PSI TEST FEATS ACHIEVED ALONE AT HOME: DO THEY DISAPPEAR UNDER LAB CONTROL?

Suitbert Ertel

Georg-Elias-Müller-Institut für Psychologie, Göttingen

ABSTRACT

Extraordinary hit rates from multiple choice tests, obtained by participants alone in their homes, are ambiguous. On the one hand, their feats might reflect psi power manifesting itself better under informal home than under lab conditions. Yet hit excesses obtained without control might also be due to negligent or fraudulent conduct. One way out of this dilemma is to let participants complete 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 I) and N = 47 (sample II) of student participants. Sample I (female 84%) completed the ball drawing test, version I. Table tennis balls are drawn from an opaque bag on which numbers 1 to 5 are written, each number on ten balls. Participants guess and draw numbers blind and record their guessed and drawn numbers (hit expectancy 20%). One test unit consists of six or eight runs comprising 60 trials each (total 360 or 480 trials). Participants shake the bag prior to each trial and put drawn balls back into the bag. Sixteen high scoring participants of sample I were also tested, using the same test, under lab control. Sample II (female 73%) completed the ball drawing test, version II. This test resembles test I except that green or red dots are sprinkled over the balls, participants guess numbers (five targets) and colours (two targets), the combined expectancy being 10%. Thirteen high scorers of sample II were also tested under lab control using a pearls drawing test where they draw one of five colours (no numbers, expectancy 20%). Hypotheses: 1. Hit rates of high scorers in home tests decline (due to less psi-conducive conditions under control and regression towards the mean). 2. Hit rates of high scorers under control score still significantly above chance (due to genuine psi which was also effective at home). Both hypotheses were confirmed with sample I and replicated with sample II. *Sample I* obtained an average ES = 0.369 (sd=.126) under home condition and an average ES = .122 (sd = .207) under control (ES = $Z/\sqrt{(N_{trials})}$), the difference is significant ($t_{(corr.smpls.)}$ = -3.02, df = 15, p = .004, one-tailed). The significance of hit rates under control is p = $<10^{-15}$ by Chi^2 = 721.7, df = 16 ($Chi^2 = \Sigma(Z)(Z)$). *Sample II* obtained an average ES = .275 (sd= .162) under home condition and an average ES = .098 (sd = .153) under control, the difference is significant (t = -3.00, df = 12, p = .006, one-tailed). The significance of hit rates under control is p = $<10^{-15}$ by Chi^2 = 223.9, df = 13. Surprisingly, three participants obtained significantly *higher* hit rates under control compared with their home performance. The issue of fraud and bias loses relevance in view of such finding. It is recommended to introduce the "*first-home-then-lab-test*" strategy in parapsychological research on a broader scale. Once this strategy were generally applied, the widespread lamenting of psi researchers and their critics about tiny and elusive experimental psi effects might come to an end.

INTRODUCTION

Out of necessity, not out of choice (funds and assistants were lacking) 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 complete psi tests at home. The students, participating at this project out of curiosity and/or with intent to obtain obligatory credit points, were instructed in class, took the material home, ran the required 360 – 480 trials within 8 days (total testing time 1 ½ hours), returned their record sheets and waited for feedback from the experimenter who analyzed the data individually and who informed each participant in detail, by phone or email, about his or her results. Five annual cohorts of student beginners (total N = 238, female 84%) were tested using the ball drawing 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, female 73%): Each target ball carried numbers 1 to 5 plus green or red dots, the participants guessed numbers and colors. The last student cohort # 7 (N = 48, female 86%) had pearls to draw from a box, five different colors (no numbers) served here as targets. The median age of the total is 23.2 years, age differences between cohorts are negligible.

The ball drawing test yielded considerable hit surpluses from chance (see Table 1 and Figure 1, open circles). Hit rates declined with cohorts 2000–2001, but cohort 2002 regained the level of cohorts 1998/1999. Each cohort's hit surplus was highly significant. The results of ball test II exceeded those of ball test I. An analysis of the pearls drawing data is more complex[1], 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).

Table 1: Results of home tests by student beginners across seven cohorts (columns). Balls tests I and II and pearls test (see row #02). Rows ## 08 - 11 provide effect size and inferential results.

01	Cohort no.	1	2	3	4	5	6	7	
02	Cohort year	1998	1999	2000	2001	2002	2003	2004	
03	Test	Balls I	BallsI	Balls I	Balls I	Balls I	Balls II	Pearls	
04	Targets	Numbers	Numbers	Numbers	Numbers	Numbers	Ns. & colours	Colours	
05	N Ss	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	% exp	20.00	20.00	20.00	20.00	20.00	10.00	20.00	
09	% obs	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/√Trials	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	3.54*	7.6
11	p	10^{-8}	10^{-6}	10^{-6}	10^{-4}	10^{-13}	$<10^{-15}$	0.0002	10^{-13}

Note: % exp = expected hit percentage. % obs = observed hit percentage. % surplus = percent of hits above expectancy. Example: 22 hits, 20 expected = 10% surplus. Expectancy is different in column, 6, see row 06. Col 7 left = Hit average. Col. 7 right = Combined hit avg. and differential color effect. * $Z_{Poisson}$ = 3.50.

Yet all such results might be doubted because the observed deviations from chance might be brought about by negligent participants who might gain leaking sensory information or who might deceive themselves by inadvertently recording wishful hits while actually having misses. Without control, even intentional deception is feasible. Unlike ordinary psi test data where bias and fraud are meticulously ruled out, the present data base is not safe.

[1] The significance of hit rate deviations of the pearls test (hit counts fixed, trial counts variable) should generally be obtained by applying the Poisson test (Timm, 1968, 1994). For large trial counts, however, as for the present total of 70,121 trials, Binomial test results are numerically almost equal to Poisson test results.

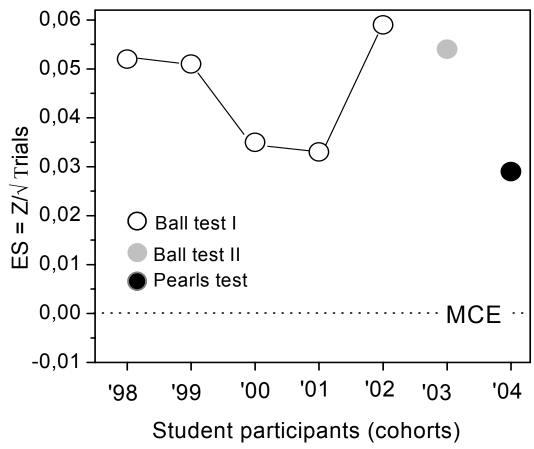

Figure 1: Effect size of hits for seven cohorts of student beginners, cohorts 1998-2002 took the ball drawing test I, cohort 2003 the ball drawing test II, and 2004 took the pearls drawing test.

However, unlike aseptically clean data where bias and fraud *cannot* have occurred, the present data are suitable for investigating and possibly finding out, post hoc, whether or not bias, e.g., by sensory leakage or memory help or fraud *have actually* occurred.

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 pearls must be perceived, stored and applied so as to become helpful subsequently for better hitting. The same applies to wishful errors at record taking that might eventually increase due to gradually fading attention. But the students' hit rates, those of high and low scorers alike, exceed chance expectancy already with first trials, there is no indication of increasing performance over time (Ertel, in print). [2]

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.[3] Student participants do *not have to* complete these tests, most participants take the opportunity to find out whether they score well. They cannot satisfy such curiosity by manipulating their records.[4] Self report information by students (14-15 ys) about cheating in class (Anderman et al., 1998) shows that 39% admitted having cheated occasionally when taking exams or completing homework. Extrinsic motivation (getting grades etc.) and expected reward (escaping effort) turned out to be the prime motivation. The sample of students used in the present ESP studies are adults, no grades or other extrinsic reward could be expected for good hit rates, rather intrinsic motivation was manifested in responses to a post-experimental questionnaire of the pearls test sample (N = 48): Among 28 items addressing attitudinal and affective reactions to this test the first eight of an average rank order of affirmation are (1) accuracy, (2) interest, (3) curiosity, (4) joy when getting hits,

[2] An earlier general objection to ESP methods raised by Goodfellow (1938), Feller (1940), Brown (1953), Pöppel (1967), and Gatlin (1979) and revived by Brugger&Taylor (2003) has been scrutinized elsewhere (Ertel, submitted). This objection drawing upon patterns of calls and hypothetical patterns of targets matching with call patterns did not find support by various empirical tests using the author's ball test data.

[3] Hansen (1990) advocates considering not only the possibility of cheating, but the probability of its occurrence in the first place *("...One must first assess the likelihood of fraud being attempted... [Considering probability,] even critic C. E. M. Hansel (1966) has stated: 'It is unlikely that more than a small number of experiments on ESP are affected by cheating'* " (Hansel, 1990, p. 32).

[4] Self report information by students (14-15 ys) about cheating in class (Anderman et al., 1998) shows that 39% admitted having cheated occasionally when taking exams or completing homework. Extrinsic motivation (getting grades etc.) and expected reward (less effort) turned out to be the prime motivation. The sample of students used in the author's ESP studies are adults, no grades or other extrinsic reward could be expected for good hit rates, rather intrinsic motivation was manifested. Moreover, regarding the proportion of cheaters at exams it should be considered that occasional cheaters do not cheat at every exam, the average proportion of cheaters at individual exams is much less than 39%.

(5) conscientiousness, (6) eagerness, (7) doubt of extrasensory perception, (8) fun. The eight most frequently negated items are (28) concern, (27) competition with fellow students, (26) lack of concentration, (25) fears, (24) strange self perceptions, (23) confusion, (22) thoughts about possible cheating, (21) statistical calculations.

Moreover, the participants had been informed by instruction that high scorers at home would be invited to subsequent lab tests under control. If a student under home condition 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. [5]

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

METHODS

Materials

Ball drawing test I (numbers only): An opaque bag containing 50 table tennis balls. The numbers 1 to 5 are written on them, each ball carries one number. Ten balls are used for each number. The face used is inaccessible to kinesthetic perception. Record forms for the home tests are provided where participants have to fill in their guessed and drawn numbers.

Ball drawing test II (numbers and colors): This test differs from the foregoing test only in that each ball is sprinkled, in addition, with either red or green dots. Red and green colors are used equally often. The record forms provide blank space to be filled in with guessed and drawn numbers and colors.

Pearls drawing test: A box containing 1,500 little handicraft pearls (diameter: 4 mm) whose colors are red, green, blue, yellow or white, 300 pearls for each color. Record forms provide blanks to be filled in with frequencies of blind draws which are made to pick a desired color (see procedure).

Procedure

Ball drawing tests: Each trial starts off with first putting down, on the record sheet, a guessed number (or number plus color, respectively) that the participant supposes to draw on the next trial. He then shakes the bag, closes his eyes, reaches into the bag, picks and draws a ball, checks the actually drawn number (and color) and writes it (them) down. He puts the ball back into the bag. Each run requires 60 trials, one ordinary test series requires six runs.

For the ball drawing tests, the number of trials is fixed while the number of hits is open to variation as is common with multiple choice procedures. For the pearls drawing test, the number of *hits* is fixed while the number of *trials* is open to variation, as follows:

The pearls drawing test: The record form is made up of rows with 10 blanks each. The participant fills them in with *frequencies of blind draws which were needed until a pearl with the desired target color was picked.* Prior to starting off with one row of blanks, the participant decides the target color for that row. Supposing he chooses "red" and picks a red pearl at his first trial, he fills in "1", if another pick of a red pearl requires, say,

[5] The idea of "reducing experimenter control in a study of special subjects" is not new. Bierman & Gerding (1992) advocate more relaxed ("sloppy") testing conditions in case of "special" subjects, control might be "reduced to theoretically relevant" factors. "It is argued 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" (Gerding & Bierman, 1997, p. 2).

10 draws, he fills in "10" in the second blank etc. Each draw of a pearl is preceded by stirring the pearls in the box with stretched five fingers. The pearls are drawn with eyes closed. Drawn pearls are put back into the box. One sub-series consists of 10 rows, thus 100 hits are required for one sub-series, MCE being 500 draws. Three sub-series, i.e. 300 hits, MCE 1,500 draws, are needed for one complete unit. The participants are told to use the five colors row-wise as targets within one sub-series equally often.

Testing under lab control

The participant completes the test sitting opposite the experimenter. The participant's guesses and draws are filled in the blanks by the experimenter who also observes the participant's behavior. Attempts to peep into the box or other such fraudulent actions would hardly go unnoticed, apparently they have not occurred.

Participants and experimenters

The ball drawing test I (numbers) was completed by 238 students, the ball drawing test II (numbers and colors) by 47 students. From ball test I participants 16 high scorers completed the *same test* under control. Among them were 5 non-students, two boys (8 and 12 years) and three adults who had shown high hit rates at home. Of the balls test II sample, students only, 13 high scorers completed the *pearls test* under control (second control sample). High scorers among students who completed the pearls test at home (N = 49, the most recent cohort), have not (yet) been tested under control. The author acted as experimenter for 15 of 16 participants of the first control sample and for 8 of 13 participants of the second control sample. Six control experiments were conducted with student assistants serving as experimenters.[6]

Data analysis

Mean chance expectancies (MCE) are as follows: 0.2 for the balls test I, 0.1 for the balls test II, and 0.2 for the pearls test (see Table 1, row #06). The ball test data are subjected to the Binomial test. The pearls test data (hit count fixed, trial count variable) might be subjected to the Poisson test (as advocated by Timm 1968 and 1994). But there are grounds to apply the Binomial test consistently.[7] Each individual Binomial Z value is transformed into effect size $ES = Z/\sqrt{T}$ [T = Number of trials]. The significance for samples of individuals is obtained by $Chi^2 = \Sigma Z^2$, df = N [N = number of individuals].

RESULTS

Figure 2 shows results of sample #1 participants, i.e. of those who completed ball test I under home and lab conditions. Effect size measures of hit rates obtained under home (left) and control condition (right) are plotted. Observed effect size of the total (lower dot, left) is plotted for comparison. Necessarily, 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 large. More important, the effect size under control (right) is remarkably lower than the effect size obtained under home condition (left). However, even though some participants obtained hit rates under control at chance level only (N = 6, = 38% of 16), hit rates of the high scorers' total remained highly significant under that condition ($Chi^2 = 721.7$, df = 16, $p < 10^{-15}$).

[6] Three student experimenters had completed the test under home and lab conditions earlier (control by SE). They tested two students each.

[7] The pearls' test condition is atypical for applying the Poisson procedure which is applicable in cases of very rare events (p<.20) and when the sample size is unknown. The pearl's test probability is .20 and the number of hits is fixed (N = 300), hence the range of necessary trials is also *roughly* known, empirically (1100 – 1600) (see also Footnote 3).

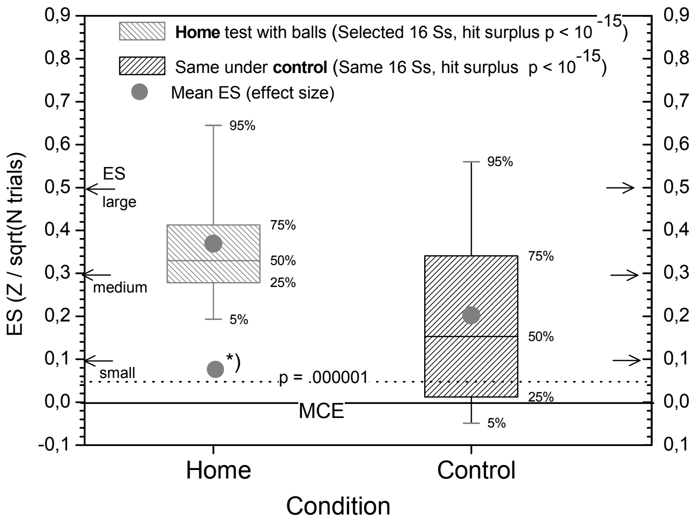

Figure 2: Sample #1 results (N = 16). Effect size from ball test I obtained at home (left) and under control in the lab (right). The lower left full circle, marked *), represents the observed effect size of the total from which sample #1 had been selected (19,442 hits, exp. 0.2, by 85,200 trials of 238 Ss). The dotted p-line is based on the total.[8]

Figure 3 shows results of sample #2 participants (N = 13), i.e., of those who completed ball test II at home (left) and the pearls test under lab conditions (right). The mean hit rate of high scorers of sample #2 is less pronounced than that of sample #1 (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 control, four of 13 participants obtained hit rates at chance level (= 31% of 13), nevertheless, the hits total of high scorers remained highly significant under that condition ($Chi^2 = 223.9$, $df = 13$, $p < 10^{-15}$).

[8] Box plots show means (here circles) medians (here solid lines within boxes), 25% and 75% centile values (lower and upper boundaries of the boxes) and 5% and 95% centile values (lower and upper ends of vertical lines). The end of the lower vertical (5%) of controls (right) extends below ES = 0 because a small number of participants with significant hit surpluses at home obtained hit scores below expectancy under control.

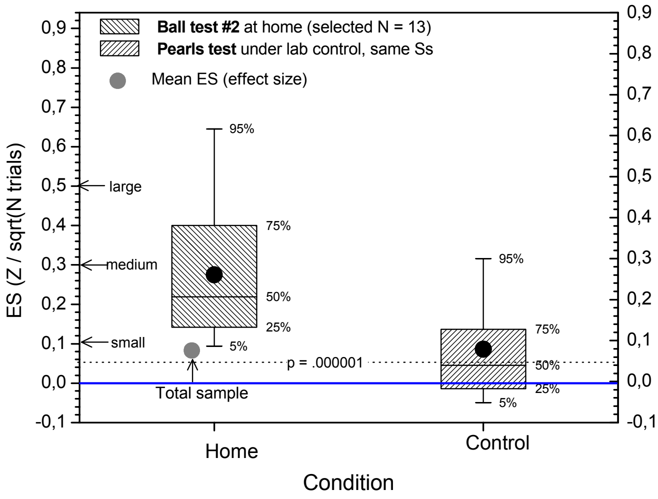

Figure 3: Sample #2 results (N = 13). Effect size from ball test II obtained at home (left) and from the pearls test under control (right). The lower left full circle, marked *), represents the observed effect size of the total from which sample #2 had been selected (2,620 hits, exp. 0.1, by 22,560 trials of 47 Ss). The dotted p-line is based on the total.

For the above data analysis individual scores were taken as units, the box plots of Figures 1 and 2 are based on them. Alternatively, hits may be summed across participants ignoring individual hit rate variation. This procedure has been applied with Table 1 and is taken up again with Table 2 which summarizes results for the two control samples (see Table 2).

Table 2: Results of samples #1 and #2 obtained under home condition (ball-drawing I for sample I, ball drawing II for sample II) and under control condition (ball drawing I for sample I) and pearls drawing test (for sample II).

	Sample	I		II	
	N participants	16		13	
01	Condition	*Home*	Control	*Home*	Control
02	Test	Balls I	Balls I	Balls II	Pearls
04	Trials	11,040	9,360	7,800	15,244
05	Hits	3,996	2,628	2,033	3,400
06	% exp	20	20	20	20
07	% obs	36.20	28.08	26.06	22.30
08	% surplus	81.00	40.40	30.03	11.50
09	ES=Z/√Trials	0.405	0.201	0.152	0.053
10	$Z_{Binomial}$	42.53	19.5	13.38	7.1 *
11	p	$<10^{-15}$	$<10^{-15}$	$<10^{-15}$	10^{-12}

Note: * $Z_{Poisson}$ = 6.73
For other clarifications see Notes of Table 1

This mode of analysis confirms that hit rates of the two samples tested under control are extraordinary.

Discussion

We set out by asking: "(1). Do high scorers under home test conditions produce random hit rates under control? If they do, hit surpluses obtained at home may be explained as due to bias, fraud or other such 'normal' means, psi hypotheses would be dispensable." The results show that high scorers under home conditions, when tested under control, did not produce hit rates at random. Paranormal abilities must have been effective when they completed the tests under control, hence it would be absurd to surmise that such ability was not effective when control was absent.

We also asked: "2. If hit rates under control are lower than without control, but not random, two additional questions arise. 2.1 Are the remaining hit surpluses still strong enough to allow for assuming paranormal factors? 2.2 If they are strong enough, why do hit rates under control decline?" The results presented in Figures 2 and 3 showed that hit rates of high home scorers, even though lower on average under control, were in fact strong, hence no other than paranormal factors (such as statistical flukes etc.) need consideration. The only issue of uncertainty is why hit rates declined in the lab. Two main factors must be considered.

One straightforward factor is regression towards the mean (RTM). For simple stochastic reasons, exceptional hit rates which may be due to or boosted by chance fluctuation tend to be followed by less outstanding performance, hence the observed change in hit rates from *much* to *less* might at least partially be explained by RTM.

More important, a decline by control 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 control, you might think that my home records were fudged.*" Other lab-tested participants might have felt similarly. 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). Yet stress-free conditions as are generally experienced at home can hardly be brought about in the lab by kind words with offering cakes and drinks.

The issue of whether situational conditions affect psi performance, however, is thornier. 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 control even increased immensely as is shown in Figure 4 and Table 3. This observation can hardly be explained in terms of diminishing psi-conducive conditions since for them effects by control cannot have surpassed those of a relaxed home atmosphere. The conclusion remains ambiguous: The presence of experimenters exerts either hit diminishing (more often) or hit enhancing effects (less often) on psi-gifted participants. In any event, the occurrence in three cases of psi-enhancing effects by control shows that fraud and bias and the like are largely irrelevant when it comes to explain hit score differences between controlled and uncontrolled conditions.

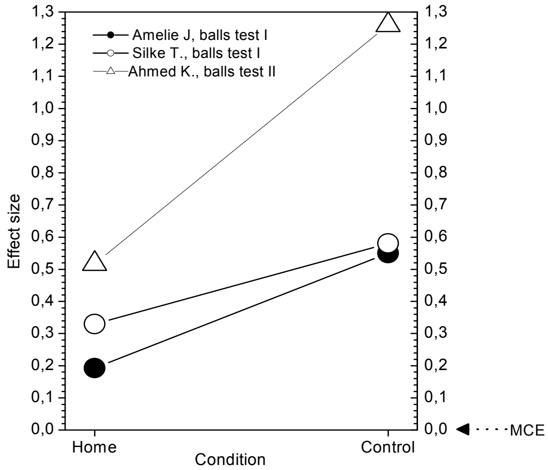

Figure 4: Effect size of hit rates obtained by three participants under home (left) and lab control (right) conditions. Curiously, their hit rates increase under lab control.

Table 3: Original and derived data of three participants whose hit rates increased significantly and extraordinarily under control. Significance tests for Binomial Z's and fir differences of proportions.

	Participant	Amelie J.		Silke T.		Ahmed K.	
01		Home	Control	Home	Control	Home	Control
02	Test	Balls I	Balls I	Balls I	Balls I	Balls II	Balls II
03	Trials	240	960	360	480	480	480
04	Hits	67	436	120	204	123	230
05	% exp	20.0	20.0	20.0	20.0	10.0	10.0
06	% obs	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/√Trials	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}		.0035		10^{-12}	

Note: If clarifications are needed, see Notes of Table 1

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 if misconduct occurred at all, its contribution to the overall deviation of hit rates from chance is negligible.

One final comment on the difference between samples #1 and #2: Hit rates of sample #2 were less pronounced than those of sample #1, under both, home and lab conditions (see Figures 2 and 3). The difference in home test performance is simply due to applying different selection criteria, they were less strict for sample #2 than for sample #1. Part of the difference under the lab condition may therefore be explained by less pronounced psi abilities of sample #2 participants. In addition, the test used in the lab for sample #2 high scorers (the pearls drawing test) differed from the test that they had used at home (the ball drawing test II). Procedural changes effectuating some loss of familiarity with conditions might have an unfavorable 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 completing ESP tests alone at home, without control, do indicate paranormal power since the majority of high scorers at home are capable of producing significant hit deviations as well under control.

One might also conclude that the test strategy applied in these tests, a *first-home-then-lab* strategy, might 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 inviting successful lab performers subsequently to participate as subjects at further parapsychological research.[9] The para-community's and their critics' common lamenting about "tiny" and "elusive" experimental psi effects might eventually lose its grounds and henceforth be regarded as symptom of a methodological fallacy of the past.

Contemporary researchers should have followed, in principle, model designs of pioneers of parapsychological research who gave many examples of continuously testing psi-gifted subjects over long periods who showed considerable consistency of performance (Martin & Stribic, 1938, 1940[10] , Birge, 1942, Ryzl, 1966, review of results from star subject Pavel Stepanec, Kelly, 1972, Musso, 1973, Houtkooper & De Diana,1977, Schouten &Kelly, 1978, Sasaki & Ochi ,1982 , Hearne, 1984, Jacobs, 1985, Sijde, 1987). Replicability of performance, provided that the testing conditions do not change, has also been claimed elsewhere (example: Schmeidler, 1964).[11] Recommendations to conduct psi-tests with high-scoring participants whose performance is more reliable than that of low-scoring participants have been given early on (Schmeidler, 1948, Murphy, 1969) [12]. The complementary advice to forego unselected participants altogether if "fact-finding repeatability" is needed is not novel either (Whiteman, 1972). [13] It seems as if insights of past generations have been unduly sacrificed in favor of fashionable developments, concurrent detrimental effects on experimental psi manifestation went unnoticed.

References

Anderman, E. M., Griesinger, T., & Westerfield, G. (1998). Motivation and cheating during early adolescence. *Journal of Educational Psychology 90*, 84-93.

Bierman, D. J. & Gerding, J. L. F. (1991). Towards a reduction of experimenter control in the study of special subjects. *The Parapsychological Association 35th Annual Convention. Proceedings of Presented Papers.*

---

[9] This rule implies that high scorers in the balls or pearls drawing tests exhibit generalized psi ability that should manifest itself as telepathy, precognition, PK etc. under variable conditions. The predictive power (validity) of ball drawing test results has been investigated with a smaller number of high scorers and has been sufficiently confirmed (to be published elsewhere).

[10] J. B. Rhine 1977) was aware of the importance of testing psi-gifted subjects, but he apparently thought that "special" subjects would show up by luck: "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* (italics by SE)….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." (Rhine, 1977, p. 44). But Rhine did not actually teach us such lesson.

[11] "*The author concludes that there are detailed, meaningful, and self-consistent patterns within ESP data which can be obscured if all of a subject's responses are pooled*" (Schmeidler, 1964, abstract). "*In regard to future work in parapsychology, we must move flexibly to new problems that are related to three main issues: First, finding the most favorable conditions under which gifted subjects can do their best work …*" (Murphy, 1969).

[12] "…it is suggested that any experiments with gifted individuals could probably be completed with greater clarity and efficiency than those with subjects chosen at random. It may even be that sensitive subjects could disclose energy relations or physiological correlates of psi that otherwise would be imperceptible because our instruments were not sufficiently delicate or our research accurately enough aimed to detect such faint effects without them." (Schmeidler, 1948, abstract).

[13] …Unselected subjects should not be used in experiments aimed at fact-finding repeatability…" (Whiteman, 1972, abstract). Repeatable experimental conditions are indispensible for replicating experimental outcomes (Hövelmann, 1984). Recruiting participants with sufficient psi-ability is one of the foremost – and repeatable – experimental conditions enhancing replicability of paranormal effects.

Birge, W. R., & Rhine, J. B. (1942). Unusual types of persons tested for ESP. I. A professional medium. *Journal of Parapsychology, 6*, 85-94.

Bleksley.A.E.H. (1984). Summary of an experiment with a gifted subject on long-distance ESP during sleep. *Parapsychological Journal of South Africa, 5,* 85- 99.

Brown, G. S. (1953). Probability and psychical research. *Nature, 172,* pp.154-156.

Brugger, P. & Taylor, K. I. (2003), 'ESP. Extrasensory perception or effect of subjective probability?', *Journal of Consciousness Studies,10*, pp. 6-7. Editors: Alcock, J. E., Burns, J. E., & Freeman, A., Title of issue: Psi wars. Exeter, U.K.: Imprint Academic.

Ertel, S. (submitted). Are ESP test results stochastic artifacts? Brugger & Taylor's claims under scrutiny.

Ertel, S. (in print). 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.*

Feller, W. K. (1940). 'Statistical aspects of ESP'. *Journal of Parapsychology,* 4, pp. 271-298.

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.

Goodfellow, L.D. (1938), A psychological interpretation of the results of the Zenith radio experiments in telepathy. *Journal of Experimental Psychology, 23,* 601-632 (reprinted in *Journal of Experimental Psychology: General*, *121*, 1992, pp. 130-144).

Hansen, G. P. (1990). Deception by subjects in psi research. *The Journal of the American Society for Psychical Research, 84,* 25-80.

Hearne, K. M. T. (1984). A survey of reported premonitions and of those who have them. *Journal of the Society for Psychical Research, 52*, 261-270.

Houtkooper, J. M., & De Diana, I. P. F. (1977). Relationships between PK and physiological variables in a gifted subject. *European Journal of Parapsychology, 2*, 68-79.

Jacobs, J. C. (1985). Psi-guided awakening from sleep. 1: The original experiments of W. Van Vuurde. *Journal of the Society for Psychical Research, 53*, 159-168.

Kelly, E. F., & Kanthamani, B. K. (1972). A subject's efforts toward voluntary control. *Journal of Parapsychology, 36*, 185-197.

Martin, D. R., & Stribic, F. P. (1938). Studies in extra-sensory perception: II. An analysis of a second series of 25,000 trials. *Journal of Parapsychology, 2*, 287-295.

Martin, D. R., & Stribic, F. P. (1940). Studies in extra-sensory perception: III. A review of all University of Colorado experiments. *Journal of Parapsychology, 4*, 159-248.

Musso, J. R., & Granero, M. (1973). An ESP drawing experiment with a high-scoring subject. *Journal of Parapsychology, 37*, 13-36.

Pöppel, E. (1967). Signifikanz-Artefakte in der experimentellen Parapsychologie. *Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie*, 10, pp.63-72.

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

Schmeidler, G. R. (1948). Research projects in parapsychology. *Journal of Parapsychology, 12*, 107-113.

Schmeidler, G. (1964). An experiment on precognitive clairvoyance. Part II. The reliability of the scores. *Journal of Parapsychology, 28*, 15-27.

Sijde, P. C. (1987). Pavel Stepanek, a gifted subject. *Tijdschrift voor Parapsychologie, 55*, 23.

Ryzl, M. (1966). A model of parapsychological communication. *Journal of Parapsychology, 30*, 18-30.

Sasaki, K. S., & Ochi, Y. (1982). Observation of the deformation of pure aluminum plates by psychokinesis. *Journal of the Psi Science Institute of Japan, 7*, 7-13. 7 illus; 13 ref.: Journal of the Psi Science Institute of Japan.

Schouten, S. A., & Kelly, E. F. (1978). On the experiment of Brugmans, Heymans, and Weinberg. *European Journal of Parapsychology, 2*, 247-290.

Timm, U. (1968) Das „Round-Table-Experiment" – Ein neuer Versuchsplan für ASW-Gruppenversuche. . *Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie.* 11, 34-43.

Whiteman, J. H. M. (1972). On the concept of repeatability in scientific experimentation. *Journal of the American Society for Psychical Research, 66*, 227-229.

Address for correspondence: Suitbert Ertel,
Georg-Elias-Müller-Institut für Psychologie,
Gossler Strasse 14, D-37073 Göttingen, Germany.
Email: sertel@uni-goettingen.de