Australian Journal of Parapsychology
2005, Volume 5, Number 2, pp. 149-164

Technical Paper No. 10

Psi Test Feats Achieved Alone at Home: Do they Disappear under Lab Control?[1]

SUITBERT ERTEL

ABSTRACT: Extraordinary hit-rates from multiple-choice tests, obtained by participants alone in their homes, are ambiguous. On the one hand, their feats might in fact reflect psi power manifesting itself better under informal home than under formal lab conditions. Yet hit surpluses obtained without lab control might also be due to negligent or fraudulent conduct. One way out of this dilemma is to let participants run psi tests at home and to invite high scorers thereafter to do additional runs under lab control. This strategy has been endorsed using *N* = 238 (Sample 1) and *N* = 47 (Sample 2) of student participants. Sample 1 took the ball-selection test (version I).[2] Fifty numbered table tennis balls (10 of each, numbered 1 to 5) are drawn from an opaque bag.

1 Based on a poster presentation at the 47th Convention of the Para-psychological Association, August 5-8, 2004 Vienna, Austria.

2 *Editor's Note:* Previously, the author used the term "ball drawing test" to describe his experimental paradigm (see Ertel, 2005). He now defers to the less ambiguous term "ball-selection test".

Participants guess and draw out numbered balls blind, recording the data as they go (P_{MCE} = 0.20). Participants put drawn balls back into the bag and they shake the bag prior to each trial. One test unit consisted of six or eight runs comprising 60 trials each (total: 360 or 480 trials). Sixteen high scoring participants of Sample 1 were also tested under lab control, again using ball test procedure (version I). Sample 2 took the ball-selection test (version II). This test resembles ball test I in almost every respect except that green or red dots are sprinkled over the balls, and participants guess numbers (five targets) *and* colours (two targets; *MCE* = 10%). Thirteen high-scorers of Sample 2 were also tested under lab control using the bead-selection test where each participant draws one of five colours (no numbers, *MCE* = 0.20). It was hypothesised that (i) hit-rates of high scorers in home tests decline (due to less psi-conducive conditions under lab control and regression towards the mean), and (ii) hit-rates of high scorers under lab control are still significantly above chance (due to genuine psi as was effective at home). Both hypotheses were confirmed with Sample 1 and replicated with Sample 2. Three participants obtained significantly *higher* hit-rates under lab control compared with their home performance. The issue of fraud and bias loses relevance in view of such findings. It is recommended that the *first-home-then-lab-test* strategy be introduced into parapsychological research on a broader scale.

INTRODUCTION

Out of necessity, not of choice (due to lack of funds and assistants), I have been conducting experimental research into the paranormal ignoring, with qualms, the community's methodological 'Demand #1': Fraud must be ruled out in the first place (Milton, 1996). I invited students to conduct psi tests at home. The students, participating in this project out of curiosity and/or out of desire to obtain obligatory credit points, were instructed in class, took the material home, ran the required 360 to 480 trials within 8 days (total testing time 1 ½ hours), returned their record sheets, and waited for feedback from the experimenter, who analysed the data individually and informed each participant in detail, by phone or email, about his or her results. Five annual cohorts of student beginners (total *N* = 238, 84% female) were tested using the ball-selection test (Version I): From an opaque bag, participants drew numbers 1 to 5 written on table-tennis balls (each ball one number) after having guessed at each trial which number they would draw next. Another ball test, somewhat more complex (Version II) was used for Cohort 6 (*N* = 47, 73% female)—each target ball carried numbers 1 to 5 plus green or red dots, the participants guessed numbers and colours. The last student Cohort, #7 (*N* = 48, 86% female), had beads to draw from a box, five different colours (no numbers) serving as targets. The median age of the total is 23 years (age differences between cohorts are negligible).

The ball-selection test yielded surprisingly positive hit deviations from chance (see Table 1 and Figure 1, open circles).

Table 1.
Ball-Selection Test Results of Home tests by Student Beginners across Seven Cohorts.

01	Cohort No.	1	2	3	4	5	6	7	
02	Cohort Year	1998	1999	2000	2001	2002	2003	2004	
03	Test	Ball I	Ball I	Ball I	Ball I	Ball I	Ball II	Bead	
04	Targets	Numbers	Numbers	Numbers	Numbers	Numbers	Numbers & Colours	Colours	
05	*N*	57	38	56	36	44	47	48	
06	Trials	13,680	9,120	20,160	12,960	15,840	22,560	70,121	
07	Hits	3,021	2,011	4,316	2,759	3,539	2,620	14,400	–
08	% expected	20.00	20.00	20.00	20.00	20.00	10.00	20.00	–
09	% observed	22.08	22.05	21.41	21.29	22.34	11.61	20.54	–
10	% surplus	10.40	10.25	7.05	6.45	11.70	16.10	2.70	–
11	$ES = Z/\sqrt{n}$	0.052	0.051	0.035	0.033	0.059	0.054	0.013	0.029
12	$Z_{binomial}$	6.1	4.9	5.0	3.7	7.4	8.1	3.5	7.6
13	*p*	10^{-8}	10^{-6}	10^{-6}	10^{-4}	10^{-13}	$<10^{-15}$	0.0002	10^{-13}

Notes: % surplus = percentage of hits above expectancy (e.g., 22 hits, 20 expected = 10% surplus);
Expectancy rate is different in Column 6;
Column 7 (left): Hit average; Column 7 (right): Hit average with differential colour effect combined

Hit-rates declined with cohorts 2000-2001, but Cohort 2002 replicated the hitting rates of Cohorts 1998 and 1999. Each cohort's hit surplus was highly significant. The results of ball test II exceeded those of ball test I. An analysis of the bead-selection data is more complex (for details, see Appendix). The result may be condensed by saying that deviations from chance were as large as those of the best sample of ball test I participants (i.e., of Cohort 2002).

Yet, all such results might be doubted because, in principle, the observed deviations from chance might be effected by negligent participants who might gain access to some sensory information, or might deceive themselves by inadvertently recording wishful hits while they had actually misses. Unlike ordinary psi test data, where bias and fraud are meticulously ruled out, the present database is not reliable. However, unlike aseptically clean data, where bias and fraud *cannot* have occurred, the present data are most suitable for investigating and possibly finding out, post hoc, whether or not bias *has actually* occurred, either by sensory leakage, or memory help, or fraud, etc.

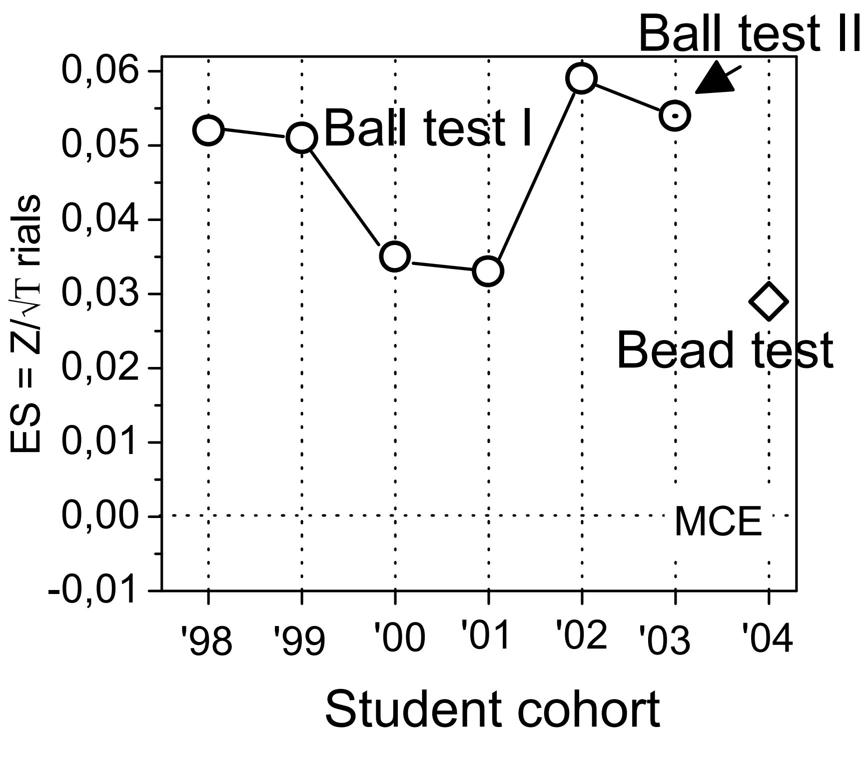

Figure 1.
Effect size of hits for seven cohorts of student beginners, cohorts 1998-2002 took the ball test I; cohort 2003 the ball test II; and 2004 took the bead test.

The issue of sensory leakage may be solved, above all, by checking for signs of ordinary learning. Sensory leakage cannot take effect on first trials—tactile or temperature information from balls and bead must be perceived, stored and applied to be of subsequent help in hitting. The same applies to wishful errors during record-taking that might eventually increase due to gradually fading attention. But the students' data, those of high and low scorers alike, show that hit-rates already exceed chance expectancy with first trials—there is no indication of increasing performance over time (Ertel, 2005).

The issue of fraudulent misconduct may be addressed with blunter or sharper means. A blunter test, nevertheless indispensable in professional criminal practice, as is well known, is to unravel motives. Student participants generally want to *know* whether they score well, they cannot satisfy such curiosity by manipulating their records. Moreover, they have been informed by instruction that high scorers at home would be invited to subsequent lab tests under lab control. If a student under home conditions were tempted, say, by deviant motives, to deceive the experimenter, he or she would have to consider that deception might eventually be discovered, a risky undertaking for fresh participants of academic courses.[3]

A keener way of addressing the fraud issue is to actually invite good home-scorers to perform additional runs in the lab. The present paper gives an account of this endeavour. Two questions need to be answered: (1) Do high scorers under home test-conditions produce random hit-rates under lab control? If they do, hit surpluses obtained at home may be due to bias, fraud or other such "normal" means, and the psi hypotheses would be unsupported; and (2) If hit-rates under lab control are lower than without lab control, but *not* random, two additional questions arise: (2.1) Are the remaining hit surpluses still strong enough to allow for assuming paranormal factors?; and (2.2) If they are strong enough, why do hit-rates under lab control decline?

---

[3] The idea of "reducing experimenter control in a study of special subjects" is not new. Bierman and Gerding (1992) advocate more relaxed ("sloppy") testing conditions in case "special" subjects' control might be "reduced to theoretically relevant" factors. Gerding and Bierman (1997) argue that "strong emphasis on this issue [control over all factors] has hindered progress in parapsychological research. . . . One sometimes gets the impression that the demand for control is not determined by legitimate concerns about methodological rigor, but instead by the conservative attitude of so-called skeptics" (p. 2).

METHOD

Participants and Experimenters

The ball-selection test I (numbers only) was taken by 238 students (84% female), and ball-selection test II (numbers and colours) by 47 students (73% female). Eleven high-scoring participants from ball test I took the *same test* under lab control conditions. Also included in the first lab control sample were 5 non-students, two boys (8 and 12 years of age), and three adults who happened to become engaged as participants with high hit-rates at home.

Of the ball test II sample (students only), 13 high-scorers took the *bead test* under lab control (second lab control sample). High scorers among students who took the bead test at home (N = 49, the most recent cohort) have not (yet) been selected for testing under lab control.

The author acted as experimenter for 15 of 16 participants of the first lab control sample and for 8 of 13 participants of the second lab control sample. Six lab control experiments were conducted with student assistants serving as experimenters.[4]

Materials

The following sets of apparatus were used: (1) Ball-selection test I (numbers only): An opaque bag (originally for sports shoes, 30 cm x 40 cm) containing 50 table tennis balls. One of five digits (1, 2, 3, 4, or 5) is written on each ball, ten times, evenly spread on its surface. There are 10 balls for each number (Total: 50 balls). Record-forms for home testing are provided where participants have to put down their guessed and drawn numbers, for each run of 60 trials. (2) Ball-selection test II (numbers and colours): This test differs from the foregoing test only in that each ball is sprinkled with either red or green dots. Red and green colours are used equally often. The record forms provide blank space to be filled in with guessed and drawn numbers and colours. (3) Bead-selection test: A box containing 1500 little handicraft beads (diameter: 4 mm) whose colours are red, green, blue, yellow or white, 300 beads for each colour. Record-forms provide blanks to be filled in with frequencies of blind draws that were necessary to pick a desired colour (see Procedure).

---

[4] Three student experimenters had been tested themselves under home and lab conditions earlier (control by SE). They tested two students each.

Procedure

The ball-selection test: Each trial starts off with first putting down, on the record sheet, a guessed number (or number plus colour, respectively) that the participant thinks he would draw next. The participant then shakes the bag, closes their eyes, reaches into the bag, picks and draws a ball, checks the actually drawn number (and colour) and writes them down. The participant puts the ball back into the bag. Each run requires 60 trials, one ordinary test series requires six runs. In the ball-selection tests, the number of trials is fixed while the number of hits is open to variation as is common with multiple-choice procedures.

The bead-selection test: In this test, the number of hits is fixed while the number of trials is open to variation. The record form is made up of rows with 10 blanks each. The participant fills them in with frequencies of blind draws until he or she picks a bead with the desired target colour. Prior to starting off with one row of blanks, the participant decides the colour, which remains the target for that row. Supposing the participant chooses “red” and picks a red bead at his or her first trial. The participant would fill in “1”, if another pick of a red bead requires, say, 10 draws, the participant would fill in “10” in the second blank, etc.

Each draw of a bead is preceded by stirring the beads in the box using one’s out-stretched fingers. The beads are drawn with eyes closed. Drawn beads are put back into the box. One test sub-series consists of 10 rows, thus 100 hits are required for one sub-series. Three sub-series (i.e., 300 hits) are needed for one complete test unit. Mean chance expectation (*MCE*) for 1,500 trials (draws) is 300 hits. The participants are told to use as targets all five colours (row-wise) equally often.

Lab tests under control: The participant takes either test while sitting at a table opposite the experimenter. The experimenter fills in the blanks with the participant’s guesses and draws and observes the participant’s behaviour. Attempts to peep into the box or other such fraudulent actions would hardly go unnoticed, and apparently such attempts have not occurred, though it is not claimed here that a trained magician could not succeed in deceiving the experimenter once in a while. The magician’s trick most likely deployed would be to keep a just-drawn ball craftily in the hand without putting it back, and to call that ball’s number on the next trial. But the number of hits for just drawn numbers (which might have been kept in the hand) is not larger than the number of hits for not-just-drawn numbers (see Ertel, 2005). The ball test material has been examined by German skeptics (GWUP), of whom I asked whether they could tell me how participants might obtain hits by deception. They were suspicious of the results, but they lacked an explanation, despite my insistence that they tell me by which tricks a person could obtain hits.

Data analysis

The data were analysed using the Binomial test. Mean chance expectancies (*MCE*) are as follows: $P = 0.2$ for ball test I; $P = 0.1$ for ball test II; $P = 0.1$ and $P = 0.2$ for the bead test (see Table 1, Row 8).

Each individual Binomial Z is transformed into effect size $ES = Z/\sqrt{n}$, where n = number of trials. Significance across individuals is tested using a chi-square formula, where $\chi^2 = \Sigma Z^2$, $df = k$ (where k = number of Z-values—in the present case, $k = N$ for one-sided tests, $df = 2k$ for two-sided tests).

RESULTS

Figure 2 shows the results of Sample 1 participants (those participants who used ball test I under home and lab conditions). Effect size measures of hit-rates obtained under home (left) and lab control condition (right) are plotted—Home: $ES = 0.369$ ($SD = .126$); Lab Control: $ES = .122$ ($SD = .207$). The difference is significant, $t(15) = -3.02$, $p = .004$ (one-tailed). Observed effect size of the total (lower dot, left) is plotted for comparison. The difference between the effect size of selected high scoring participants ($N = 16$, full circle upper left) and that of the total ($N = 238$, full circle lower left) is necessarily large. More importantly, the effect size under lab control (right) is remarkably lower than the effect size obtained under home conditions (left). However, even though individual participants obtained hit-rates under lab control at chance level only ($N = 6$, = 38% of 16), hit-rates of the high-scorers' total remained highly significant under that condition, $\chi^2(16, N = 16) = 721.7$, $p < 10^{-15}$.

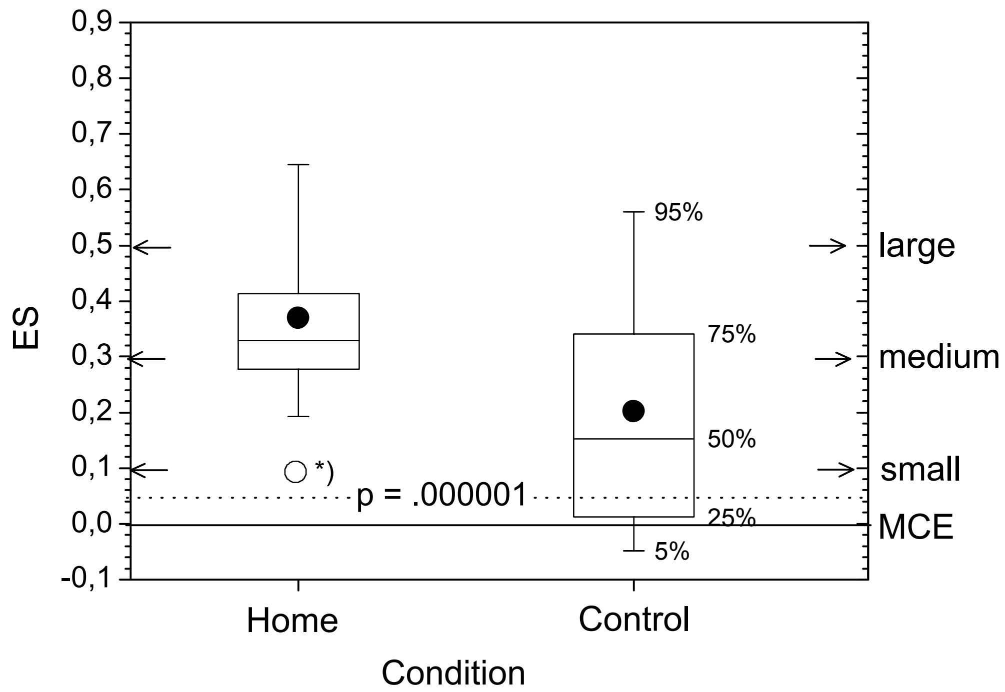

Figure 2.
Results for Sample 1 (N = 16). Effect size ($ES = Z/\sqrt{n}$) from ball test I obtained at home (left) and under lab control in the lab (right). The lower left circle marked *) represents the observed effect size of the total from which Sample 1 had been selected (19,442 hits in 85,200 trials of 238 participants; $P_{exp.}$ = 0.20). The dotted p-line above the *MCE* line is based on the total.

Figure 3 shows results of Sample 2 participants (N = 13)—that is, of those who used ball test II at home (left) and the bead test under lab conditions (right). Effect size measures of hit-rates obtained under home (left) and lab control condition (right) are plotted—Home: ES = .275 (SD = .162); Lab Control: ES = .098 (SD = .153), the difference between the two being significant, $t(12) = -3.00$, p = .006 (one-tailed).

The mean hit-rate of high-scorers of Sample 2 is less pronounced than that of Sample 1 (see Figure 2) because the selection criteria for Sample 2 were less strict. More important, the average effect size for hit-rates of Sample 2 participants decreased under lab control, as was the case with Sample 1 participants. The home-lab difference between hit proportions is highly significant (Z = 6.36, $p = 10^{-10}$). Under lab control, 4 of 13 participants obtained hit-rates at chance level (31% of 13). Nevertheless, total hit-rates of high scorers remained highly significant under that condition, χ^2 (13, N = 13) = 223.9, p = <10^{-15}.

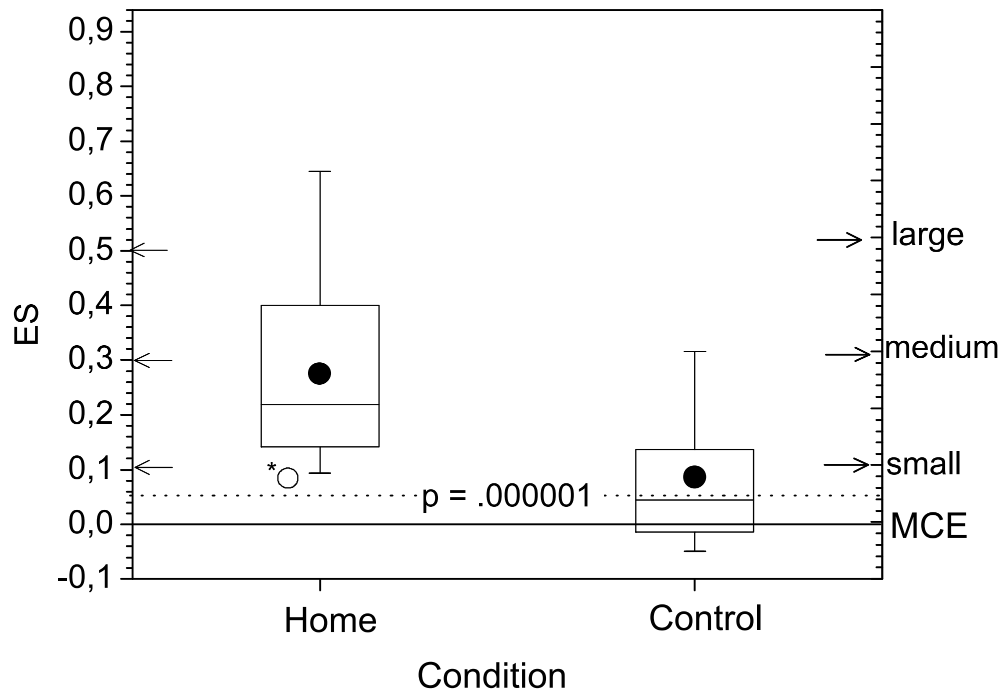

Figure 3.
Results for Sample 2 (N = 13). Effect size ($ES = Z/\sqrt{n}$) from ball test II obtained at home (left) and from the bead test under lab control (right). The lower left full circle marked * represents the observed effect size of the total from which Sample 2 had been selected (2,620 hits in 22,560 trials of 47 Ss; $P_{exp.}$ = 0.10). The dotted p-line above the *MCE*-line is based on the total.

The data above have been analysed while treating individual scores as units; the box plots of Figures 1 and 2 are based on those units. Alternatively, they may be accounted for by summing hits and trials over participants. This procedure has been applied in Table 1 and is taken up again in Table 2, which summarises results for the two lab control samples (see Table 2).

Table 2.
Ball Test and Bead Test results of selected Samples 1 & 2 (Home and Lab-Control Conditions).

01	Sample	Sample 1 (N = 16)		Sample 2 (N = 13)	
02	Condition	Home	Control	Control	Home
03	Test	Ball I	Ball I	Bead	Ball II
04	Trials	11,040	9,360	15,244	7,800
05	Hits	3,996	2,628	3,400	2,033
06	% expected	20	20	20	20
07	% observed	36.20	28.08	22.30	26.06
08	% surplus	81.00	40.40	11.50	30.03
09	$ES = Z/\sqrt{n}$	0.405	0.201	0.053	0.152
10	Z_{binomial}	42.53	19.5	7.1	13.38
11	p	$< 10^{-15}$	$< 10^{-15}$	10^{-12}	$< 10^{-15}$

Note: If clarification is needed, see the Notes beneath Table 1

This mode of analysis confirms that hit-rates of the two samples tested under lab control are extraordinary. The issue of situational conditions upon psi, however, is a thornier issue. We noticed that conducting psi tests in the lab did not have detrimental effects for all participants alike. Curiously, for three high scorers at home, for Ahmed, Amelie, and Silke, hit-rates under lab control were vastly greater, as is shown in Figure 4 and Table 3. This observation can hardly be explained in terms of diminishing psi-conducive conditions, nor did control for them improve such conditions. The conclusion remains ambiguous: The presence of experimenters might exert either hit-diminishing (more frequently) or hit-enhancing effects (less frequently) on psi-gifted participants. In any event, the occurrence in three cases of psi-enhancing effects by lab control shows that fraud and bias and the like are largely irrelevant when it comes to explaining hit-score differences between controlled and uncontrolled experiments.

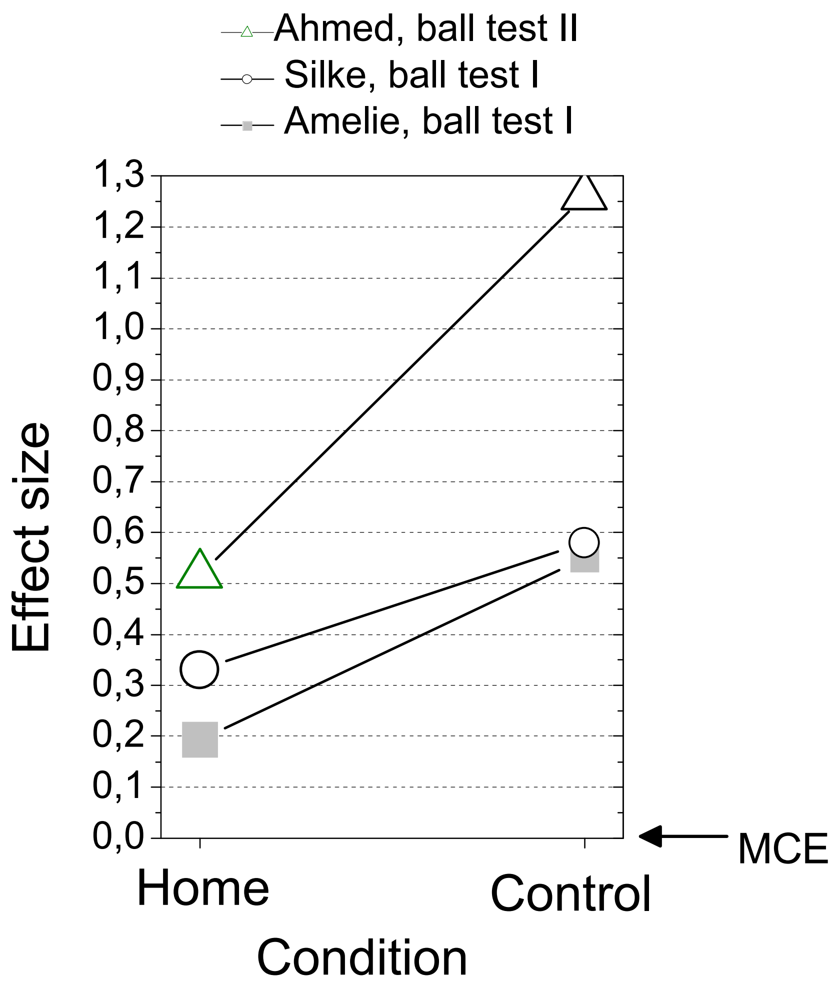

Figure 4.
Effect size of hit-rates obtained by three participants under home (left) and lab control (right) conditions. Curiously, their hit-rates increase under lab control.

It is not claimed here that negligent or fraudulent actions cannot have occurred under home test-conditions in any single case. The claim merely is that that if misconduct occurred at all, its contribution to the overall deviation of hit-rates from chance is negligible.

Table 3.
Original and Derived Data of Three Participants Whose Hit-Rates Increased Significantly and Extraordinarily Under Lab Control.

	Participant	Amelie J.		Silke T.		Ahmed K.	
01		Home	Control	Home	Control	Home	Control
02	Test	Ball I	Ball I	Ball I	Ball I	Ball II	Ball II
03	Trials	240	960	360	480	480	480
04	Hits	67	436	120	204	123	230
05	% expected	20.0	20.0	20.0	20.0	10.0	10.0
06	% observed	27.9	45.4	33.3	42.5	25.6	47.9
07	% surplus	39.5	127.0	66.5	112.5	156.0	279.0
08	$ES = Z/\sqrt{n}$	0.193	0.550	0.330	0.580	0.517	1.260
09	$Z_{binomial}$	7.83	29.5	6.26	12.27	11.33	27.6
10	p	10^{-14}	$<10^{-40}$	10^{-9}	10^{-34}	10^{-29}	$<10^{-40}$
11	Z (of difference)	4.91		2.70		7.16	
12	p	10^{-6}		3.50×10^{-3}		10^{-12}	

DISCUSSION

We set out by asking, Do high-scorers under home test-conditions produce random hit-rates under lab control? If they do, hit surpluses obtained at home may be due to bias, fraud or other such ‘normal’ means, and the psi hypotheses would be unsupported. However, the results show that high-scorers under home-conditions, when tested under lab control, did not produce hit-rates at random. Paranormal abilities must have been effective when they took the tests under lab control, hence it would be absurd to surmise that such ability was *not* effective when lab control was absent.

We also asked, If hit-rates under lab control are lower than without lab control, but *not* random, then two additional questions must arise: Are the remaining hit surpluses still strong enough to allow for assuming paranormal factors?; and, If they are strong enough, why do hit-rates under lab control decline? The results presented in Figures 2 and 3 showed that hit-rates of high home-scorers, even though lower on average under lab control, were in fact strong; hence, nothing other than paranormal factors (such as statistical flukes, etc.) need be considered. The only issue of uncertainty is why hit-rates declined in the lab. Two main factors must be considered.

One straightforward factor is regression to the mean (RTM). For simple stochastic reasons, exceptional performances tend to be followed by more normal performances, hence the observed change in hit-rates from *much* to *less* might at least partially be explained by RTM.

More important, a decline by lab participants might be due to the lessening of psi-conducive conditions. One of the participants of Sample 1 expressed concern: "I am afraid that if my score comes out random under lab control, you might think that my home records were fudged." Other lab-tested participants might have expressed similar sentiments. Researchers of the paranormal regard stress, tension, and anxiety as psi-inhibiting factors. An undisputed requirement for psi research is to provide "social ambience" in order to make fearful participants feel more comfortable: "Interaction with the subject should be calm, friendly, positive, and unhurried" (Reinsel, 1999, p. 2). Yet stress-free conditions, as are generally experienced at home, can hardly be brought about in the lab by calmly offering coffee and cool drinks.

One final comment on the difference between Sample 1 and 2. Hit-rates of Sample 2 were less pronounced than those of Sample 1, under both, home and lab conditions (see Figure 2). Apparently, the difference in home test performance is simply due to applying different selection criteria, which were less strict for Sample 2 than for Sample 1. Part of the difference under the lab condition may therefore be explained by an overall less pronounced psi ability of Sample 2 participants. In addition, the test used in the lab for Sample 2 high scorers (the bead-selection test) differed from the test that they had used at home (ball test II). Procedural changes leading to some loss of familiarity with conditions might have an unfavourable effect on performance. In short, comparing the two sample's performances cannot lead to safe conclusions and may therefore be abandoned.

Conclusion

The results suggest that hit-rate excesses in multiple-choice tests obtained by samples of participants alone at home, without lab control, do indicate paranormal power since quite a few high scorers at home are capable of producing significant hit deviations under lab control as well.

One might also conclude that the *first-home-then-lab* test strategy may be recommended as a methodological rule: Experimental psi research should preferentially be conducted with selected samples of psi-gifted individuals. Psi-gifted individuals are rare. Nevertheless, they may be detected, as this study shows, almost without costs, by using appropriate tests, by letting them test themselves at home, by testing high home-scorers in the lab, and by subsequently inviting successful lab performers to participate in further parapsychological studies.[5] The common lament of the para-community and their critics about "tiny" and "elusive" experimental psi effects might eventually lose its grounds and henceforth be remembered as symptom of past methodological deficiency.[1]

REFERENCES

Bierman, D. J. & Gerding, J. L. F. (1992). Towards a reduction of experimenter control in the study of special subjects. *The Parapsychological Association 35th Annual Convention. Proceedings of Presented Papers.*

Ertel, S. (2005). The ball drawing test: Psi from untrodden ground. In M. A. Thalbourne & L. Storm (Eds.), *Parapsychology in the 21st century: The future of psychical research* (pp. 90-123). Jefferson, NC: McFarland.

Gerding, J. L. F., & Bierman, D. J. (1997). Theory-based control in an unattended forced-choice psi study with a special subject. *The Electronic Journal of Anomalous Phenomena*, *97*.1.

Milton, J. (1996). Establishing methodological guidelines for ESP studies: A questionnaire survey of experimenters' and critics' consensus. *Journal of Parapsychology*, *60*, 289-334.

Reinsel, R. (1999). Advice for newcomers to parapsychology. I: Preliminary notes on experimental procedure for GESP studies. Retrieved

---

[5] This rule implies that high-scorers in the ball- or bead-selection tests exhibit generalised psi ability effective with telepathy, precognition, PK, etc. under variable conditions. The predictive power (validity) of ball-selection test-results has been investigated with a smaller number of high scorers and has been sufficiently confirmed (to be published elsewhere).

[1] J. B. Rhine (1977) was well aware of the importance of testing psi-gifted subjects, but he apparently thought that "special" subjects would show up by chance: "Among other advantages, the use of a special subject considerably reduces the *uncertainty and loss of time required by the search for a selection of subjects from the general population*. . . . Perhaps the main lesson to be learned . . . is how best to obtain and prepare subjects from the general population, and, on the other hand, how best to find the good subjects and keep them good." (p. 44, emphasis added). However, Rhine did not actually teach us that lesson.

December 4, 2005, from <http://www.parapsych.org/PDF/GESP_research_guidelines.pdf>

Rhine, J. B. (1977). History of experimental studies. In B. B. Wolman (Ed.) *Handbook of parapsychology* (pp 25-55). New York: Van Nostrand Reinhold.

Georg-Elias-Müller Institut für Psychologie
Georg-August Universität
37073 Göttingen
Germany

Email: sertel@uni-goettingen.de

APPENDIX

An analysis of the bead-selection test (Cohort 7, 2004) showed an unexpected peculiarity. Preceding ball test data had not shown notable hit differences among the five number targets. Surprisingly, the bead test data did reveal, for 11 individuals out of the total (N = 48), significant hit differences among the five colour targets. Another unexpected observation was that average hit excesses, summed across colours, were much less marked than average hit excesses, summed across numbers in the ball test I condition (hit-rate = 20.54%, where *MCE* hits for beads = 20%, percent deviation from *MCE* = 2.7%)—the difference is less than half of the lowest hit surplus among the ball test I results of Cohort 2001. It thus seems as if, for the bead test, psi had bifurcated into two partial effects, one accounting for an average hit level and another one for variance across colours (differential effect). Average hit level and differential effects showed almost equal occurrences (significant average hit-rates: 12 cases, significant differential cases: 14), and they did not correlate with each other (significant average hit-rates only: 8 cases, significant differential cases only: 6; both effects combined: 6 cases). A comparison between ball and bead test results in Figure 2 has been effected by combining, for the bead test data, average and differential effects (by adding log p) and by transforming the result to a Z value, which comes to 7.57 ($p = 10^{-13}$). Note that in terms of significance, the results of the bead test (combined analysis) compares with the result of the best ball-selection cohort, while the effect size is considerably lower. This seeming discrepancy can be explained by the fact that the number of trials for the bead test is much larger. The average testing time for the ball and the bead tests is roughly equal though, due to procedural differences. The bead test requires less record taking (see *Procedure* section above).