

Fortgesetzte Diskussionen zu früheren Beiträgen

Kommentare zum Aufsatz von Suitbert Ertel

„Außersinnliche Wahrnehmung unter Kontrolle organisierter Skeptiker“

In: *Zeitschrift für Anomalistik* 7 (2007), 236-269

ECKHARD ETZOLD ¹

Preisgeld blockiert Forschung

Mit dem Ballziehtest präsentiert Suitbert Ertel einen originellen Versuch, der in vielfacher Hinsicht von den üblichen parapsychologischen Experimenten abweicht. Ist es sonst üblich, durch Kontrolle und Aufsicht Vorkehrungen gegen Täuschung und Betrug zu treffen, so werden die Teilnehmerinnen und Teilnehmer hier dazu angehalten, die Versuche zuhause ohne Fremdkontrolle durchzuführen, im Vertrauen darauf, dass die „Spielregeln“ eingehalten werden. In einem zweiten Gang werden die besonders „Erfolgreichen“ dann ausgewählt und setzen ihre Versuchsreihen im Göttinger Georg-Elias-Müller Institut unter Aufsicht fort. Das entgegengebrachte Vertrauen, dass die Versuchsteilnehmer den Ballziehtest gemäß den Spezifikationen durchführen, ist somit konstitutiv für diesen Versuch. Gemäß den Befunden zum Sheep-Goats-Effekt, der von Schmeidler (1943) erstmals beschrieben und von anderen Forschern bestätigt wurde (Lawrence 1993), darf bei einem solchen Versuchsdesign auch mit erhöhten Effektstärken gerechnet werden, da schmälernde Einflüsse durch ein skeptisches Umfeld weitgehend ausgeblendet werden. Das scheint auch zunächst der Fall zu sein. Mit Trefferraten, die umgerechnet $Z = 6,4$ und sogar $Z = 19,3$ ergeben, werden Signifikanzen erzielt, die weit jenseits fast jeder Zufallserwartung liegen. Wenn da nicht die Frage wäre, was denn an den Ergebnissen wirklich anomal ist und was konventionell erklärt werden könnte.

Parapsychologische Studien mit solch hohen Z -Werten kamen auch schon früher vor. In Ganzfeld-Versuchen wurden Ergebnisse mit bis zu $Z = 5,20$ erzielt (Bem et al. 2001). In die PK-Meta-Analyse von Bösch et al. (2006) sind 14 Studien eingegangen, deren Ergebnisse Z -Werte zwischen $Z = 4$ und $Z = 10,68$ aufweisen, u.a. Kelly & Kanthamani (1972) mit $Z > 10$,

¹ Eckard Etzold, Dipl. theol., ist Pfarrer der evangelisch-lutherischen Kirchengemeinde Broitzem bei Braunschweig.

desweiteren Honorton und Barksdale (1972) mit $Z = 6,15$ oder Helmut Schmidt mit neun Studien im Bereich zwischen $Z = 4$ und $Z = 5,66$. Bei der Mehrheit dieser Studien wurden die hohen Signifikanzen ebenfalls mit selektierten Versuchspersonen erzielt. Insofern sind die von Suitbert Ertel berichteten hohen Z -Werte zunächst nicht außergewöhnlich, sondern reißen sich ein in die Befunde, die wir bereits aus der Vergangenheit kennen.

Nun hat sich Ertel zusammen mit erfolgreichen Versuchspersonen bei dem Preisausschreiben der GWUP in Würzburg angemeldet. Bevor es dort jedoch zu den eigentlichen Versuchen kam, wurden unter Annahme teilweise phantastischer Täuschungsmöglichkeiten² mehrere Veränderungen des Versuchsdesigns vorgenommen, um die Möglichkeit von Betrug oder die Wirksamkeit konventioneller Phänomene (unterschwellige Wahrnehmung und verdeckte Manipulationen) zu verhindern. Damit entsprach der Versuch jedoch nicht mehr dem originalen Göttinger Ballziehtest. Gerade der Ausschluss direkter Berührung der Bälle (durch Handschuhe) und des Feedbacks dürften nach bisherigen Erfahrungen in der Parapsychologie zu einem drastischen Einbruch der Effekte führen. Aber es sollte in Würzburg ja die Möglichkeit ausgeschlossen werden, wegen eines womöglich hochsignifikanten Ergebnisses, das vielleicht auch konventionell zustande kommen könnte, das Preisgeld zahlen zu müssen.

Wie zu erwarten, fielen die Ergebnisse in Würzburg deutlich geringer aus. Der einzige Erkenntnisfortschritt war, dass der Versuch offenbar unter den – sachlich begründeten – Göttinger Voraussetzungen hohe Effekte erzielt, aber bei gravierenden Veränderungen des Versuchsdesigns offenbar nicht mehr funktioniert. All das hätte man jedoch schon früher wissen können.

Diese Vorgehensweise macht m.E. deutlich, dass ein Wettbewerb mit einem Preisgeld kein geeignetes Mittel der Forschung ist. Vorkehrungen gegen Betrug werden nämlich durch das Preisgeld geradezu „erzwungen“. Die Wettbewerbssituation verfremdet derart die Versuchsbedingungen, dass die erwarteten Effekte gemindert, ja sogar ausgelöscht werden. Sinnvoller und vielversprechender wäre m.E. eine Vorgehensweise in Würzburg gewesen, bei der man sich zunächst strikt an das Originaldesign des Göttinger Ballziehtests gehalten hätte. Dann sähe man, ob sich an einem anderen Ort mit anderen Beobachtern auch dieselben Effekte einstellen. Wenn sich die zuvor beobachteten Effekte unter diesen Bedingungen nicht einstel-

2 Von skeptischer Seite wurde behauptet, der Beutel sei schwach durchsichtig, so dass man die Zahlen auf den Bällen von außen lesen könne. Ich habe das mit einem Beutel von Herrn Ertel getestet, aber selbst unter wechselnden Bedingungen und Lichtverhältnissen war es mir nicht gelungen, die Zahlen auf den Bällen im (Nike)-Beutel zu erkennen. Selbst der Versuch, einen Ball in den Beutel zu legen, ihn loszulassen und bei dem von Herrn Ertel beschriebenen Mischen der Bälle so „am Ball zu bleiben“, dass es unentdeckt bleibt, und man ihn wieder herausziehen kann, stellt sich als derart schwierig dar, dass das m.E. nicht zur Erklärung der gefundenen Effekte dienen kann.

len, wäre zu fragen wieso. Wenn sich aber dieselben Effekte mit derselben hohen Signifikanz einstellen, wäre die weitere angemessene Vorgehensweise gewesen, diese Effekte (konventionell) zu erklären, und zu dieser Erklärung die „erfolgreichen“ Versuchspersonen zu hören. Die Betrugshypothese darf dabei nicht durch eine Veränderung des Versuchsdesigns ausgeschlossen werden, weil damit – wie oben beschrieben – elementare Versuchsvoraussetzungen nicht mehr erfüllt werden, die nichts mit Betrug zu tun haben. Sie muss den tatsächlichen Nachweis von Betrug erbringen, wenn sie bestätigt werden soll.

In einem weiteren Versuch unter Beibehaltung des Göttinger Versuchsdesigns müsste dann eine dem Versuch gegenüber skeptisch eingestellte Testperson versuchen, in identischer Weise wie die Versuchspersonen Suitbert Ertels und unter Nutzung des neuen Wissens, wie es konventionell gehen könnte, dieselben hochsignifikanten Effekte zu produzieren. Gelingt das, so wäre das ein Anlass, eine konventionelle Erklärung für möglich zu halten. Gelingt es nicht, die hochsignifikanten Versuchsergebnisse unter Anwendung einer konventionellen Erklärung zu replizieren, so folgert daraus das Eingeständnis, dass wir es mit einem zur Zeit noch nicht erklärbaren Effekt zu tun haben. Ob es sich dabei um außersinnliche Wahrnehmung handelt, wie der Titel der Studie, „Außersinnliche Wahrnehmung unter Kontrolle organisierter Skeptiker“, nahelegt, muss m.E. offen bleiben.

Unbeantwortet bleibt bisher die Frage: Was zeichnet die „erfolgreichen“ Versuchspersonen gegenüber den anderen aus? Über welche Fähigkeiten verfügen sie, über die andere nicht verfügen? Als eine mögliche alternative „anomale“ Erklärungshypothese könnte auch ein „paranormales“ Markieren der Bälle in Frage kommen (vgl. Radin et al. 2007).

Für den Ausschluss möglicher „konventioneller“ Erklärungen halte ich folgende Testvarianten für sinnvoll, obgleich ich auch weiß, dass es teilweise schwer ist, geeignete Testpersonen und Mittel zu finden. Grundsätzlich sollte dabei der Hautkontakt zwischen Testperson und Ball gegeben sein, so wie man sich ja Heilen durch Handauflegen auch nicht ohne direkte Berührung vorstellen kann:

- Es wäre interessant, solch einen Ballziehtest mit einem oder mehreren Erblinden durchzuführen, bei denen nachgewiesenermaßen eine hohe taktile Sensibilität in den Fingerspitzen vorhanden ist.

- Anstelle von fünf Zahlen könnten die Bälle auch mit fünf farbigen Strichen (rot, orange, gelb, grün, blau) gleicher Form markiert werden. Bei diesem Versuch sollte vorher bekannt sein, ob Versuchspersonen farbenblind sind.

Es würde in der Sache sicher weiter führen, wenn die GWUP und Suitbert Ertel mit seinen Versuchspersonen diese gemeinsamen Studien fortführen könnten, allerdings unter Verzicht auf die Wettbewerbssituation und mit möglichst originalgetreuem Göttinger Versuchsdesign im oben beschriebenen Sinn. Das setzt aber voraus, dass bei allen Beteiligten wirklich der

Wille vorhanden ist, etwas Neues zu lernen.

Literatur

- Bem, D.J.; Palmer, J.; Broughton, R.S. (2001): Updating the Ganzfeld database: A victim of its own success? *Journal of Parapsychology* 65, 207-218.
- Bösch, H.; Boller, E.; Steinkamp, F. (2006): Examining Psychokinesis: The Interaction of Human Intention With Random Number Generators – A Meta-Analysis. *Psychological Bulletin* 132, 497-523.
- Honorton, C.; Barksdale, W. (1972): PK performance with waking suggestions for muscle tension vs. relaxation. *Journal of the American Society for Psychical Research* 66, 208-214.
- Kelly, E. F.; Kanthamani, B. K. (1972): A subject's efforts toward voluntary control. *Journal of Parapsychology* 36, 185-197.
- Lawrence, T. R. (1993): Gathering in the Sheep and Goats. A Meta-Analysis of Forced-Choice Sheep-Goat ESP Studies, 1947-1993. In: Schlitz, M.J. *The Parapsychological Association 36th Annual Convention. Proceedings of Presented Papers*. Toronto, 75-86.
- Radin, D.; Hayssen, G.; Walsh, J. (2007): Effects of Intentionally Enhanced Chocolate on Mood. *Explore: The Journal of Science and Healing* 3, 433-546.
- Schmeidler, G. R. (1943): Predicting good and bad scores in a clairvoyance experiment: A preliminary report. *Journal of the American Society for Psychical Research* 37, 103-110.

ANDREAS HERGOVICH ³

Der Ballzieh-Test oder das Design der zahlreichen Betrugsmöglichkeiten

Suitbert Ertels zentrale Aussagen sind: Bei Ballzieh-Tests mit ausgewählten Versuchspersonen (Vpn) wurden in Heimtests und am Institut in Göttingen hochsignifikante Trefferquoten erzielt. Bis auf ein anfänglich signifikantes Testergebnis in einer Testreihe (Testreihe 7) konnten bei den nachfolgenden Tests der GWUP in Würzburg 2005 und 2006 keine signifikanten Trefferquoten mehr erzielt werden. Ertel beklagt sich im Folgenden vor allem über die seiner Ansicht nach unnötigen Änderungswünsche der GWUP gegenüber seiner Versuchsanordnung: „Eine Begutachtung der Beutel-Bedingung durch unabhängige professionelle Fachmagier wäre sinnvoll. Erwartung: Die Beutelbedingung wird, möglicherweise mit leichten Veränderungen, von ihnen als unbedenklich erklärt werden. Auf eine Box, auf Handschuhe,

³ Prof. Dr. Andreas Hergovich ist Psychologe an der Universität Wien.

auf das Nichtzurücklegen der Bälle könnte man ggf. nach ihrem Urteil in Zukunft verzichten“ (S. 130).

Ich möchte in meinem Kommentar darlegen, wie sehr sich Ertel in diesem Punkt irrt. Wenn man der Beschreibung der Standard-Versuchsanordnung von Ertel folgt, bieten sich viele Trickmöglichkeiten an. Bei allen Methoden wird davon ausgegangen, dass sich die Vp an das Standard-Testverfahren hält, und überdurchschnittliche Trefferraten erzielt.

Methode 1: Der Beutel weist eine kleine Präparation in Form einer kleinen, gerade einen Tischtennisball fassenden Extratasche auf. Die Vp steckt den Ball beim Zurücklegen in die Extratasche. Nach Belieben wird die betreffende Nummer vorhergesagt und der Ball aus der Extratasche gezogen, wenn die Vp das Gefühl hat, durch bloßes Raten nicht mehr weiter zu kommen.

Schlussfolgerung: Der Beutel wird besser ersetzt, weil es unzumutbar ist, vor jeder Testreihe zu kontrollieren, ob der Beutel gegen einen präparierten ausgetauscht wurde oder eine kleine Extratasche (z.B. mit Klettverschluss) beim unpräparierten Beutel angebracht wurde.

Methode 2: Die Bälle sind präpariert, sodass sie unterschiedlich glatte/raue Oberflächen aufweisen, z.B. durch Bearbeitung mit Schleifpapier (für Vpn ohne viel Fingerspitzengefühl) oder mit Fixierspray. Zum Beispiel könnten alle Bälle mit der Nummer 2 und 5 aufgeraut sein. Die Vp sagt 2 oder 5 voraus und greift nach einem rauhen Ball.

Schlussfolgerung: Mitgebrachte Bälle sind nicht akzeptabel. Es kann nicht garantiert werden, dass der Magier oder der Versuchsleiter genügend Fingerspitzengefühl aufweisen, feine Präparationen zu entdecken.

Methode 3: Bälle und Beutel können kontrolliert werden. Die Vp hat nicht unbedingt eine betrügerische Absicht. Die Vp weist klebrige Finger auf und verklebt die Bälle. Intuitiv lässt sie die Finger vom klebrigen Ball, wenn sie eine noch nicht genannte Nummer vorhergesagt hat.

Schlussfolgerung: Die Handschuhbedingung ist vorzuziehen. Die Ratewahrscheinlichkeit wäre entsprechend zu korrigieren.

Methode 4: Bälle und Beutel können kontrolliert werden. Die Vp hat die Augen beim Hineingreifen in den Beutel doch nicht ganz geschlossen (Jeder von uns weiß aus Kindertagen, dass man die Augen nur ganz wenig öffnen muss, um genügend zu sehen, ohne dass das bemerkt wird) oder guckt durch die Augenbinde (sofern eine verwendet wurde?) hindurch und nimmt sich so einfach den Ball, den sie braucht (oft genug wird ja ein Ball mit der gesuchten Nummer ins Blickfeld fallen).

Schlussfolgerung: eine Box-Vorrichtung, wie in Würzburg vorgesehen, ist gegenüber dem Beutel prinzipiell zu bevorzugen.

Methode 5: Bälle und Beutel können kontrolliert werden. Die Vp verwendet eine sogenannte Halbschale (d.h. einen halben Tischtennisball, der genau auf die anderen Bälle passt. Auf der Schale steht eine Nummer). Die Halbschale wird in der Handfläche palmiert (der Handrücken kann gezeigt werden, ebenso die Finger). Die Vp greift in den Beutel, setzt dabei die Halbschale auf einen Ball, holt ihn aus dem Beutel und zeigt ihn dem Versuchsleiter. Vorher wurde natürlich die Nummer auf der Halbschale vorhergesagt.

Schlussfolgerung: Es ist möglichst zu vermeiden, dass die Vp selbst in den Behälter greift. Der Versuchsleiter sollte sich den Ball nicht nur zeigen lassen, sondern ihn auch in die Hand nehmen.

Methode 6: Bälle und Beutel können kontrolliert werden. Die Vp palmiert beim Hineingreifen in den Beutel einen Ball (oder einfacher: lässt diesen im Sakko-Ärmel verschwinden) und produziert sichtbar einen zweiten Ball, der dem Versuchsleiter zur Protokollierung hingelegt oder gezeigt wird. Der herzeigbare Ball wird zwischen Daumen und Zeigefinger gehalten, der geheime Ball wird vom gekrümmten Mittel-, Ring- und kleinem Finger umfasst (sogenannte Fingerpalmage). Während der Versuchsleiter den produzierten Ball notiert, hat die Vp Zeit, die Nummer des palmierten (oder im Ärmel befindlichen) Balles zu registrieren. Befindet sich der Ball im Ärmel, kratzt man sich z.B. am Ohr und kann in den eigenen Ärmel hineinsehen. Die Nummer des geheimen Balles wird als nächstes vorhergesagt. Beim Zurücklegen wird der palmierte Ball zwischen Daumen und Zeigefinger genommen oder man lässt ihn bei gestrecktem Arm einfach aus dem Ärmel in die Hand rollen. Ein neuer Ball wird heimlich palmiert. So ist der Betrüger immer einen Schritt voraus.

Schlussfolgerung: Das Zurücklegen der Bälle ist keine gute Idee. Mehrere Personen sollten den Versuchsablauf kontrollieren.

Methode 7: Bälle und Beutel können kontrolliert werden. Die Vp hat nicht unbedingt eine betrügerische Absicht. Sie spürt einfach, ob ein Ball vor kurzer Zeit berührt wurde oder nicht. In eigenen Versuchen mit Tischtennisbällen habe ich bemerkt, dass diese eine bemerkenswert kühle Oberfläche haben, sich aber, wenn sie nur wenige Sekunden in der Hand gehalten werden, spürbar erwärmen. Wenn man den Ball in den Beutel zurücklegt, den Beutel danach schüttelt und wieder in den Beutel hineingreift, ist es leicht zu sagen, ob der Ball eben gezogen wurde oder nicht. Ich habe den Versuch zu Hause bei Zimmertemperatur mit drei vorhandenen Bällen durchgeführt und konnte zu 100% sagen, welchen Ball ich zuletzt in der Hand hatte. Bei 50 Bällen scheint diese Möglichkeit vielleicht abwegig zu sein, wenn man aber bedenkt, dass beim „Pfannkuchen-Wenden“ die komplette Durchmischung der Bälle sicher

nicht hundertprozentig garantiert ist (das Mischproblem wäre eine eigene Abhandlung wert), zudem Vpn immer die Tendenz haben, die Bälle an der selben Stelle abzulegen und zu ergreifen, besteht die reale Möglichkeit, den gerade gezogenen Ball wieder zu ziehen. Hat die Vp die gleiche Nummer wie beim vorigen Versuch vorhergesagt, wird sie intuitiv auch eher zum wärmeren Ball greifen, wenn sie ihn erwischt. Hat sie aber eine andere Nummer genannt, dann wird sie den wärmeren Ball loslassen und einen unberührten, kühlen wählen. Warum Ertel meint, die „im Alleinversuch sich möglicherweise einschleichenden Tast- oder Wärmeempfindungen hätten unter den Augen von Versuchsleitern weniger Chancen, gespürt und beim Suchen eines Balles im Behälter genutzt zu werden“ (S. 126), bleibt ein Rätsel. Man sieht ja nicht, ob die Vp etwas spürt oder nicht.

Schlussfolgerung: Mit bloßen Händen sollte die Vp nicht in den Beutel greifen. Handschuhe sind vorzuziehen. Die Ratewahrscheinlichkeit wäre zu korrigieren.

Methode 8: Bälle und Beutel können kontrolliert werden. Alle Bälle wurden zuvor mit weißer Leuchtfarbe präpariert. Die Präparierung ist mit bloßen Augen nicht zu sehen. Wird ein Ball ans Licht geholt, leuchtet er nach dem Zurücklegen im dunklen Beutel. Öffnet man die Augen einen kleinen Spalt oder schaut man unter der Augenbinde durch, kann man den Ball leicht identifizieren. Hat man vor, zu tricksen, wird die Nummer des eben zurückgelegten Balles vorhergesagt. Solche Trickmöglichkeiten erscheinen vielleicht „an den Haaren herbeigezogen“ zu sein, tatsächlich gibt es aber einen Händlertrick, der auf diesem Prinzip basiert.

Schlussfolgerung: Die Beutelbedingung ist abzulehnen. Bloße Augenscheinkontrollen reichen bei weitem nicht aus. Es dürfen keine mitgebrachten Bälle verwendet werden.

Dies sind nur einige vorstellbare und m.E. praktikable Trickmöglichkeiten. Professionelle Magier werden sich viele weitere und raffiniertere Möglichkeiten ausdenken können. Vor diesem Hintergrund sehe ich derzeit keinerlei Notwendigkeit, von Psi-Kräften auszugehen und kann auch die Kritik Ertels an den strengen Bestimmungen der GWUP nicht nachvollziehen. Warum hat Ertel, der sich doch jetzt bereits seit ca. 10 Jahren mit den Ballzieh-Tests befasst, tatsächlich nie professionelle Magier zu seiner Versuchsanordnung befragt? Offenbar hat er die Wahrscheinlichkeit, betrogen zu werden, völlig unterschätzt.

Das Hauptargument Ertels gegen die Annahme von Tricks ist offenbar das Ergebnis, wonach unter seiner Kontrolle z.T. noch mehr Treffer erzielt wurden als daheim ohne Kontrolle. Aber was ist von einer Kontrolle wirklich zu halten, bei der der wohlgesonnene Kontrolleur keinerlei Erfahrungen mit Tricktäuschungen hat und zusätzlich die Aufgabe der Protokollierung übernimmt?

Einer Powerpoint-Präsentation im Internet entnehme ich folgendes Zitat Ertels: „Aber gibt es einen Grund, den Studierenden eher *Betrugsfähigkeit* als *Psi-Fähigkeit* zuzutrauen“? (ver-

füßbar unter <http://www.parapsychologie.ac.at/programm/ws200001/mulacz/ertel/> am 18.10.2007). Ja, den gibt es, und das ist m.E. der gesunde Menschenverstand, der einen lehrt: Betrug gibt es auf allen Ebenen in der Wissenschaft, aber Psi-Fähigkeiten sind äußerst rar gesät. Und zum Betrügen braucht man in der Regel auch keine großen Fähigkeiten. Ertel schreibt weiter: „Weshalb sollten die Studierenden mogeln, zumal das Mogeln zeitlich viel aufwendiger ist als eine ehrliche Testdurchführung.“ Nein, ist es nicht. Bei den Heimversuchen fülle ich einfach das Protokollblatt aus, ohne mich mit dem Mischen, Ziehen und Zurücklegen der Bälle abzumühen.

Aus meiner Sicht hat sich die GWUP bei den Testungen der Vpn Ertels wenig vorzuwerfen. Es mutet auch etwas wehleidig an, wenn Ertel sich über den generellen Betrugsverdacht, den aus seiner Sicht die GWUP ihm gegenüber zum Ausdruck brachte, entrüstet. Es geht schlicht darum, routinemäßig Betrug auszuschließen, und wie soll das funktionieren, wenn bei wissenschaftlichen Kollegen genau diese Möglichkeit tabuisiert wird. Dies wurde schon von Price (1955) thematisiert: „when we try to imagine knavery and folly in connection with a particular individual, we encounter a surprising emotional blockage, and the possibility seems unreasonable. And thus we find skeptics searching for every other conceivable sort of explanation – proposing absurd systems of involuntary whispering, or indulging in the metaphysical acrobatics of arguing that ESP cannot occur because it involves a ‚negative hypothesis‘ – while the one explanation that is simplest and most in accord with everyday experience is dismissed as inconceivable“ (Price 1955, S. 155).

Betrug ist eine Seite des Menschen, die leider nicht halt macht vor persönlichem Vertrauen (wie die Statistik der Ehescheidungen beweist). Man stelle sich vor, jemand ist mit einem Angestellten der Gepäckskontrolle am Flughafen befreundet. Würde es Ertel für gerechtfertigt halten, wenn der Angestellte bei all seinen Freunden und Verwandten auf die Kontrolle verzichtet, weil er ihnen persönlich vertraut? Ich vermute, er würde sich darüber empören. Bei den GWUP-Tests ist die Gefahr geringer, aber immerhin geht es um viel Geld und die Antwort auf die spannende Frage, ob es Psi-Kräfte gibt. Niemand kann es m.E. daher der Testleitung verübeln, wenn sie ungeachtet der Person (und selbst dann, wenn sie sich sicher ist, es wird nicht betrogen), strenge Sicherheitsvorkehrungen vorsieht.

Die Klagen Ertels über die unerwartete vertraglich verlangte Aufteilung des potentiellen Preisgeldes verraten schließlich mehr über das professorale Selbstverständnis in Bezug auf die Vereinnahmung der Leistungen von Studierenden für die eigenen Ziele als sie eine Überzeugungskraft als psi-hemmender Faktor aufweisen können. Ertel möge sein Design für wesentlich halten, außenstehende Beobachter werden sich sagen, wenn jemand echte Psi-Begabungen hat, dann sollte diese Person Bälle genauso wie Karten oder ESP-Symbole raten können. Ihr sollte dies unter Box- oder Beutelbedingungen, mit oder ohne Handschuhe selbst unter Anwesenheit von Skeptikern gelingen, sofern diese freundlich und offen sind (aber das

war in Würzburg offenbar der Fall). Würde dann nicht ein Gutteil des Geldes der Person, die tatsächlich über Psi-Fähigkeiten verfügt, gebühren?

Erwähnen sollte man noch die trivialerweise fehlende Objektivität der Heimversuche. Jeder Psychologe/jede Psychologin weiß, dass man Fragebögen den Vpn besser nicht mit nach Hause zum Ausfüllen gibt. Die Durchführungsobjektivität, also die Unabhängigkeit der Befragung von der Vorgabemodalität ist nicht gegeben. Der eine füllt den Fragebogen leicht angeheitert mit Freunden aus, die zweite überlässt die langweilige Tätigkeit ihren Kindern und der dritte beantwortet alle Fragen ganz gewissenhaft. Warum sollte es bei den Balltests zu Hause anders sein? Nicht unterschätzen würde ich auch die menschliche Komponente des Geschehens. Ist es nicht wahrscheinlicher, dass mitfühlende Studierende, die den Professor vielleicht verehren und dessen langjährige und intensive Bemühungen um Psi-Effekte kennen, ihm den Gefallen machen und hohe Trefferraten vorgaukeln als dass echte Fähigkeiten vorliegen. Oder dass böartige Studierende, deren Professor gar so überzeugt von der Existenz von Psi ist, diesem einen Streich spielen, um ihn in der irrigen Überzeugung zu bestärken.

Ich erwarte nicht, dass Ertel meinen Schlussfolgerungen viel abgewinnen kann. Vermutlich wird er für jede potentielle Trickerklärung Argumente parat haben, warum die Vpn so nicht geschwindelt haben können. Vielleicht wird bei zukünftigen Untersuchungen wenigstens erwogen, vorab kritische und professionelle Magier zu kontaktieren, um der zu erwartenden Kritik besser begegnen zu können.

Literatur

Price, G. R. (1955): Science and the supernatural. *Science*, 122, 359-367.

Autorenantwort:

SUITBERT ERTEL⁴

Betrugsverdacht und sensorische Schlupflöcher

Die Kritik von Hergovich am Balltest geht ins Leere. Dass die Testteilnehmer zu gewissenloser Täuschung und zu komplizierter Trickanwendung bereit und fähig seien, ist Spekulation ohne empirische Basis. Kontrolle darüber ist zudem möglich, ohne dass eine psi-hemmende Verkünstelung der Standard-Testbedingungen vorgenommen wird. Von den beiden sensorischen Schlupflöchern, die der Kritiker für denkbar hält, ist nur eins diskutabel. Gezielte Ana-

⁴ Prof. Dr. Suitbert Ertel ist emeritierter Professor für Psychologie an der Universität Göttingen.

lysen vorliegender Daten mit N=238 zeigten, dass auch hier mit Sicherheit kein messbarer Einfluss vorliegt. Die Kommentare von Etzold sprechen für eine Fortsetzung der Balltestprüfungen durch die GWUP unter Verzicht auf das für Forschungszwecke störende Preisgeldangebot. Seinem Vorschlag entsprechend wurde der Test mithilfe psi-begabter Personen mit Verwendung von Handschuhen und mit Farben als Zielobjekte durchgeführt. Ergebnis: Die Trefferzahlen nehmen mit Benutzung von Handschuhen bei einigen Testpersonen erwartungsgemäß stark ab, bei anderen allerdings nehmen sie signifikant zu. Die Benutzung von Farben statt Zahlen führt bei einigen Personen zu einem starken Trefferanstieg.

Zunächst zum Kommentar von Hergovich: Seine Einwände gegen meinen Balltest, die die Position der GWUP unterstützen sollen, lassen sich besser evaluieren, wenn man sie unterteilt (siehe Tabelle 1).

Tabelle 1: Betrugsvarianten beim Balltest nach Hergovich, differenziert nach Labortest mit zwei Betrugsbedingungen, und nach Heimtest. Die M-Nummerierung ist die von Hergovich. 0 = indiskutabel, völlig absurd. 1 = diskutabile Denkbarekeit.

		Labortest		Heimtest
		L1	L2	H
Betrüger →		Testperson	Testperson und -leiter	Testperson
Nr.	Betrogene →	Testleiter und GWUP	GWUP	Testleiter
1 (M1)	Extratasche im Beutel anbringen	0	1	0
2 (M2)	Bestimmte Bälle aufrauhern	0	1	0
3 (M4)	Zahlen an Beutelöffnung erspähern	0	1	0
4 (M5)	Halbschale über Bälle stülpen	0	1	0
5 (M6)	Fingerpalmage praktizieren	0	1	0
6 (M8)	Leuchtfarbe auf Bällen verwenden	0	1	0
7	Beobachtungs- protokoll fälschen	0	0	1

Meine Replik ist ausführlich, weil nicht nur die methodischen Forderungen der GWUP, sondern die Güteeigenschaften des Balltests in seiner Standardform zur Debatte stehen. Zunächst sind die bewusst inszenierten Täuschungen, von denen Hergovich sieben aufführt (sechs von ihm als „Methoden“ durchnummeriert, eine ohne Nummerierung), zu trennen von zwei unbemerkten Schlupflöchern der Sinneswahrnehmung („sensory leakage“), die Hergovich irreführenderweise auch unter „Methoden“ (M) subsumiert.

Betrug

Tabelle 1 unterscheidet L1 von L2. L1 ist Betrug in Labortests, bei denen der Testleiter und die GWUP Betrugsopfer wären – die Testperson würde dann allein betrügen. L2 ist Betrug in Labortests unter Kontrolle, bei denen nur die Beobachter der GWUP Betrugsopfer wären, während der Testleiter zusammen mit der Testperson die GWUP betrügen würden. Auch wird Betrug im Heimtest mit aufgeführt (H), bei dem die Testperson allein betrügt und der Testleiter Betrugsopfer wäre.

Die Einträge „0“ in Tabelle 1 besagen, dass Betrug nicht völlig unmöglich wäre (nicht undenkbar). Doch die Vorstellung, die jeweilige Betrugsvariante sei unter gegebenen realen Bedingungen praktisch umsetzbar, ist in den mit „0“ bezeichneten Fällen absurd. Absurd wäre vergleichsweise z. B. die Vorstellung, dass morgen ein Himmelskörper auf der Erde einschlägt und sie zerstört, obgleich dies denkbar und nicht völlig unmöglich ist. Der Eintrag „1“ bedeutet: Die betreffende Konstruktion eines Betrugs ist weniger absurd und sollte diskutiert werden, auch wenn er immer noch „an den Haaren herbeigezogen“ (Hergovich) erscheint.

Unter der Bedingung L1 habe ich nur 0-Einträge vorgenommen. Sie zeigen an, dass es nach meiner Meinung absurd ist zu denken, die Testperson könne trotz Kontrolle einen Beutel mit Extratasche mitbringen und diese heimlich verwenden, sie könne die Bälle aufrauhern usw. (Näheres dazu im nächsten Abschnitt). Der Testleiter müsste sich schon mit der Testperson zusammen getan haben, um ein Betrugsmanöver einzuüben und die GWUP hinteres Licht zu führen (Fall L2).

Unter der Bedingung L2 – Testleiter und Testperson kooperieren als Betrüger – ist der Betrugsverdacht nicht völlig absurd, Betrug ist nur höchst unwahrscheinlich. Er würde zunächst voraussetzen, dass sich der Versuchsleiter beim Anwerben eines Komplizen über die Gefahr hinwegsetzt, dass dieser seinen Betrugsplan aus ethischen Gründen ablehnt und durch Weitererzählen ein Verfahren der universitären Ethikkommission gegen ihn auslöst. Auch würde der Testleiter, dem es gelingt, mit einem Komplizen vor der GWUP den Balltest vorzuführen, Gefahr laufen, dass er, wenn der geplante Betrug aufgedeckt wird, seinen persönlichen Ruf in der Öffentlichkeit und den der Parapsychologie als Wissenschaft insgesamt ruiniert.

Die „1“-Einträge unter L2 behandeln also *nahezu* absurde Fälle von Betrug. Sie sind aber auch ohne den Kontrollaufwand auszuschalten, mit dem die GWUP bei Einführung meines Balltests ins Würzburger Testprogramm die Standard-Testbedingungen verdorben hat, was im Folgenden ausgeführt wird:

(1) Würde man befürchten müssen, dass der Testleiter im mitgebrachten Beutel eine Extratasche angebracht hat, würde diese von der kontrollierenden GWUP vor dem Test durch Inspektion des Beutelinners leicht entdeckt werden. Auch kann man die Testperson dazu anhalten, den jeweils gezogenen Ball sichtbar über der Öffnung des in der Hand hängenden Beutels und damit auch hörbar in den Beutel hinein plumpsen zu lassen. Letzteres wird meist ohnehin routinemäßig praktiziert.

(2) Würde man befürchten müssen, dass der Testleiter Bälle mit zwei der fünf Zielzahlen aufgeraut hat, dann würde auch dies von der kontrollierenden GWUP vor dem Test durch Inspektion der Bälle entdeckt werden können. Die GWUP könnte auch selbst besorgte und beschriftete Bälle verwenden, was aber für sie aufwändig wäre. Zudem müsste der betrügerische Testleiter der Testperson aufgetragen haben, von diesem Trick sehr sparsam Gebrauch zu machen, da ja dann Trefferüberhänge nur bei höchstens zwei Zahlen erwartet werden können, was leicht auffällt und zu gezielten zusätzlichen Kontrollen Anlass geben würde. Zudem würde man mit einem notwendigerweise gering zu haltenden Trefferüberhang das von der GWUP gesetzte Soll an Treffern nicht erreichen.

(3) Würde man befürchten müssen, dass der Testleiter die Testperson im Erspähen von Zahlen auf Bällen an der Beutelöffnung so trainiert hat, dass dies in einer mehrstündigen Testreihe von den Beobachtern nicht bemerkt wird, kann man die Testperson bitten, den Ball immer nur tief im Innern des Beutels zu greifen und heraus zu ziehen. Darauf werden die Testpersonen in der Standard-Instruktion ohnehin hingewiesen. Man kann sie zusätzlich bitten, beim Ballziehen ihre geschlossenen Augen zur Decke zu richten, was viele Probanden von sich aus tun.

(4) Würde man befürchten müssen, dass die Testperson eine heimlich mitgebrachte und im Beutel verwendete Halbschale, auf der eine bestimmte Zahl steht, über einen im Beutel ergriffenen beliebigen Ball überstülpt, dann würde dieser Trick ausgeschlossen, wenn man die Testperson bittet, den gezogenen Ball dem GWUP-Beobachter zwischen Daumen (unter dem Ball gehalten) und Zeigefinger/Mittelfinger (über dem Ball) zu zeigen – was die meisten Teilnehmer ebenfalls unaufgefordert tun. Denn dann würde der Testleiter früher oder später bemerken, dass die Zahlen der echten und falschen Ballhälfte nicht übereinstimmen. Zudem ist es abwegig zu denken, dass die Testperson zunächst unbemerkt eine solche Halbschale in den Beutel befördern könnte, dass sie dann diese Halbschale, da sie nach jeder Benutzung unter die echten Bälle gemischt und mit ihnen geschüttelt wird, bei späterem Bedarf wieder herausuchen könnte, um das nächste Überstülpen vorzunehmen. Nicht zuletzt würde dies

einen Treffer-Überhang nur bei einer einzigen Zahl ermöglichen mit der Gefahr, dass dies bemerkt wird und zu weiteren Überprüfungen Anlass gibt (wie bei Trick 2).

(5) Würde man befürchten müssen, dass die Testperson darauf trainiert worden ist, aus dem Beutel einen Ball unbemerkt herauszuziehen und in ihrem Ärmel verschwinden zu lassen, dann würde man diesen unwahrscheinlichen schwierigen Trick verhindern, indem man sie bittet, den Test mit ärmelloser Bekleidung oder mit aufgekrempten Ärmeln durchzuführen.

(6) Würde man befürchten müssen, dass die Zahlen auf den Bällen mit einer unsichtbaren Farbe präpariert wurden, so dass sie im Dunkel des Beutels leuchten, und würde man meinen, die Testperson könne das wahrnehmen, wenn sie den geöffneten Beutel etwas hochhebt und in die dunkle Tiefe hinein linst, dann würde man dies spätestens beim zweiten oder dritten Versuch bemerken. Dieser Trick wäre auszuschließen, indem man die Testperson bittet, beim Herausziehen der Bälle die geschlossenen Augen zur Decke zu richten (wie beim Verdacht von Trick 3).

(7) Das Mogeln beim Eintragen gezogener Zahlen durch Testpersonen ist nur unter der H-Bedingung (Heimtest) möglich, da die Testpersonen nur daheim das Versuchsprotokoll selbst führen. Dass sie dabei mogeln, ist indessen von vorne herein unwahrscheinlich. Zum einen: Den Teilnehmern wird vor dem Test angekündigt, dass sie zu einem Nachttest unter Kontrolle gebeten werden, wenn sie im Heimtest signifikante Abweichungen von der Zufallserwartung zeigen. Personen, die etwa aus einem Drang zum Täuschen oder aus Bequemlichkeit mogeln (den Fall erwähnt Hergovich), müssten befürchten, dass sie mit ihrer Fälschung bei der angekündigten späteren Laborkontrolle auffliegen.

Zudem wird routinemäßig eine Prüfung der von den Testpersonen abgelieferten Daten durchgeführt, wovon die Testpersonen nichts wissen: Ein willkürliches Eintragen „gezogener“ Zahlen ins Protokollblatt, ohne dass tatsächlich Bälle gezogen werden, lässt sich in der Regel auch ohne Labortest-Kontrolle ermitteln. Dies aus folgendem Grund: Probanden, die den Zufall simulieren möchten, unterliegen bei ihren Trials durchweg einem Schätzfehler (Bias): Die überwiegende Mehrheit (89% von $N = 238$) unterschätzt mehr oder weniger die Wahrscheinlichkeit, mit der eine Zahl, die bei Trial t_i gezogen wurde, beim Trial t_{i+1} noch einmal gezogen wird. Man neigt zu einem überzufälligen Zahlenwechsel. Vom Zufall erwartet werden 20% Sofortwiederholungen, beobachtet wurden bei $N = 238$ Testpersonen nur 10.9% Sofortwiederholungen ($SD=9.67$) (Wiederholungsvermeidung beim Zahlenraten).

Eine Testperson, die das Protokollblatt mit geratenen und gezogenen Zahlen ausfüllt, ohne Bälle zu ziehen, wird demnach nicht nur für die *geratenen* Zahlen den typischen Bias der Wiederholungsvermeidung zeigen – der ist für *geratene* Zahlen normal. Sie wird den gleichen

Bias auch beim Erfinden „gezogener“ Zahlen zeigen. Dies ist Indikator 1 für Mogelei, das Vermeiden von Sofortwiederholungen auch bei den „gezogenen“ Zahlen.

Hinzu kommt der subtilere Indikator 2, der sich bei der Aufeinanderfolge der *geratenen* Zahlen zeigt: Testpersonen neigen normalerweise dazu, beim Raten einer Zahl auch die gerade *gezogene* Zahl nicht sofort zu nennen. Beispiel: Geraten wird bei t_i die Zahl 1 und gezogen die Zahl 2. Bei t_{i+1} neigen die Probanden dann normalerweise dazu, die Zahlen 3, 4 oder 5 zu raten, nicht die 1 (Ratezahlwiederholung) und nicht die 2 (Ziehzahlwiederholung). Erwartet werden 20% Nennungen zuletzt *gezogener* Zahlen beim Raten, beobachtet wurden nur 12.4% (SD = 6.4). Wenn der Bias der verminderten Ziehzahlwiederholung *fehlt* (Indikator 2) und auch Indikator 1 *vorliegt*, wird der Verdacht erheblich verstärkt, dass die „gezogenen“ Zahlen tatsächlich nicht gezogen, sondern willkürlich hingeschrieben wurden.

Unter den 238 Testpersonen der Göttinger Standardisierungsstichprobe, die den Balltest zuhause durchführten, gibt es nur zwei Testpersonen (Thea und Marie⁵), die beide Indikatoren eines willkürlichen Eintragens *gezogener* Zahlen aufweisen. Nur Thea hat darüber hinaus einen sehr signifikanten Trefferüberhang. Zwar lassen sich auch Treffer während eines ansonsten instruktionsentsprechenden Testdurchgangs – oder nach dessen Abschluss – durch vereinzelte „Verbesserungen“ ermogeln. Diese wären mit einer Datenanalyse nicht zu ermitteln. Doch werden die Heimtest-Ergebnisse in der Regel nur zur Auslese geeigneter Probanden für weitere Balltest-Versuche unter Kontrolle verwendet. Verlässliche Schlussfolgerungen werden nur aus Versuchsergebnissen gezogen, die nicht gemogelt werden können.

Tabelle 2: Zwei sensorische Schlupflöcher (sensory leakage) nach Hergovich, differenziert nach Labor- und Heimtest. Die M-Nummerierung ist die von Hergovich. 1 = diskutabile Denkbarkeit

Nr.		Labortest	Heimtest
1 (M3)	Tastspuren auf den Bällen nach Betastung	1	1
2 (M7)	Wärmereste auf den Bällen nach Betastung	1	1

5 Die Namen wurden verändert.

Sensorische Schlupflöcher

Für den Heimtest als auch für den Test unter Kontrolle sind die beiden von Hergovich aufgeführten Möglichkeiten *sensorischer Schlupflöcher* zu bedenken (siehe Tabelle 2).

Zu hypothetischen *Tastspuren* konstruiert Hergovich den Fall, dass eine Testperson unsaubere Finger hat und beim Anfassen der Bälle klebrige Substanz an diese abgibt. Die Testperson würde zu Testbeginn z. B. eine „1“ ziehen und den betreffenden Ball in den Beutel zurücklegen. Wenn sie anschließend z. B. eine „2“ rät und dann den Ball mit der klebrigen Substanz tastet, dann weiß sie mehr oder weniger unbewusst, so meint der Kritiker, dass auf diesem Ball nicht die „2“ geschrieben steht, so dass sie „intuitiv die Finger vom klebrigen Ball lässt“. Angenommen es sei möglich, mit klebrigen Händen das Vorhandensein von Klebe an den Bällen vom Nichtvorhandensein zu unterscheiden, dann würde sich die Trefferwahrscheinlichkeit beim zweiten Zug, bei dem also unbewusst nur 49 statt 50 Bälle zur Auswahl zugelassen werden, von .200 auf .204 erhöhen. Bei den anschließenden Zügen wiederholen sich zwar nach dieser Konstruktion noch ähnliche Minimalgewinne an Trefferwahrscheinlichkeit, so lange bis die Testperson noch nicht alle Zahlen wenigstens einmal gezogen hat. Doch selbst wenn die Erwartung bei jedem Zug des ganzen Run mit 60 Trials konstant .204 bliebe, würde dabei noch nicht einmal *ein* Treffer über dem Zufallsdurchschnitt herauspringen (genau: 12.24 statt 12 Treffer). Mit dem Vorkommen von schmutzigen oder schwitzigen Fingern kann man also nicht begründen, dass die Testperson Handschuhe tragen müsse.

Lässt sich die Forderung des Handschuhtragens begründen, weil sich Testpersonen unbewusst anhand von *Wärmespuren* an zurückgelegten Bällen orientieren können? Hergovich meint: „*Hat die Vp die gleiche Nummer wie beim vorherigen Versuch [Trial] vorhergesagt, wird sie intuitiv auch eher zum wärmeren Ball greifen, wenn sie einen erwischt.*“ An das Schlupfloch der Wärmeorientierung habe ich bei der Entwicklung des Tests von vorne herein gedacht. Ich habe geprüft, ob Information hindurch sickert. Wenn eine Prüfung ergeben sollte, dass das „Loch“ dafür zu klein ist, braucht man es nicht zu stopfen.

Eine ausführliche Darstellung der Untersuchungen zu dieser Frage findet sich in der methodischen Darstellung des Balltests in Ertel (2005). Zahlreiche Befunde liefern dem Verdacht, dass Treffer durch „sensory leakage“ (Wärmeempfindungen, Tastempfindungen, auch von Schriftspuren) oder durch „memory leakage“ (unbewusstes Sich-erinnern an Positionen zurück gelegter Bälle) begünstigt würden, keinerlei empirische Basis.

Ergebnisse mit gerade zurück gelegten Bällen (just drawn numbers, JDNs): Wenn die bei Trial t_i gerade zurück gelegten Bälle beim anschließenden Trial t_{i+1} leicht wieder gefunden werden, dann hat man zu erwarten, dass Testpersonen beim Raten der gerade gezogenen Zahl (JDN) tendenziell mehr Treffer haben als beim Raten früher gezogener Zahlen. Dies aber ist nicht der Fall, die Trefferpunktzahlen beim Raten der JDNs unterscheiden sich nur minimal

und nicht signifikant (bei $N = 238$) von denen beim Raten von nicht-JDNs. Überdies meiden Testpersonen generell das Raten von JDNs (s. o.). Das wäre für sie von Nachteil, wenn beim Raten von JDNs mehr Treffer herauspringen würden. Das Raten von JDNs nimmt im Laufe eines Runs von t_1 bis t_{60} auch nicht zu, was man erwarten müsste, weil sich durch größere Trefferwahrscheinlichkeit der JDNs ein Reinforcement einstellen sollte.

Zeitreihen-Ergebnisse über einzelne Runs und Run-Serien: Wenn Informationen durch sensorische Schlupflöcher fließen sollten, z. B. durch Wärmereste, Schweißspuren oder Schrifterhebungen, dann werden sich diese im Laufe eines Runs mit 60 Trials, die zu einem Testlauf gehören, allmählich stärker auswirken. Die Testpersonen würden zunehmend sensibler tasten und dabei lernen, Empfindungen genauer zu diskriminieren. Vor allem müssten sie beim ersten Trial eines Runs, vor aller Tasterfahrung mit den Bällen und Zahlen, geringere Trefferzahlen erreichen als später. Diese Hypothesen haben sich nicht bestätigt. Die ersten zehn Trefferzahlen eines Runs mit 60 Trials, summiert über die $N = 238$ Probanden, sind von Trial-Nr. 1 bis 10: 333, 299, 332, 340, 306, 311, 320, 335, 311, 336, die letzten zehn, von Nr. 51 bis 60, betragen 333, 306, 340, 316, 335, 333, 300, 333, 337 – kein Lernanstieg liegt hier vor.

Man könnte hier einwenden, dass der Effekt der sensorischen Schlupflöcher nicht bei der Gesamtstichprobe ($N = 238$) durchschlagen muss, man könne ihn ggf. nur bei sensorisch sensibleren Personen, nur bei den „Tastbegabten“ finden. Überaus wichtig ist deshalb der Befund, dass ein Anstieg der Treffer von t_1 nach t_{60} auch dann nicht zu finden ist, wenn man die Berechnung nur für die Testpersonen durchführt, deren Trefferzahlen über dem Median der Gesamtstichprobe liegen ($N = 119$).

Zur Verkünstelung der Standard-Testbedingungen

Eckhard Etzold teilt erfreulicherweise meine Sorge, dass durch Verkünstelung der Versuchsanordnung, die mit den methodischen Änderungen der GWUP entstand, die für Psi-Effekte günstigen Bedingungen des Balltests in seiner Standardform verloren gehen. So solle eine Testperson z. B. weiterhin mit den Bällen „Hautkontakt“ behalten, das Verwenden von Handschuhen werde die Trefferzahl vermindern. Dies habe ich mit sechs „High hitters“ geprüft: Unter Kontrolle absolvierten sie den Balltest einmal ohne, einmal mit Handschuhen, in dieser Folge an verschiedenen Tagen mit zeitlichen Abständen, die beträchtlich waren (für Testpersonen 1-3) oder nur wenige Tage betragen (Testpersonen 4-6). Die Testlänge betrug mindestens 8 Runs (480 Trials). Tabelle 3a enthält die Ergebnisse der Tests ohne Handschuhe, Tabelle 3b die Ergebnisse der Tests mit Handschuhen.

Tabelle 3a: Ergebnisse der Balltests unter der Standard-Bedingung „ohne Handschuhe“.

	1	2	3	4	5	6
	Galina	Tanya	Vanya	Christine	Julia	Veronika
1 N Trials	960	960	960	480	480	540
2 Prob. erwartet	.10	.10	.10	.20	.20	.20
3 Treffer erwartet	96	96	96	96	96	108
4 N Treffer	184	133	139	117	141	183
5 Z	9.41	3.93	4.57	2.34	5.08	8.01
6 ES ₁	.30	.13	.15	.11	.23	.34

Tabelle 3b: Ergebnisse der Balltests unter der Standard-Bedingung „mit Handschuhen“.

	1	2	3	4	5	6
	Galina	Tanya	Vanya	Christine	Julia	Veronika
1 N Trials	480	480	480	480	480	540
2 Prob. Erwartet	.10	.10	.10	.20	.20	.20
3 Treffer erwartet	48	48	48	96	96	108
4 N Treffer	114	101	108	158	127	120
5 Z	9.97	7.99	9.05	7.02	3.48	1.24
6 ES ₂	2.02	3.49	3.81	2.93	-1.01	-4.27

Tabelle 3c: Differenzen der Balltest-Ergebnisse unter der Bedingung „mit Handschuhe“ (ES₁) im Vergleich zur Standard-Bedingung „ohne Handschuhe“ (ES₂).

	1	2	3	4	5	6
	Galina	Tanya	Vanya	Christine	Julia	Veronika
1 ES ₂ -ES ₁	.16	.23	.26	.21	-.07	-.29
2 Z (der Differenz)	2.02	3.49	3.81	2.93	-1.01	-4.27
3 P	.02	.0002	.0001	.003	n.s.	.00001

Tabelle 3c zeigt die Differenzen der Effektstärken, ES mit Handschuhen minus ES ohne Handschuhe, sowie die Z-Werte der Proportionsdifferenzen und deren Signifikanzen.⁶

Die Ergebnisse fallen überraschend aus, denn Etzolds Erwartung, dass sich die Trefferzahlen unter der Handschuhbedingung signifikant vermindern, trifft nur für Veronika zu (Probandin Nr. 6).⁷ Eine Verminderung der Treffer (nicht signifikant) kommt sonst nur noch bei Julia vor. Bei den übrigen Testpersonen steigt das Trefferniveau sogar signifikant oder sehr signifikant an. Warum die Handschuhbedingung eine derart starke Wirkung auf die Trefferhäufigkeit haben kann und diese bei verschiedenen Personen mal zur Verminderung, mal zur Steigerung der Treffer führt, darüber kann man derzeit nur spekulieren.

Zum Preisgeld und Vorschläge für weiteres Forschen

Etzold beanstandet die Preisgeld-Ausschreibung, da sich die Herstellung verkünstelter Versuchsbedingungen durch die Preisgeld-Verwalter, die ihr Geld nicht verlieren wollen, auf Psi-Effekte effekthemmend auswirke. Seinen Vorschlag, gerichtet an die GWUP, die von mir unter natürlichen Standardbedingungen vorgeführten Effekte durch mogelnde Skeptiker herstellen zu lassen und an ihrem Erfolg oder Misserfolg die Validität meiner Befunde mit ausgewählten Testpersonen zu bemessen, halte ich für einfallsreich und sehr bedenkenswert. Allerdings würde der Erfolg eines Magiers beim Balltest, der durch einen Trick zustande kommt, den die Beobachter nicht durchschauen⁸, immer noch die Frage offen lassen, ob naive Testpersonen, die in solchen raffinierten Künsten nicht bewandert sind, ihren eigenen Erfolg ebenfalls durch unerkannte Tricks erzielt haben können.

Ein Balltest mit Blinden, den Etzold vorschlägt, wäre wegen der Abwesenheit visueller Cues interessant. Er ist noch nicht durchgeführt worden. Der Test wurde aber mit 30 naiven

6 In dieser Tabelle sind nur Effektstärken ($ES = Z/\sqrt{N}$) und keine Trefferproportionen aufgeführt, da sich die unterschiedlichen Testbedingungen bei den Probanden 1-3 (Trefferwahrscheinlichkeit .10) und 4-6 (Trefferwahrscheinlichkeit .20) für einen Vergleich nicht eignen. Die Trefferwahrscheinlichkeit .10 ergibt sich, weil die Probanden außer den fünf Zahlen gleichzeitig zwei Farben auf den Bällen zu raten hatten.

7 Veronika hatte am Test der GWUP teilgenommen, ihre Treffer waren dort bis aufs Zufallsniveau gesunken. Dies kann durch eine Psi-Hemmung bedingt sein, die u. a. durch den Handschuh-Zwang bedingt wurde.

8 Ein trainierter Magier wird naive Zuschauer, die auf ihn keine Kontrolle ausüben, mit dem einen oder anderen von Hergovich beschriebenen Trick, etwa in einer Show, mit einem Ballzieh-Test erfolgreich täuschen können. Ein Versuchsleiter, der ihm unmittelbar gegenüber sitzt und über Stunden Hunderte von Wiederholungen der Ballzüge beobachtet, wird auch auf die raffinierteren Tricks à la Hergovich kaum hereinfliegen.

Kindern im Kindergartenalter unter Kontrolle durchgeführt – mit kindertümlichem Material und nur mit 100 Trials pro Kind, denn kleine Kinder halten nicht so lange durch (Hricikova 2003). Der Vorteil von Tests mit Kindern dieser Altersstufe ist u.a., dass es unter ihnen keine Trickkünstler geben kann. Ihr Trefferüberhang hatte eine mehr als doppelt so hohe Effektstärke ($ES = Z\sqrt{N} = .19$) als die von $N = 238$ Studierenden ($ES = .07$), die den Balltest zuhause durchführten.

Etzold schlägt auch einen Ratetest mit Farben vor. Ein solcher wurde mit vier psi-begabten Testpersonen unter Kontrolle durchgeführt. Rote, grüne, gelbe, blaue bzw. schwarze Punkte waren auf weiße Pingpongbälle gemalt worden. Farben statt Zahlen waren zu raten. Die Trefferprozent der Farben liegen bei zwei der vier Personen sehr signifikant höher als mit Zahlen (siehe Abbildung 2). Bei den beiden anderen Teilnehmern zeigten sich keine Unterschiede. Bei weiteren Balltests sollte man demnach ggf. Farben als Zielobjekte statt Zahlen verwenden. Das geeignetere Zielobjekt wäre für die Probanden individuell vorher zu ermitteln.

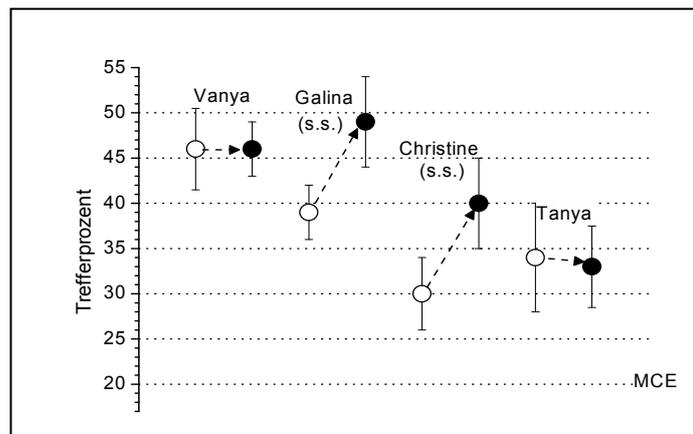


Abbildung 2: Trefferprozent von vier Testpersonen unter der Bedingung „Zahlen raten“ (offene Punkte) und „Farben raten“ (volle Punkte). Konfidenzgrenzen: .95. Signifikanzangaben zu den Unterschieden.

Literatur

- Ertel, S. (2005): The ball drawing test: Psi from untrodden ground. In: Thalbourne, M.A.; Storm, L.: Parapsychology in the Twenty-first Century. Essays on the future of Parapsychological Research. Jefferson N.C.: McFarland.