierenden Verstorbenen“ – Mattiesen zieht ja aus dieser Lebendigkeit den Schluss, dass es sich bei den Kommunikatoren des „Transdramas“ eben um „Lebende“, nämlich um die im Jenseits *fortlebenden* Verstorbenen handelt. So eindrucksvoll sich das auch alles darstellt, bei Licht besehen ändert es nichts an dem Grundprinzip, dass wir es mit Informationen zu tun haben – ich bezeichne die konkreten Aussagen der sich angeblich manifestierenden Verstorbenen als eine Information erster Ordnung und die Modalitäten ihrer Bekundung als eine Information zweiter Ordnung, um dieser Überlegung Rechnung zu tragen. Schlussendlich ist jedoch beides Information, was vielleicht deutlicher wird, wenn wir an die Möglichkeit einer Digitalisierung denken: sowohl der Inhalt der Aussage wie auch die Modalitäten wie individuelle Sprachfärbung, Klang der Stimme etc. lassen sich in gleicher Weise in Form von Folgen der Ziffern 1 und 0 kodieren. Daraus erhellt, dass die Betrachtung dieser Aspekte kein neues Argument einführt – weder pro noch contra.

Nun ist es natürlich mit der von mir als Titel dieses Kommentars gewählten Feststellung nicht getan. Daher begrüße ich die Anregungen des Autors für zukünftige Forschung auf diesem Gebiet, möchte diese aber noch ergänzen mit dem Hinweis, dass insbesondere ein differenzierender Ansatz innerhalb experimenteller Forschung zur AKE Not tut, und zwar entlang zweier Forschungslinien: Studien von spontanem Auftreten von AKE einerseits und Experimente mit begabten Versuchspersonen andererseits, wie dies etwa Karlis Osis mit dem Ziel der Gewinnung objektiver Messdaten begonnen hat.

### *Nachtrag*

Zwischenzeitlich sich eine kleine Verbesserung der von Andreas Sommer zu Recht beklagten wissenschaftssoziologischen Situation in Hinblick auf die Diskussion der Survival-Frage ergeben, insofern, als bei der letzten Convention der Parapsychological Association in Stockholm (2006) diese Problematik Gegenstand eines „Panel“ gewesen ist, das zu organisieren ich die Ehre hatte (einem Vorschlag Erlendur Haraldssons folgend). Ob das der Beginn einer nachhaltigen Veränderung des „Klimas“ ist, bleibt abzuwarten, aber jedenfalls ist die Bedeutung dieses Themas im Rahmen der doch mehrheitlich experimentell orientierten PA diskutiert worden und es ist somit sozusagen „back on the agenda“. Die Abstracts finden sich unter [http://parapsych.org/pa\\_abstracts\\_2006.html](http://parapsych.org/pa_abstracts_2006.html).

### **Literatur**

- Beloff, J. (1989): Dualism: A parapsychological perspective. In: Smythies, J.R.; Beloff, J. (Eds.): *The Case for Dualism*. University Press of Virginia, Charlottesville, VA.
- Beloff, J. (1990): *The Relentless Question*. McFarland, Jefferson, N.C.
- Grossman, N. (2005): Some Thoughts on Superpsi (That Bogus Pseudo-Scientific Hypothesis). *Journal of Spirituality and Paranormal Studies*, in print.
- Grosso, M. (2006): *Experiencing the Next World Now*. Paraview, N.Y.
- Grunewald, F. (1925): *Mediumismus. Die physikalischen Erscheinungen des Okkultismus*. Ullstein, Berlin.
- Fontana, D. (2005): *Is there an Afterlife? A comprehensive Overview of the Evidence*. O-Books, Hants.
- Mulacz, P. (1976): Der sogenannte wissenschaftliche Spiritismus als parapsychologisches Problem. In: Schatz, O. (Hrsg.): *Parapsychologie. Ein Handbuch*. Styria, Wien/Graz/Köln.
- Mulacz, P. (2006): The Importance of Survival research. Basic issues of the 'survival' question. The Parapsychological Association, 49th Annual Convention, Proceedings of Presented Papers.
- Oesterreich, T.K. (1920): *Der Okkultismus im modernen Weltbild*. Sibyllen, Dresden.
- Oesterreich, T.K. (1924): *Die philosophische Bedeutung der mediumistischen Phänomene*. Kohlhammer, Stuttgart.

Wiesner, H. P.; Thouless, R. H. (1947): The Psi Processes in Normal and 'Paranormal' Psychology. *Proceedings of the Society for Psychical Research*, part 174, vol. XLVIII.

Woerlee, G.M. (2005): *Mortal Minds: The Biology of Near Death Experiences*. Prometheus Books, N.Y.

### **Autorenantwort:**

ANDREAS SOMMER<sup>5</sup>

### **Die FüÙe des Kolosses – weder tönern noch ehern**

Peter Mulacz versteht meine Arbeit offensichtlich als den Versuch einer empirischen Untermauerung eines kartesischen Substanz-Dualismus. Verantwortlich für dieses Missverständnis ist möglicherweise mein Hinweis auf die Unvereinbarkeit eines kartesischen Seelenbegriffs mit gewissen neurowissenschaftlichen Befunden gleich im ersten Satz. Es ist anzunehmen, dass damit auch bei anderen Lesern der Eindruck erweckt wurde, der nachfolgende Text verteidige einen Dualismus kartesischer Spielart, während ich lediglich auf das Spannungsfeld zwischen dem folkloristischen (substanzdualistischen) Seelenbegriff und dem reduktionistischen Bewusstseinsverständnis wichtiger neurowissenschaftlicher Theoretiker hinzuweisen beabsichtigte.

Für die Gelegenheit der Ausräumung dieses Missverständnisses bin ich Prof. Mulacz dankbar, da ich mich nicht als Vertreter eines ontologischen Dualismus verstehe. Der aufmerksame Leser wird feststellen, dass im ganzen folgenden Text keine Rede mehr von „Substanz“ (der Begriff „kartesisch“ fehlt wohlweislich im gesamten Text) im Zusammenhang mit dem auch weiterhin gelegentlich verwendeten Begriff „Seele“ ist, den ich – synonym mit „Psyche“ und „Bewusstsein“ – pragmatisch für die Summe individuellen Wollens, Verhaltens, Erlebens, Erinnerns, Identitätsempfindens usw. verwende.

Was die Epistemologie und Ontologie der Seelenfrage angeht, so stimme ich darin zu, dass eine Übersicht zum Stand der philosophischen Diskussion zum Leib-Seele-Problem in einer erschöpfenden Behandlung der Survival-Forschung wünschenswert ist. Eine Übersicht der philosophischen Positionen und Argumente – geschweige denn deren „philosophische Durchdringung“ – als bloßer Abschnitt eines Zeitschriftenbeitrages ist jedoch schwierig zu bewerkstelligen und mir ist bewusst, dass der Umfang meiner Arbeit so wie sie ist bereits eine Zumutung an viele chronisch zeitarme Leser darstellt.

Der Hauptgrund meiner Auslassung einer philosophischen Diskussion des Leib-Seele-Problems liegt jedoch darin, dass ich folgende Auffassung von Mulacz nicht teilen kann: „So-

---

5 Andreas Sommer ist am Centre for the Study of Anomalous Psychological Processes der University of Northampton in Großbritannien tätig. E-mail: sommer@survival-research.net.

lange diese Grundfragen nicht geklärt sind, bleibt jede Diskussion der ‚Survival‘-Problematik notwendigerweise konditional“. Demnach wäre die philosophische Lösung des Leib-Seele-Problems Voraussetzung zur Interpretation der Daten aus der Survival-Forschung (oder in Prof. Mulacz Gleichnis: Die Füße des Daten-Kolosses müssen eine einwandfreie theoretische Basis sein – wenn auch wenige Zeilen vorher noch der empirischen Forschung der Primat zugeschrieben wird). Dieser Meinung vermag ich mich schon darum nicht anzuschließen, weil sie in letzter Konsequenz so lange die Einstellung empirischer Forschungen zur Survival-Frage fordern muss, bis sich jemand ein theoretisches Modell ausdenkt, das hypothetisches Weiterleben zweifelsfrei erklärt und voraussagt. Merkwürdigerweise wird gleichzeitig von der Existenz parapsychologischer Effekte ausgegangen, da „Psi-Phänomene unter Lebenden nie ausgeschlossen werden können“, wobei die eigenen wissenschaftstheoretischen Grundsätze vergessen zu werden scheinen. Denn obiger Forderung zufolge dürften eigentlich überhaupt keine parapsychologischen Phänomene vorkommen, weil es in der Vielfalt theoretischer Modelle der Parapsychologie kein einziges gibt, das sich als das zweifellos richtige bewährt hat.

Eine wichtige wissenschaftshistorische Erkenntnis darf bei alledem nicht vergessen werden, nämlich dass empirische Daten eben auch Theorien *schaffen* können – Physiker gehen z.B. von der Existenz von Gravitonen aus, masselose Teilchen, die zum Verständnis der Gravitation notwendig sind, ohne je in der Lage gewesen zu sein, die Existenz eines Gravitons empirisch zu beweisen (Walach 2005, S. 156). Stephen Braude (2003) gibt in diesem Zusammenhang zu bedenken, dass sich in der Wissenschaft z.B. kaum jemand bei der Entdeckung radikal neuer Stoffeigenschaften um frühere philosophische Debatten zur Ontologie der Materie zu kümmern pflegt. Darum, so Braude, sei es im Falle einer entsprechenden Datenlage wissenschaftlich legitim, „Seelen“ zu postulieren, auch wenn wir über noch keine sichere Theorie verfügen, die hypothetisches Weiterleben verständlich machen kann: „If the best survival data, and the totality of good evidence, can't be handled satisfactorily along non-survivalist lines, we may have no choice but to entertain a variety of metaphysical theses we would previously have dismissed out of hand“ (Braude 2003, S. 297). Die Betonung liegt hier freilich auf dem „Wenn“ der ausreichend großen Summe und hohen Qualität der Daten, die zwar von den meisten Kritikern der Survival-Hypothese für gewöhnlich unterschätzt (soweit überhaupt zur Kenntnis genommen) werden, aber – und hier stimme ich mit Peter Mulacz wieder überein – das Weiterleben nach dem Tod bisher nicht zwingend beweisen.

Widersprechen möchte ich in Bezug auf eine rein philosophische Auseinandersetzung mit der Frage des Fortlebens auch der Auffassung, wonach „nur auf der Basis der Annahme eines substantiellen Seelenbegriffs [...] die ganze Diskussion überhaupt möglich“ sei. Eine neutralmonistische Auffassung des Leib-Seele-Problems im Sinne Whiteheads (1929), der ich selber zuneige, ist mit der Hypothese des Fortlebens mindestens kompatibel (zur Brauchbarkeit eines Whiteheadschen Bewusstseinsmodells zum theoretischen Verständnis hypothetischen Weiterlebens siehe z.B. Griffin 1997).

Meine Wahl des Begriffs „Survival“ wird bemängelt, da es in der ganzen Diskussion eigentlich nicht konkret um die Frage des Fortlebens gehe, sondern „um die (angebliche) Kommunikation zwischen ‚Jenseitigen‘ und Diesseitigen“. Dies mag zwar auf einen gewissen Teil der Evidenzklassen zutreffen, vor allem auf mediumistische Phänomene, macht aber eine Einordnung anderer wichtiger Kategorien, z.B. angebliche Erinnerungen an frühere Leben, nicht-responsive Erscheinungen Lebender und Verstorbener, AKE mit verifizierbaren Wahrnehmungen, sowie gewisse phänomenologische und konzeptuelle Überlegungen zum Thema usw., problematisch.

Mein Stillschweigen zur – meines Erachtens mit großer Vorsicht zu genießenden – katholischen „Arme-Seelen“-Literatur und ähnlichen anekdotischen Berichten beruht auf der Einsicht, dass diese allenfalls heuristischen Wert haben können und in einem Zeitschriftenbeitrag zu empirischen Aspekten der Survival-Forschung höchstens als Fußnote Erwähnung finden sollten.

Verblüfft bin ich über die Feststellung, wonach ich versäumt haben soll, die Problematik der Ausschlussmöglichkeiten von hypothetischem Psi unter Lebenden darzustellen, was doch ausführlich in der Diskussion geschehen ist (Sommer 2005, S. 211-223). Stattdessen hätte ich mir von Prof. Mulacz eine kritische Stellungnahme zu meiner Einschätzung von Braudes „Argument aus der lähmenden Komplexität“ als erstes brauchbares theoretisches Eingrenzungskriterium von Superpsi gewünscht.

Dass ich zwei deutschsprachige Autoren aus den 1920ern Jahren, Grunewald und Oesterreich, in der Behandlung der Superpsi-Hypothese nicht berücksichtigt habe, liegt daran, dass keiner der beiden Autoren – die in anderen Zusammenhängen besser zu würdigen sind – je etwas Entscheidendes zur Superpsi-Problematik beigetragen hat, die schon seit den frühen 1880ern Forschern auf unserem Gebiet Kopfzerbrechen bereitet.

Die Vermutung, ich habe den im Kommentar zitierten Satz zur Relativierung der Argumente Mattiesens zugunsten der Eigenständigkeit mancher Trancepersönlichkeiten als „Salvierungsklausel“ eingesetzt, ist unzutreffend, da sich in der Diskussion auch meine Warnung vor der Verwechslung dramatisierter Inhalte aus dem Unbewussten mit eigenständigen Persönlichkeiten findet, die ich relativ ausführlich und in Verbindung mit der Superpsi-Hypothese darstelle und diskutiere (Sommer 2005, S. 211-214, 218-222). Zwar freue ich mich über Prof. Mulacz Lob darüber, dass ich Mattiesens Werk in Erinnerung zu bringen versucht habe, aber wenn Mattiesens drei Bände umfassende Feinanalyse und Argumente anscheinend mit wenigen Sätzen zu entkräften sind – wie schon vor dreißig Jahren geglaubt zu werden schien (Mulacz 1976, S. 208, 243, 253-256) – warum ist Mattiesens Ansatz dann „viel zu wenig rezipiert“?

Angenommen, die originelle Unterscheidung zwischen Informationen erster und zweiter Ordnung ist zutreffend und auch eine Digitalisierung der Idiosynkrasien menschlichen Verhaltens ist möglich (wobei mir sowohl der verwendete Informationsbegriff als auch letztere

Annahme als beliebig und simplizistisch erscheinen): Die Gültigkeit von Braudes „Argument aus der lähmenden Komplexität“ vorausgesetzt – und bisher ist nichts ins Feld gebracht worden, um das Argument zu entkräften –, würden Kategorisierungsversuche außersinnlich erworbener Informationen in unterschiedliche Ordnungen wenig an den inhärenten theoretischen Schwächen von Superpsi ändern.

Die Überschrift „150 Jahre Stagnation“ beschreibt die Situation meiner Ansicht nach nicht präzise. Zwar muss man sicher auch nicht den Optimismus von William Roll teilen, einem der bewährtesten Forscher auf unserem Gebiet, demzufolge eine Entscheidung der Survival-Frage in Aussicht steht (Roll 2006, S. 170), doch haben sich die Daten seit Gründung der Society for Psychical Research schon alleine durch die vielen durch unabhängige Replikationen bestätigten Befunde Ian Stevensons zu angeblichen „früheren Leben“ erheblich vermehrt, und auch gewisse konzeptuelle Fortschritte sind zu verzeichnen (so z.B. durch Stephen Braudes und andere wichtige philosophische Arbeiten). Setzen wir die kümmerliche Zahl der heute auf unserem Gebiet tätigen Forscher in Relation zur Summe der neueren Fakten und Gedanken, kann von einer Stagnation meines Erachtens kaum die Rede sein, wenn sich die forschungspolitische Situation auf unserem Gebiet auch eher verschlechtert hat.

### Literatur

- Braude, S.E. (2003): *Immortal Remains: The Evidence for Life After Death*. Rowman & Littlefield, Lanham.
- Griffin, D.R. (1997): *Parapsychology, Philosophy, and Spirituality: A Postmodern Exploration*. State University of New York Press, Albany, NY.
- Mulacz, P. (1976): Der sogenannte wissenschaftliche Spiritismus als parapsychologisches Problem. In: Schatz, O. (Hrsg.): *Parapsychologie. Ein Handbuch*. Styria, Graz, 178-283.
- Roll, W. (2006): On apparitions and mediumship: an examination of the evidence that personal consciousness persists after death. In: Storm, L.; Thalbourne, M. (Hrsg.): *The Survival of Human Consciousness. Essays on the Possibility of Life After Death*. McFarland, Jefferson, 142-173.
- Sommer, A. (2005): Im Niemandsland der Bewusstseinsforschung: Survival Research. Eine Übersicht. *Zeitschrift für Anomalistik* 5, 178-237.
- Walach, H. (2005): *Psychologie – Wissenschaftstheorie, philosophische Grundlagen und Geschichte. Ein Lehrbuch*. Kohlhammer, Stuttgart.
- Whitehead, A. N. (1929): *Process and Reality: An Essay in Cosmology*. Macmillan, New York.

**Kommentar zum Aufsatz von Suitbert Ertel  
„Astrologie und Psi. Eine Fallstudie verstärkt die Zusammenhangshypothese“  
In: *Zeitschrift für Anomalistik* 4 (2004), 52-68  
und zur Antwort von Suitbert Ertel (Bd. 5, 290-295) auf den Kommentar  
von Volker Guiard (Bd. 5, 282-289)**

VOLKER GUIARD<sup>6</sup>

**Das „größte methodische Missverständnis“ sollte endgültig vom Tisch**

Im Anschluss an den oben genannten Aufsatz von Ertel (2004a) – ich übernehme hier weiterhin die von Ertel verwendete Abkürzung AuP – kritisierte ich (Guiard 2004) Ertels Anwendung der Statistik. Inzwischen wurde die Diskussion mit den in der Überschrift genannten Artikeln fortgesetzt (Ertel 2004b; Guiard 2005; Ertel 2005). Im weiteren werde ich mich auf die letzte Antwort von Ertel beziehen.

Bevor ich zu den Einzelheiten komme, will ich zunächst etwas in Erinnerung rufen, was vermutlich in Vergessenheit geriet. Ertel (2005) schreibt in seiner Fußnote 19 über mich, „dass er sich entschieden hat, bei seiner Bewertungsarbeit aus der Nische seiner Spezialisierung nicht herauszutreten und Inhaltliches auszublenden“. Das ist durchaus korrekt. Weiter schreibt Ertel: „Doch dann sollte er seine Enthaltbarkeit offen bekennen“. Diese berechtigte Forderung hatte ich in meinem ersten Kommentar erfüllt. Ich schrieb dort auf der ersten Seite: „In meinem Kommentar gehe ich nur auf methodisch-statistische Aspekte ein. Es gäbe sicherlich viele weitere Dinge zu diskutieren, ich hoffe jedoch, dass dieses von anderen Kommentatoren übernommen wird.“ Da ich mich auf dem von Ertel behandelten Gebiet durchaus nicht kompetent fühle, wäre es unangebracht, hier auch inhaltliche Kommentare zu geben, zumal es in der Gesellschaft für Anomalistik genügend Fachleute gibt, die den Inhalt viel besser beurteilen können, so wie es dann auch mit dem Kommentar von Boller (2005) geschah. Ich halte es aber durchaus für sinnvoll, die statistische Methodik zu bewerten, denn auch dann, wenn die Aussagen eines Artikels noch so interessant sind, wird der Wert des Artikels mitunter eingeschränkt, wenn die Beweiskraft auf Grund der problematischen Handhabung der statistischen Methodik gemindert ist. In diesem Fall war aber scheinbar auch der Inhalt fragwürdig, wenn ich den Kommentar von Boller (2005) richtig verstanden habe.

---

6 PD Dr. Volker Guiard arbeitet am Forschungsbereich Genetik und Biometrie des Forschungsinstituts für die Biologie landwirtschaftlicher Nutztiere in Dummerstorf bei Rostock. E-Mail: [guiard@anomalistik.de](mailto:guiard@anomalistik.de).

*Konfirmatorische oder exploratorische Tests?*

Da ich die von Ertel verwendeten statistischen Tests im konfirmatorischen Sinne verstand, wies Ertel in seiner letzten Antwort (Ertel 2005) darauf hin, dass er die im Untertitel zu AuP genannte Zusammenhangshypothese nur im exploratorischen Sinne untersucht hat. Sie wurde nicht als bestätigt hingestellt, sondern es wurden Vorschläge gemacht, wie diese Hypothese noch systematischer untersucht werden könne. Zur Zusammenhangshypothese schreibt Ertel im Abstract von AuP: „...wird die weiter zu prüfende Hypothese nahe gelegt, dass herausragende Fälle von astrologisch richtigen Deutungen ... möglicherweise auf hellseherischem oder telepathischem Wege zustande kommen ...“. Zu dieser Hypothese gibt es in AuP jedoch keinen statistischen Test, weswegen ich auf diese Hypothese aus statistischer Sicht nicht einging. Getestet werden nur die beiden Seiten dieses Zusammenhanges, also die astrologischen Fähigkeiten einerseits und die Psi-Fähigkeiten andererseits. Ob diese Tests nun konfirmatorisch oder exploratorisch gemeint sind, ist damit noch nicht gesagt. In meinem ersten Kommentar (Guiard 2004) schrieb ich, dass ich bei solchen Tests, bei denen Ertel die Bonferroni-Korrektur anwendete, annahm, dass Ertel diese Tests im konfirmatorischen Sinne verwendete. In seiner damaligen Antwort dazu hätte Ertel mir widersprechen können, was er aber nicht tat. Insofern musste ich auch weiterhin von der Regel ausgehen, dass bei vorhandener Bonferroni-Korrektur ein konfirmatorischer Test vorliegt. Aus dieser Sicht war die Anwendung oder Nicht-Anwendung der Bonferroni-Korrektur aber nicht immer konsequent. Ausnahmen wurden nicht regelmäßig begründet, wie Ertel (2005) in seiner Fußnote 15 schreibt. Außerdem war seine in dem einleitenden Text zu den Ergebnissen des Ballzieh-Tests in AuP beschriebene Regel zur Anwendung der Bonferroni-Korrektur falsch, wie ich bereits erläuterte (Guiard 2004, S. 72). Auch die statistischen Tests zur astrologischen Zuordnung interpretierte ich als konfirmatorische Tests. Hierzu gab es von Ertel bisher auch keinen Widerspruch.

*Die Richtung eines einseitigen konfirmatorischen Tests ist vor dem Versuch festzulegen*

Ertel (2005, S. 291) schreibt: „Das offensichtlich größte methodische ‚Missverständnis‘ unter Punkt 4 von Guiards jetzigem Beitrag sollte endgültig vom Tisch, zumal auch andere Forscher sich gelegentlich vor die dort behandelte Frage gestellt sehen können: *Wann darf man bei beobachteten Abweichungen von der Zufallserwartung, deren Richtung der Alternativ-Erwartung widerspricht, einen einseitigen Signifikanztest anwenden?*“

Diese Bemerkung bezieht sich auf einen Test, der von Ertel (in AuP) als rechtsseitiger Test geplant wurde. Als die untersuchte Trefferanzahl nun aber nicht, wie erwartet, zu groß, sondern eher zu klein ausfiel, führte Ertel unter Berufung auf Kimmel (1957) einen linksseitigen Test durch. Unter explorativem Gesichtspunkt könnte man das tolerieren. Der damalige Text von Ertel erweckte aber eher den Eindruck, dass die Tests im konfirmatorischen Sinne gemeint seien. Nun könnte Ertel sicherlich diesen Eindruck korrigieren und betonen, dass es

sich hierbei um exploratorische Tests handelte. Auch dann wäre aber die folgende Diskussion von Bedeutung, wenn auch nicht so sehr in Bezug auf den konkreten Artikel AuP, so doch aus allgemeinem methodischen Interesse gemäß der obigen Formulierung von Ertel: „zumal auch andere Forscher sich gelegentlich vor die dort behandelte Frage gestellt sehen können“. Ich gehe also in der folgenden Diskussion davon aus, dass es sich um einen konfirmatorischen Test handelt, auch wenn dieses in dem konkreten Artikel AuP nicht zutreffen sollte.

Ich versuchte auch in meinen vorhergehenden Kommentaren darzulegen, dass die Richtung eines konfirmatorischen einseitigen Tests vor dem Versuch festzulegen ist. Das dabei verwendete Beispiel bezeichnete Ertel als tendenziös. Ein anderes Beispiel, welches ich in der direkten Kommunikation mit Ertel nutzte, gab Ertel in seinem letzten Kommentar „verkürzt“ wieder, wobei mein ursprüngliches Beispiel in seiner Darstellung nicht mehr zu erkennen war. Da also meine Beispiele aus unerfindlichen Gründen von Ertel nicht akzeptiert werden, werde ich weiterhin nur Beispiele von Kimmel (1957) verwenden, in der (unsicheren) Hoffnung, dass Ertel nicht auch diese als tendenziös bezeichnet.

Was ist nun der Inhalt des Artikels von Kimmel? Kimmel diskutiert die Sinnhaftigkeit einseitiger Tests, wobei er vorrangig für zweiseitige Tests wirbt. Er akzeptiert aber, dass es seltene Fälle gibt, bei denen ein einseitiger Test sinnvoll ist. Weiterhin gibt er drei Kriterien an, um solche Fälle zu erkennen. Zum Schluss schreibt er: „Opponents of one-tailed tests ... should welcome this attempt to limit the use of one-tailed tests to those infrequent situations provided for by the proposed criteria“.

Um vor einseitigen Test zu warnen, schreibt Kimmel (S. 352): „A two-tailed null hypothesis can be rejected by a large observed difference in either direction but a one-tailed null hypothesis cannot be rejected by a difference in the unpredicted direction, no matter how large this difference may be. This means that an experimenter using a one-tailed hypothesis cannot conclude that an extreme difference in the unpredicted direction is reliable different from zero difference. This limitation cannot be shrugged off by the comment 'We have no interest in a difference in the opposite direction,' scientists are interested in empirical fact regardless of its relationship to their preconceptions.“ Im letzten Satz (welcher ebenfalls von Ertel zitiert wurde, jedoch mit anderer Schlussfolgerung) betont Kimmel, dass ein Wissenschaftler sich auch für ein unerwartetes Ergebnis interessieren sollte. Da ein einseitiger Test, wie er in den vorhergehenden Sätzen schreibt, eine Prüfung solcher Ergebnisse aber nicht zulässt, wäre die Konsequenz, von vornherein einen zweiseitigen Test zu verwenden, um für solche Fälle gewappnet zu sein.

Anschließend schreibt Kimmel, dass ein einseitiger Test z.B. die Frage beantworten soll, ob ein neues Produkt besser ist als das bisher hergestellte. Er hält es aber auch hier für wünschenswert, wenn man sich nicht nur für die Fälle „besser“ und „nicht besser“ entscheiden kann, sondern auch für „schlechter“. Für die Entscheidung, ob das neue Produkt vermarktet wird, würde zwar die Unterscheidung zwischen „besser“ und „nicht besser“ genügen, der Fall „schlechter“ würde aber zusätzliche Konsequenzen nach sich ziehen, die von Kimmel jedoch

nicht weiter ausgeführt werden. Um aber auch das Testergebnis „schlechter“ zu ermöglichen, ist also von vornherein ein zweiseitiger Test vorzusehen.

Es folgen nun bei Kimmel die Beschreibungen der drei Situationen, bei denen er einen einseitigen Test akzeptieren würde. Von einer nachträglichen Veränderung der Testrichtung in Abhängigkeit von den Daten ist aber nirgends die Rede. Im Gegenteil: am Ende der ersten Situationsbeschreibung verbietet er sogar die datenabhängige „Metamorphose“ eines einseitigen Tests in einen zweiseitigen. Die Festlegung des Testtyps hat vor dem Versuch stattzufinden.

Die Beschreibung der zweiten Situation beginnt mit dem Text, welcher von Ertel in AuP ebenfalls zitiert wurde (in seiner Fußnote 5): „Use the one-tailed test when results in the unpredicted direction will, under no conditions, be used to determine a course of behavior different in any way from that determined by no difference at all“. Diese Situation würde also dem obigen Beispiel von Kimmel entsprechen, sofern der Fall „schlechter“ – anders, als oben angenommen – keinerlei zusätzliche Handlungskonsequenzen nach sich ziehen würde, so dass als gemeinsame Konsequenz der Fälle „schlechter“ und „gleichwertig“ lediglich die Nichtvermarktung des neuen Produktes folgt. Die beiden Fälle „schlechter“ und „gleichwertig“ können dann zu „nicht besser“ zusammengefasst werden, womit der Fall „nicht besser“ dann als „einseitige“ Nullhypothese dienen kann. Ertel leitete hieraus jedoch ab, dass ein rechtsseitiger Test in einen linksseitigen verwandelt werden dürfte, um eine unerwartete Abweichung zu testen. Davon ist aber bei Kimmel keine Rede. Einseitige Tests werden bei ihm nie dazu genutzt, eine Abweichung in der unerwarteten Richtung zu testen.

Oft wird das Hypothesenpaar für einseitige Tests in der Form  $H_0: \Delta=0$  und  $H_A: \Delta>0$  formuliert, wenn mit  $\Delta$  der Erwartungswert des Qualitätsunterschiedes zwischen den beiden Produkten bezeichnet wird. Wenn nun dieser Qualitätsunterschied negativ ausfällt, also das neue Produkt bedeutend schlechter zu sein scheint als das alte, dann ist es sicherlich schwer einzusehen, dass man trotzdem die Nullhypothese  $\Delta=0$ , also die Gleichwertigkeit beider Methoden akzeptieren soll. Dieses ist aber ein Missverständnis, welches durch die ungünstige Formulierung der Nullhypothese verursacht wird. Es ist sinnvoller, eine einseitige Hypothese in der Form  $H_0: \Delta \leq 0$ <sup>7</sup> anzugeben, was unserem obigen „nicht besser“ entspricht. Auch wenn das neue Produkt bedeutend schlechter sein sollte als das alte, so kann trotzdem diese einseitige Nullhypothese nicht abgelehnt werden, da der Fall  $\Delta < 0$  in dieser Nullhypothese bereits enthalten ist. Diese Hypothese setzt sich also zusammen aus den Fällen „gleichwertig“

---

7 Bei dieser Hypothese ist das Risiko erster Art von dem tatsächlichen Wert von  $\Delta$  abhängig. Da dieses Risiko aber in jedem Fall nicht größer als ein vorgegebenes  $\alpha$  sein soll, muss diese Forderung auch im ungünstigsten Fall gelten, welcher vorliegt, sofern  $\Delta=0$  gilt. Damit dient also die erste Formulierung der einseitigen Nullhypothese zur Berechnung des größten Wertes des Risikos erster Art, welcher dann mit  $\alpha$  identifiziert wird.

und „schlechter“. Mit der (vor dem Versuch erfolgten) Wahl dieser einseitigen Nullhypothese legt sich der Experimentator darauf fest, dass die Unterscheidung zwischen „gleichwertig“ und „schlechter“ nicht Gegenstand seiner Untersuchung ist, ihm geht es nur um die Unterscheidung zwischen „besser“ und „nicht besser“. Ob diese Wahl sinnvoll ist, kann man diskutieren. Kimmel hält sie nur selten für sinnvoll und gibt dafür drei Situationen an.<sup>8</sup> Sollte der Experimentator nach entsprechender Diskussion nun doch zu dem Entschluss kommen, sich auch für die Unterscheidung zwischen „gleichwertig“ und „schlechter“ zu interessieren, dann muss er also von vorneherein einen zweiseitigen Test vorsehen.

Bei seinen Argumentationen stellte Ertel (2005, S. 291) auch die Behauptung auf, dass ein p-Wert nicht größer als 0,5 sein kann. Dieser Irrtum ist nachvollziehbar, da p-Werte meistens nur dann angegeben werden, wenn – nahezu – Signifikanz vorliegt. Würde man bei Tests jedoch in jedem Fall den p-Wert angeben, so würde man schnell sehen, dass diese auch größer als 0,5 sein können. Man könnte z.B. die p-Werte auch für Metaanalysen verwenden. Solche Analysemethoden „leben“ davon, dass unter der Nullhypothese die p-Werte zwischen 0 und 1 gleichverteilt sind (und nicht zwischen 0 und 0,5). Eine solche Methode wäre z.B. die Zusammenfassung der p-Werte einzelner Tests nach der Formel von R.A. Fisher:  $\chi^2 = (-2) \times \sum(\log(p))$ . Diese Methode betrachtete auch Ertel (2004) in seiner Fußnote 11 als geeignet.

---

8 In einer der Situationen, bei denen Kimmel einen einseitigen Test für sinnvoll erachtet, geht er davon aus, dass der Fall  $\Delta < 0$  für absolut unmöglich gehalten wird. In diesem Fall wäre auch die Formulierung  $H_0: \Delta = 0$  sinnvoll. Wegen dieser Annahme wäre dann ein Test nur in positiver Richtung angebracht, auch wenn die noch von Zufallseffekten abhängigen Qualitätsunterschiede negativ ausfallen sollten. Möchte man jedoch auch prüfen, ob der Fall  $\Delta < 0$  eventuell doch möglich ist, so müsste man sich vor dem Versuch für einen zweiseitigen Test entscheiden. Man könnte allerdings die beiden Seiten der Alternativhypothese unterschiedlich wichten, da einerseits das Hauptziel des Versuches in dem „Nachweis“ von  $\Delta > 0$  bestand (rechte Seite) und andererseits der eventuelle „Nachweis“ von  $\Delta < 0$  zur Korrektur der Theorie, welche den Fall  $\Delta < 0$  verbot, führen würde, aber eine anerkannte Theorie möchte man nicht leichtfertig aufgeben. Es würde sich also anbieten, für beide Seiten unterschiedliche Irrtumswahrscheinlichkeiten  $\alpha_{\text{links}}$  und  $\alpha_{\text{rechts}}$  vorzugeben mit  $\alpha_{\text{links}} + \alpha_{\text{rechts}} = \alpha$ , wobei  $\alpha_{\text{links}}$  sehr klein gewählt wird, so dass nur bei sehr starker Evidenz gegen die Theorie diese in Frage gestellt wird. Es wäre nachdenkenswert, ob dieses Vorgehen vielleicht für das Problem von Ertel angebracht wäre.

Ertel scheint sich aber aus einem bestimmten Grunde vor dem zweiseitigen Test zu scheuen. In seinen allgemeinen Betrachtungen zur Anwendung von Tests (S. 292) schreibt er, dass das signifikante Ergebnis eines zweiseitigen Tests bedeuten würde, dass zwei sich widersprechende Theorien gelten würden. Diese Behauptung beruht jedoch auf einem Missverständnis. Die Interpretation der Signifikanz bei einem zweiseitigen Test hängt sehr wohl von der Richtung der beobachteten Abweichung ab. Aus  $\Delta \neq 0$  wird also nicht geschlossen, dass  $\Delta > 0$  und  $\Delta < 0$  gilt, sondern nur eine der Aussagen  $\Delta > 0$  oder  $\Delta < 0$  wird angenommen, wobei die Auswahl der Aussage von dem Vorzeichen der Abweichung abhängt.

Der p-Wert gibt stets unter Annahme der Nullhypothese die Wahrscheinlichkeit an, dass man das im jeweiligen Versuch vorhandene Testergebnis erhält oder ein „noch extremeres“. Dabei ist „noch extremer“ in Richtung der Alternativhypothese gemeint. Bei einem rechtsseitigen t-Test kann man also „noch extremer“ mit „noch größer“ identifizieren. Hätte man nun das Ergebnis  $t=0$  erhalten, so wären noch größere t-Werte mit Wahrscheinlichkeit  $p=0,5$  möglich gewesen. Ist jedoch  $t<0$ , so ist auch entsprechend  $p>0,5$ .

Analog verhält es sich bei einem zweiseitigen t-Test. Dabei bedeutet „extremer“, dass t weiter von der Null entfernt ist. Hier wird zunächst ein „roher“ p-Wert  $p_{\text{roh}}$  wie bei einem einseitigen Test berechnet, und zwar bei  $t>0$  rechtsseitig bzw. bei  $t<0$  linksseitig. Dabei gilt also  $p_{\text{roh}} \leq 0,5$ . Den endgültigen p-Wert errechnet man nun aus  $p=2 p_{\text{roh}}$ . Damit würde man also im Fall  $t=0$   $p_{\text{roh}}=0,5$  und damit  $p=1$  erhalten.

Was ist denn nun die Konsequenz, wenn Ertel bei einseitigen Tests einen ergebnisabhängigen Richtungswechsel zulässt? Das würde bedeuten, dass hier zwar ständig von einseitigen Tests die Rede ist, obwohl es sich eigentlich um einen zweiseitigen Test (mit einem Signifikanzniveau, also Risiko erster Art von  $2 \alpha$  anstelle von  $\alpha$ ) handelt, denn es sind jeweils in beiden Richtungen signifikante Abweichungen möglich. Gilt die Nullhypothese  $\Delta=0$ , dann ist die Wahrscheinlichkeit, diese Nullhypothese nach rechts bzw. nach links abzulehnen, jeweils  $\alpha$ . Insgesamt ist die Ablehnungswahrscheinlichkeit von  $H_0$  also gerade  $2 \alpha$ .

Ertel (2005, S. 291) fragt: „Was für einen Sinn macht eine einseitige Signifikanzprüfung, wenn die beobachteten Abweichungen die falsche Richtung haben?“ Diese Frage steht aber nicht an, denn der Entschluss, einen einseitigen Test durchzuführen, wird vor dem Versuch gefasst, wenn also noch nicht bekannt ist, ob die Ergebnisse erwartungswidrig ausfallen werden. Dabei lautet die Versuchsfrage, ob (z.B.) die neue Therapie besser ist als die alte oder nicht. Sollten die Ergebnisse in der erwarteten Richtung liegen, dann wird mit dem Testverfahren entschieden, ob diese Abweichung für Signifikanz ausreicht. Liegt die Abweichung jedoch in der unerwarteten Richtung, dann erübrigt sich die Berechnung der Testgröße, da dann das Testergebnis „Nichtsignifikanz“ (in Bezug auf die primäre Versuchsfrage) bereits klar ist. Eine Berechnung des Tests könnte dann aber auch dann noch sinnvoll sein, falls man einen p-Wert für spätere Metaanalysen angeben möchte.

Ertel (2005, S. 292) schreibt: „Hat eine Abweichung vom Zufallswert die erwartete positive Richtung, ohne aber signifikant zu sein (Nullhypothese wird beibehalten), dann führt man dies meist auf störende Bedingungen, auf zu wenige Beobachtungsfälle oder auf eine geringe Effektstärke zurück.“ Selbstverständlich wird man es dem Entwickler einer neuen Therapie zugestehen, dass er darüber nachdenkt, warum der Vorteil seiner hoffnungsvollen Therapie nicht nachweisbar war. Immerhin bedeutet die fehlende Signifikanz nicht unbedingt, dass keine Effektverbesserung vorliegt, sondern nur, dass eine eventuelle Effektverbesserung nicht nachweisbar war. Der Entwickler wird also sowohl über eine weitere Verbesserung der Therapie, als auch über einen neuen, verbesserten Versuch nachdenken, der so strukturiert ist, dass Störeffekte möglichst keinen Einfluss haben. Die oben genannten Argumente „zu wenige

Beobachtungsfälle“ und „zu geringe Effektstärke“ wird er entkräften, indem er sich überlegt, ab welcher Effektstärke die neue Therapie denn überhaupt von praktischem Interesse sein wird. Die Anzahl der Beobachtungsfälle plant er dann so, dass bei angenommener „interessanter“ Effektstärke das Risiko zweiter Art  $\beta$  vertretbar klein ist, so dass bei Nichtsignifikanz dann gefolgert werden kann, dass der Effekt der neuen Therapie gegenüber dem der alten sich – wenn überhaupt – nur unwesentlich unterscheidet.

Die kritische Betrachtung der Versuchsmethodik kann zwar zu einer Verbesserung zukünftiger Versuche führen, sie ändert aber nichts daran, dass der aktuell vorliegende Versuch keine Evidenz für einen besseren Therapieeffekt lieferte.

Außerdem ist eine solche kritische Prüfung der Versuchsmethodik nicht nur bei ausbleibender Signifikanz angebracht. Auch bei Signifikanz ist zu überlegen, ob der Verbesserungseffekt nicht durch methodische Fehler „erzeugt“ wurde, z.B. durch unzureichende Verblindung.

### Literatur

- Boller, E. (2005): Einwände gegen die Psi-Interpretation der Einzeltests in Ertels Fallstudie. *Zeitschrift für Anomalistik* 4, 68-71.
- Ertel, S. (2004a): Astrologie und Psi. *Zeitschrift für Anomalistik* 4, 52-68.
- Ertel, S. (2004b): Kritik sollte korrigieren, nicht demolieren. *Zeitschrift für Anomalistik* 4, 85-101.
- Ertel, S. (2005): Aus Unverständnis wird Missverständnis. *Zeitschrift für Anomalistik* 5, 290-296.
- Guiard, V. (2004): Statistik mangelhaft. *Zeitschrift für Anomalistik* 4, 71-80.
- Guiard, V. (2005): Missverständnisse bezüglich statistischer Auswertung, *Zeitschrift für Anomalistik* 5, 126-133.
- Kimmel, H.D. (1957): Three criteria for the use of one-tailed tests. *Psychological Bulletin* 54, 351-353.

**Autorenantwort:**SUITBERT ERTEL<sup>9</sup>**Umgang mit Überraschungen. Ergänzendes zu Guiards Kommentar über ein- und zweiseitige Signifikanzen**

Mir fällt es schwer, mit Guiards Auffassungen über die Gerichtetheit der Signifikanzprüfungen einverstanden zu sein. Guiard entwickelt eine Art Legislative: Der Forscher soll sich vor einer Untersuchung auf spätere Entscheidungen festlegen und nach Sammlung von Ergebnissen nur das tun, was ihm im Spielraum der vorausgehenden Festlegung zu tun verbleibt. Weicht das Ergebnis bei einseitiger Prüfung in der unerwarteten Richtung ab, dann soll ihm lediglich zu schlussfolgern übrig bleiben, die Nullhypothese könne nicht verworfen werden. Haben die experimentellen Maßnahmen denn dann keinen Effekt gezeigt? Darf man einen solchen einfach ignorieren?

Man stelle sich einen verantwortungsvollen Mediziner vor, der mit einem neuen Medikament eine Verminderung der Sterberate erreichen wollte und feststellen muss, dass die Sterberate der Behandelten gegenüber einer Kontrollgruppe höher statt niedriger war. Er wendet so, wie ich dies für richtig halte, den einseitigen Signifikanztest an und erhält  $p = .05$ . Er wird dieser „Signifikanz“ entsprechend konsequenterweise nicht nur das Medikament bei weiteren Behandlungen sofort absetzen, sondern die Kollegen vor möglichen schädlichen Wirkungen warnen. Hätte er mit einer zweiseitigen Prüfung  $p = .10$  erhalten, hätte er die Verminderung der Sterberate eher für nicht erheblich gehalten und nur geschlussfolgert, das Medikament hilft nicht, aber es schadet auch nicht. Er wäre weniger vorsichtig und würde durch Unterlassung möglicherweise das weitere Sterben von Menschen mit verschulden.

Gerichtete statistische Tests, die auf erwartungsgegenteilige Befunde angewendet werden, führen nicht nur manchmal zu lebenserhaltenden Entscheidungen. Sie können auch bei „nur theoretisch“ bedeutsamen Fragestellungen alarmierende Konsequenzen haben. So war ich angesichts der auffällig niedrigen astrologischen Trefferrate meines Probanden, auf die in einem weiteren Test eine relativ hohe Trefferrate folgte, der Meinung, dass sich darin nicht etwa astrologische Deutungsfähigkeit, sondern Psi-Fähigkeit geltend machen könnte, von der man aus anderen Untersuchungen weiß, dass die Wirkungen bidirektional ausfallen können. Hätte ich mir Guiards quasi-juridikale Festlegung zu Eigen gemacht, dass die Unterscheidung zwischen Trefferdefizit und Trefferüberhang „nicht Gegenstand der Untersuchung“ sind (Guiard), dann wäre nichts passiert. Doch sind die unerwartete Erhöhung der Sterberate und das unerwartete unbewusste Ausweichen des Astrologen vor Treffern bei einem Deutetest

---

<sup>9</sup> Prof. Dr. Suitbert Ertel ist emeritierter Psychologe an der Universität Göttingen.  
E-Mail: sertel@gwdg.de

alarmierende Angelegenheiten. Diese sollten durch willkürliche Selbstbeschränkungen nicht in den Topf der unbeachteten Phänomene geworfen werden, der bei Großzügigkeit gegenüber Beta-Fehlern bedenkenlos gefüllt wird und den unsensitiven Forscher weiter ruhig schlafen lässt.

Wenn der Forscher eine einseitige statistische Prüfung vornimmt, hat ihn sein Welt- und Fachwissen dazu veranlasst. Wird er von Effekten überrascht, die in die Gegenrichtung des Erwarteten gehen und lässt er den Befund als einseitig signifikant gelten, was ich empfehle, dann bedeutet signifikant soviel wie: Die Alternativhypothese von der man ausging, hat sich aus Gründen nicht bewährt, die die inhaltliche Basis der Hypothese infrage stellen. Eine ganz andere Alternativhypothese und eine entsprechende Änderung der zugrunde liegenden Voraussetzungen erscheinen notwendig. Man war auf dem falschen Dampfer. Eine Kursänderung ist sehr bedeutsam, egal ob man den Kurs exploratorisch (noch nicht völlig sicher) oder konfirmatorisch (mit größerer Sicherheit) eingeschlagen hatte.

Signifikant bedeutet im Falle erwartungswidriger Beobachtungen nicht: Der Forscher darf sich darin bestärkt fühlen, den Kurs beizubehalten. Die Bestätigung von etwas Erwartetem ist oft *nicht* sehr bedeutsam. Das Überraschende, Nicht-Erwartete ist meist bedeutsamer, wie ja auch Popper die Falsifikation von Hypothesen für bedeutsamer erachtete als ihre Bestätigung. Man sollte einen Befund, der nicht von der Zufallserwartung abweicht und deshalb nicht zur Ablehnung der Nullhypothese führt, nicht gleichsetzen mit einem Befund mit erheblicher Gegenrichtung. Denn kann die Nullhypothese bei einseitiger Prüfung wegen der Zufallsnähe der Beobachtungen nicht abgelehnt werden, ist das nicht alarmierend, der nicht-signifikante Befund ist allenfalls enttäuschend und führt meist nur zu veränderten methodischen Überlegungen auf inhaltlich gleichbleibender Linie.

Guiard meint und lehnt sich da an Kimmel (1957) an, dass man bei einseitigem Testen und richtungswidrigem Ergebnis die Nullhypothese nicht ablehnen könne, auch wenn der Widerspruch des Beobachteten zur Erwartung noch so groß sei. Das halte ich für rein bürokratisch und nicht mit Gründen nachvollziehbar, und Kimmel scheint sich da auch selbst zu widersprechen. Das würde ja bedeuten, dass die Ablehnung einer Nullhypothese, also die Anerkennung der Wahrscheinlichkeit eines faktischen Effekts, von einer von vorne herein richtig liegenden Erwartung des Forschers abhängig sei. Hätte er von vorne herein erwartet, was sich bei seiner falschen Erwartung tatsächlich ereignete, könne er die Nullhypothese ablehnen, sonst nicht.

Das Gegenteil erscheint mir nur sinnvoll: Obgleich der Forscher mit den inhaltlichen Erwägungen, die ihn zur Hypothese führten, und mit zurecht geschneiderten methodischen Mitteln sich um Bewährung seiner Alternativhypothese bemühte, kam das Gegenteil des Erwarteten heraus. Umso mehr erweist sich dann doch das Widerständige der Tatsachen als eher zuverlässig und unverdächtig, was einen möglichen Bias eines Bestätigung wünschenden Forschers betrifft. Vor allem in einem solchen Fall wäre die Faktizität des Beobachteten ernst zu nehmen.

Man könnte sogar als Vorsichtsmaßnahme die Regel einführen, dass bei Beobachtungen, die in der Richtung vorausgehender Erwartungen liegen, grundsätzlich immer eine zweiseitige (schärfere, schwierigere) Prüfung vorzunehmen sei und dass nur dann einseitig (weniger scharf, sensibler) zu prüfen sei, wenn das Gegenteil des Erwarteten beobachtet werden sollte. Die Schärfe der Prüfung wäre ein sinnvolleres Kriterium für die Entscheidung zwischen ein- und zweiseitigem Testen als allein die Erwartung des Effekts durch den Forscher. Ob etwas Gerichtetes erwartet wird, würde so wie bisher vorher festgelegt werden, doch würde die Entscheidung zwischen ein- oder zweiseitigem Test von der Richtung der später erhobenen Befunde abhängig werden. Das wäre eine Popperianische Wende beim Hypothesentesten, die nach Akzeptanz an höherer Stelle (International Biometric Society o. ä.) und bei entsprechender Verbreitung eine heilsame Wirkung weltweit auf die statistisch arbeitende Forschung ausüben würde. Skeptiker wie Randi würden für einen Nachweis von Psi beim Vorliegen richtungspositiver Ergebnisse allemal die Anwendung einer zweiseitigen statistischen Prüfung bevorzugen.

Die Problematik, über die Guiard und ich diskutieren, hat nicht zuletzt mit der Willkür zu tun, die die Festlegung von Signifikanzniveaus hervorgebracht hat. Signifikanz ist kein wahrscheinlichkeitstheoretisch notwendiger Begriff, man könnte auf ihn ganz verzichten, was manche Autoren ja auch tun. Es ist ein behelfsmäßiges und bedenkliches Interface zwischen dem rein statistisch-mathematischen Denken und dem Denken, das bei der Lösung von Sachproblemen gefordert wird und bei dem erst die Begriffe der Wichtigkeit, Bedeutung und Bedeutsamkeit sinnvoll werden.

Diese Überlegungen gehen in die Richtung des Theorems von Bayes, doch lassen sie sich auch im traditionellen Denken nach Fischer oder Neyman & Pearson anwenden. Die Signifikanzprüfung ist ein statistisches Werkzeug, das den Zwecken des Forschers gemäß flexibel einzusetzen ist. Sie erweist sich als starr, wenn sie bürokratisch als Vorschrift aufgefasst wird und zu sinnwidrigen Konsequenzen führt. Als Quelle von Hinweis-Signalen halte ich das Werkzeug für nützlich. Doch hat es dem vernünftigen Denken zu dienen und dieses Denken weder zu bestimmen noch zu ersetzen.