ISSN: 1445-2308
Australian Journal of Parapsychology
Volume 13, Number 1, pp. 111-131

Paranormal Effects and Behavioural Characteristics of Participants in a Forced-Choice Psi Task: Ertel's Ball Selection Test under Scrutiny

LANCE STORM, SUITBERT ERTEL, & ADAM J. ROCK

Abstract: In a previous study (Storm, Ertel, & Rock, 2013), we demonstrated that Ertel's (2005b,c) Ball Selection Test (a.k.a. the Ball Test) can be used to test Reactance Theory (Brehm & Brehm, 1981), and the Sheep-Goat Effect (SGE; the tendency for believers ['sheep'] to psi-hit and non-believers ['goats'] to psi-miss). The Ball Test is a forced-choice psi task involving the blind selection of numbered ping-pong balls *after* numbers are called by the participant (a hit is a match between a 'called' number and a 'selected' ball number). According to Reactance Theory, when an individual's freedom is threatened through some form of coercion, we may expect *reactance*, which is "a motivational state aimed at restoring the threatened freedom" (Silvia, 2005, p. 277). Reactance may explain the SGE. In that study (N = 82), we found a significant forced-choice hit rate of 21.06% where P_{MCE} = 20% (p = .002). Participants were randomly assigned to a control condition (n = 42) or treatment (n = 40) condition requiring them to read a statement that induced reactance. We found a significant reactance effect, and a significant SGE. In the present study, we sought answers to some crucial questions concerning the ball test's validity, including possible displacement effects, and possible biases in number-calling and selection-behaviour. We found that participants did avoid calling numbers if they had just called them or just selected them from the bag. We found no evidence of hits being mere artifacts caused by sensory leakage or mnemonic aids.

Keywords: Ball Selection Test, ESP, PK, reactance, psi, sheep-goat effect.

INTRODUCTION

The field of parapsychology has seen much criticism of the methods used to test alleged 'psi' (i.e., psychic) or paranormal abilities (Alcock

1981; Hansel, 1980). Driven by this criticism, parapsychology for many decades now has seen the development of innovative methodologies designed to ensure valid results. To that effect, the field has carved out a niche in social science research that, in some cases, is renowned for its advances in experimental design (for examples, see Bem, 2011; Beischel, & Rock, 2009; Houran & Brugger, 2000). Ertel's (2005c) so-called Ball Selection Test (a.k.a. the Ball Test), which is described in the next section, has emerged from the sidelines of this trend as a method that is reminiscent of older 'hands-on' approaches, yet is shown to have advantages for valid laboratory-based psi research (for details, see Ertel, 2005, 2013).

Hand in hand with the practical advances in parapsychology has been the need to develop an understanding of the intricacies and complexities of human psychology as it pertains to the psi response. In light of these issues, two phenomena—the perennial *sheep-goat effect* (see Schmeidler, 1943) and the more-recent *reactance effect* (see Brehm, 1981), and particularly the relationship between the two—were investigated in a recent paper by Storm, Ertel, and Rock (2013). The present study follows up the work by Storm et al. (2013) in an attempt to further understand this relationship.

The Ball Selection Test

In Ertel's (2005b,c, 2010b) standard Ball Selection Test, participants are handed an opaque bag containing 50 table tennis balls on which the numbers 1 to 5 are written, each number on 10 balls. A trial consists of the participant shaking the bag, guessing the number that s/he will subsequently select blindly from the bag, selecting (i.e., drawing out) a ball by hand through the small opening of the bag, checking the number, and putting the ball back in the bag. Participants are told to look away from the bag when they select a ball. The numbers guessed and selected are recorded.

Studies by Ertel (2005b,c) have produced significant hit rates using the Ball Test under home test conditions (without supervision). In a home test sample of 231 students (Ertel, 2005c), the average hit-rate deviation was 9% above mean chance expectation (MCE), which is significant, $z = 12.07$ ($p < 10^{-15}$). In a follow-up study (Ertel, 2005b), various tests were run. One test unit consisted of six or eight runs comprising 60 trials each (total: 360 or 480 trials). Sixteen high scoring home test participants of Sample 1 were tested under laboratory control, again using the Ball Test procedure (Version I). Sample 2 took the Ball Test (Version II) which is essentially the same as Ball Test Version I except that green or red dots are sprinkled over the balls, and participants guess numbers (five targets) *and* colours (two targets), where $P_{MCE} = 10\%$. Thirteen high-scorers of Sample 2 under home test condition were also tested under laboratory control using the bead-selection test where each participant selected one of five colours (no numbers, P_{MCE} = 20%). As hypothesized, hit-rates of selected high scorers under home test condition were significantly above MCE under laboratory condition.

In another recent study (Ertel, 2010b), nine ‘high-scorers’ under home test conditions (Version I), retested in a laboratory using Ball Test Version II, achieved a highly significant hit rate of 17.3% (P_{MCE} = 10%). Also, a replication of Ball Test Version II was conducted by two graduating student experimenters working under the guidance of a skeptical professor at the Anomalistic Psychology Research Unit (APRU) at Goldsmiths College, London University. Forty unselected APRU participants achieved a hit rate of 10.75% (where P_{MCE} = 10%), which was significant (p = .002). It is important to note that Ertel (2005c, 2010b, 2013) has gone to considerable lengths to demonstrate that psi hitting using the Ball Test design is very likely not an artifact of sensory leakage or mnemonic aid.

The Sheep-Goat Effect and Reactance Theory

Participants in paranormal experiments are considered to be either those who think that ESP is possible under a given experimental condition (‘sheep’), or those who rejected this possibility (‘goats’—see Schmeidler, 1943, 1945). These days, sheep are those who believe that “ESP exists as a genuine phenomenon” (Thalbourne, 2003, p.114), whereas goats do not hold this belief. Furthermore, it has been observed that goats try to avoid psi targets, whereas sheep tend to identify psi targets. When differential scoring is evident between sheep and goats, the effect is known as the ‘sheep-goat effect’ (SGE)—sheep tend to ‘psi-hit’, and goats tend to ‘psi-miss’. The so-called ‘vindication’ theory (Palmer, 1972) was proposed as a means by which psi-hitting could be regarded as the product of attempts on the part of sheep to ‘prove’ the psi hypothesis, whereas psi-missing from goats could be regarded as attempts to ‘disprove’ the same hypothesis (Palmer, 1971, 1972).

In our previous paper (Storm Ertel, & Rock, 2013) we argued that most psi experiments are designed to get ‘good’ performances from sheep, but they are not designed to get ‘bad’ performances from goats. Goats are merely treated like sheep (i.e., they are instructed to psi hit), but they are not instructed to psi miss. We find this situation ironic: How can experimenters realistically expect an optimal psi-missing effect from goats after telling them to try to get a hit? For sheep, of course, the reverse of this situation is not an issue of concern—sheep merely comply with the experimenter’s instructions to get more hits than expected. What we discovered is there is a consistently larger deviation from MCE in sheep compared to goats, with sheep frequently psi hitting, but goats not as often psi missing—this discrepancy was reported in the 1970s (Schmeidler & McConnell, 1973). Decades later, Steinkamp (2005) pointed out, in her review of Lawrence's (1993) meta-analysis of forced-choice studies, that there was still no reference being made to possible asymmetrical sheep-goat scoring differences (see also, Palmer, 1977).

If psi missing in goats is the effect parapsychologists are looking for, we argue that experimenters must manipulate goats' behaviour so as to produce psi missing. One way to do this is to induce reactance in goats. Reactance is the motive to do (or believe, or feel) the opposite of what an experimenter instructs a person to do (or to feel or to believe), which is induced by opinionated argument. If this manipulation is successful with goats, then goats might have a motive that is antithetical to that of sheep.

Much evidence has been found that reactance is induced when a person's attitudinal or behavioural freedom is threatened or reduced (Kraus, 1995; Wright, 1986). This reactance effect can graduate to a so-called *boomerang effect* (i.e., the individual adopts a noncompliant attitude in order to restore his/her initial freedom). Silvia (2005, 2006) has shown that reactance was highest in the group that felt the most threatened by the content of an opinionated communication (i.e., a reactance prime). Opinionated communications are the most widely used threat manipulations in reactance research (Brehm, 1966; Brehm & Brehm, 1981). (For a more detailed review of Reactance Theory and goats' behaviour, readers are referred to Storm et al., 2013.) To test reactance theory on goats, parapsychologists might apply such a reactance treatment (i.e., reactance prime) in the form of an opinionated communication—we can expect increased noncompliant behaviour in goats under threat, which may thus yield increased target *avoidance* and therefore shifts from chance scoring to psi missing.

Evidence that goats can be manipulated into changing their psi performance comes from a study by Storm and Thalbourne (2005), whose objective in a symbol-identification task was to see if pre-experimental skepticism in goats could be raised so as to 'convert' them from a psi-dismissive disposition to a psi-supportive disposition. It was hypothesized that naïve goats, after having the implications of significance testing explained to them, would shift from chance-scoring (or psi-missing) before the manipulation to psi-hitting after the manipulation. The hypothesis was supported—goats shifted from chance scoring in the pre-treatment stage (20%, where $P_{\mathrm{MCE}} = 20\%$), to psi-hitting in the post-treatment stage (30%, $p = .047$).

More recently, Storm et al. (2013) investigated the effects of reactance on psi performance using Ertel's (2005a,b) Ball Test. Participants completed up to four runs (60 trials/run) of paranormal target-seeking (trying to predict the numbers on table tennis balls). Hit rate for the whole sample (N = 82) was significant, 21.06% (p = .002), where P_{MCE} = 20% (psi effects are generally small in the forced-choice domain—however, see Storm, Tressoldi, & Di Risio, 2012, for a defense of the domain). Participants were randomly assigned to a control condition (n = 42) or treatment (n = 40) condition requiring them to read a statement that induced reactance. A significant reactance effect was found. Also, the relationship between psi hit rates and sheep-goat scores was positive and significant. Reactant goats scored significantly lower than control sheep, as expected, but not significantly lower than control goats.

Having amassed a detailed database in our initial reactance study (Storm et al., 2013), we decided that it was imperative that we scrutinize more closely the nature of the psi effect, as well as the selection behaviour of participants during the Ball Test. Having raised a few questions about our findings, we argued that further investigation might reveal new information about Ball-Test participants' behaviour and the psi process. We based this assumption on two facts: (a) we already amassed evidence supportive of psi, and indeed, sheep-goat effects, in our initial study, and (b) we had anecdotal evidence that some participants had adopted rather unusual ball-number calling and ball-selection behaviour, but we were not able to state categorically how this behaviour was 'patterned'. Before we list our formal questions, the following terms (*good missing* and *delayed target selection*) are described and defined.

Good Missing

In an earlier study, Ertel (2005c) unexpectedly came across an apparent secondary psi effect which might be classified as a form of displacement.[1] Ertel called it "good missing", which is essentially a form of unintentional or subconscious behaviour. The phenomenon came to his attention when he analyzed the data of his standardization sample of 238 participants who took the test alone at home. Ertel noticed that, if participants missed the target (e.g., number 3), they were inclined to draw 2 or 4 rather than 1 or 5 as if they still had it in mind to actually hit target 3,

---

[1] Displacement is "a form of extrasensory perception shown by consistently obtaining information about a target at one or more remove, spatially or temporally, from the actual target designated for the trial" (Thalbourne, 2003, p. 31). In, for example, a card-symbol psi experiment, a participant might miss correct symbols and choose, instead, symbols that are only *similar* to the targets. For studies on displacement, see Crandall (1985), Crandall and Hite (1983), and Pratt and Foster (1950).

but could only come close (hence the term ‘good missing’). Good missing also re-occurred with target 2 (good misses being 1 and 3) and target 4 (good misses being 3 and 5). If target 1 was missed, the participants selected numbers with smaller differences (2 or 3) more frequently than numbers with larger differences (4 or 5). Such differences between target and misses were absent only when target 5 was missed.

Ertel (2005c) found that the total good-missing effect for targets 1 to 5 was small, only a 1.5% deviation from MCE was observed, but it was significant ($p = 1.00 \times 10^{-5}$). The good-missing phenomenon was conjectured to be a compromise between conflicting subconscious tendencies of hope of psi success and fear of such success (approach vs. avoidance).

Delayed Target Selection

In earlier studies, another apparent secondary psi phenomenon (indeed, another form of displacement) came to Ertel’s (2010a) attention, which was referred to as *delayed target selection*. Delays appeared occasionally with individual participants, but not in the standardization sample. For example, a participant called Ania, who showed considerable enthusiasm in the Ball Test, keenly wanted to achieve a significant hit score. She completed 96 runs alone (5,760 trials), but failed to produce a significant primary hit rate. However, she achieved a very significant level of secondary (‘delayed’) hitting ($p = .009$) when Ertel (2010a) calculated hits as defined by numbers drawn after being called (i.e., with a one-trial delay). Ania did not have delayed ‘hit’ scores with more extreme displacements—that is, displaced by 2, 3, 4, 5, or 6 trials (Ertel, 2010a, p. 346).

Delayed hits were also observed with three psi-gifted participants, but only when they were tested by skeptical experimenters who made them feel tense and nervous (Ertel, 2013).

Questions

The present study sought answers to the following three questions:

1. Is there evidence that *good-missing* (a form of displacement) is a form of unintentional psi avoidance?

2. Is there evidence of *delayed target selection* (a form of displacement)?

3. Do participants' responses and the psi process differ in accordance with *called* ball numbers and *selected* ball numbers?

The above analyses comprise the planned analyses of the present study. We were also obligated to investigate whether there were normal (i.e., non-paranormal) explanations for the alleged psi effects (such normal effects including sensory leakage and mnemonic aids). Tests of these possible artifacts comprise the post hoc analyses of the present study.

METHOD

Participants

The sample in the study by Storm et al. (2013) was mostly comprised of students from Deakin University, Melbourne, Victoria. Students did not participate as part of their course requirement, or to gain course credit. Participants were 'unselected' with the aim of recruiting an even number of sheep and goats. A total of 82 participants were laboratory tested. Mean age for the sample (N = 79; originally N = 82, but three participants did not give their age) was 29 years (SD = 13 years). Fifty-four percent of the sample were females (n = 44; males: n = 38, 46%).

Apparatus

An opaque bag (30 cm × 40 cm) containing 50 table tennis balls. One of five digits (1, 2, 3, 4, or 5) is written on each ball a number of times, evenly spread on the ball's surface. There are 10 balls bearing each digit (Total: 50 balls).

Procedure

All participants (N = 82) read an Instruction Sheet, and then signed a Consent Form if they decided to continue with the experiment. They completed various questionnaires the results of which are reported in Storm et al. (2013). They then read an opinionated communication—a modified version of the one used by Silvia (2005). The communication was presented surreptitiously as basic information about participating in a psi study:

> "<u>This short communication was written by a university professor</u>: This Ball Test has been developed in Germany and I claim that it is the best of all procedures that have hitherto been applied in parapsychology. I am utterly convinced that psi exists and that participants cannot avoid letting their psi power come to the fore

> when they make number predictions and then select the numbered balls. Every person, I claim, is expected to display such power. *I know I have persuaded you about this. I know you agree with my opinion. In fact, you're really forced to agree because university students can't have differing opinions on this issue*."

The additional italicized sentences were intended to function as the threatening elements; they were not italicized in the study. In the control condition, participants received the same communication without the italicized sentences. Allocation to control or treatment group was random.

Each participant completed up to four runs of 60 trials in one laboratory session, though many did not complete four runs. According to Deakin University's Human Research Ethics Committee policy, participants must be given the option to withdraw at any stage. Accordingly, the number of completed runs varied amongst participants. Some participants did not have time for four runs; others got bored. These withdrawals could not be considered feedback-based optional stopping because feedback was only given to participants after their last run was completed by which time they had already made it clear that they did not want to start another run.

The total testing time for 4 × 60 trials = 240 trials, spread over one session, amounted to approximately one hour. MCE for a run of 60 trials is 12 hits because with five numbers of which participants guess one number, 12 hits (i.e., 60/5) will be drawn on average. For 240 trials, MCE = 48 hits (i.e., 360/5). (For a more detailed account of the procedures, see Storm et al., 2013.)

RESULTS

Descriptive Statistics

The Ball Test. There were a total of 12,016 trials of which there were 2,531 hits (21.06%). A total of 82 participants completed 224 runs of 60 trials each. Of the 81 participants who completed at least one run, 28 participants (35%) completed 4 runs; 11 participants (13%) completed 3 runs; 13 participants (16%) completed 2 runs, and 29 participants (36%) completed 1 run. One participant did not complete the first run (only 16 trials were attempted).

As a preliminary test on participation rates for the control (n = 42) and reactance (n = 40) groups, we found that the fewest number of runs were completed by the reactants (91 runs), whereas the control group completed 109 runs. However, the difference was not significant, $\chi^2(1, N = 82) = 4.86$, $p = .183$ (two-tailed).

Planned Analyses

The following analyses were conducted to answer three crucial questions pertaining to patterns in number calling and ball-selection behaviour that participants may adopt in the Ball Test. These behaviours relate to possible psi avoidance or displacement effects (so-called *good missing* and *delayed target selection*).

1. Is there evidence that good-missing is a form of unintentional psi avoidance? The results of analysis on our data regarding good misses are shown in Table 1. Table 1 lists numbers called (i.e., 'guessed'), which are numbers 1 to 5 in columns, and numbers drawn (i.e., selected) which are numbers 1 to 5 in rows. Hit counts are given in bold text (in rows A to E). The good-misses hypothesis assumes that frequencies of selecting wrong numbers with smaller differences from the targets (row G) would be more frequent than selecting wrong numbers with larger differences from the targets (row H) so that the differences in row I would be positive, at least for columns 1 to 4. However, the differences are not consistently positive. The good-misses effect does not replicate in our database.

The 'good misses' phenomenon, whenever it is observed, should be regarded as a secondary psi effect, but of course only in cases where it is determined to be a psi effect based on significance testing as we tried to show. Its occurrence would strongly confirm the presence of non-intentional psi. Though failure to find such an effect does not support at least three theories that posit a form of psi that functions as a result of unconscious (non-intentional) causes (see Storm, 2013, for a review), the absence of non-intentional psi does not, and cannot, disconfirm the presence of intentional psi. Psi should be observed in the first place as a 'primary' effect (i.e., as surpluses of 'exact' hits, not by "near-hits"). For the time being, the conditions under which *good missing* occurs are not understood, particularly given that the effect is still only hypothesized to be a psi effect.

2. Is there evidence of delayed target selection? After our planned tests and analyses were conducted for Storm et al. (2013), we wondered if 'delayed hits' were observable that might indicate whether psi was unintentionally avoided in another way? Given that Ertel's (2010a) findings were for 'star subjects', we could not be sure if the delayed hits effect would replicate in our sample (N = 82), as participants were 'unselected'. Nevertheless, we present our results in Figure 1. The participants' ball selections (total of all participants' trails under reactance and control condition was N = 12,016) following their calls with no delays (delay = 0—i.e., 'on-trial') shows the largest hit percentage, P = 21.06% (p = .002), where P_{MCE} = 20%. Figure 1 also shows stepwise delayed shifts from –10 up to +25.

Table 1

Numbers Called and Numbers Drawn with Special Focus on "Good Misses" and "Other Misses"

			Number Called/Guessed					Total
	Row	Number	1	2	3	4	5	(SumR)
Number Drawn	A	1	**494**	595	581	459	350	2479
	B	2	422	**582**	555	463	355	2377
	C	3	396	536	**583**	463	349	2327
	D	4	417	566	603	**486**	402	2474
	E	5	456	551	564	401	**387**	2359
	F	Sum1	2185	2830	2886	2272	1843	12016
	G	Good	818	1131	1158	864	751	4722
	H	Other	873	1117	1145	922	705	4762
	I	**Diff**	**-55**	**14**	**13**	**-58**	**46**	**-40**
	J	Sum2	1691	2248	2303	1786	1456	9484

Note: Sum1 = Hits plus misses (i.e., A + B + C + D + E); 'Diff' = 'Good' misses minus 'Other' misses (i.e., G – H); Sum2 = 'Good' misses plus 'Other' misses (G + H).

At this point, we are mainly interested in the hit percentage of trial +1 (one trial delay) which might be expected to lie above MCE and to be larger than the majority of other values obtainable with delays of +2, +3, . . ., +25. Effects of psi at trial = 0 might be inhibited occasionally to some extent and become effective only with one trial delay. But this, apparently, did not occur (in fact, it is below MCE). The fact that all shifts of number calls to later and earlier number selections resulted in lower ‘hit’ percentages, compared with the delay = 0 condition (i.e., ‘on-trial’), demonstrates that a significant conscious effort is exerted on selecting balls immediately after calling the number for that ball. That is, psi effects are mainly ‘on-trial’ (i.e., delay = 0), and there is no evidence of ‘off-trial’ effects (i.e., a displacement effect we refer to as delayed target selection).

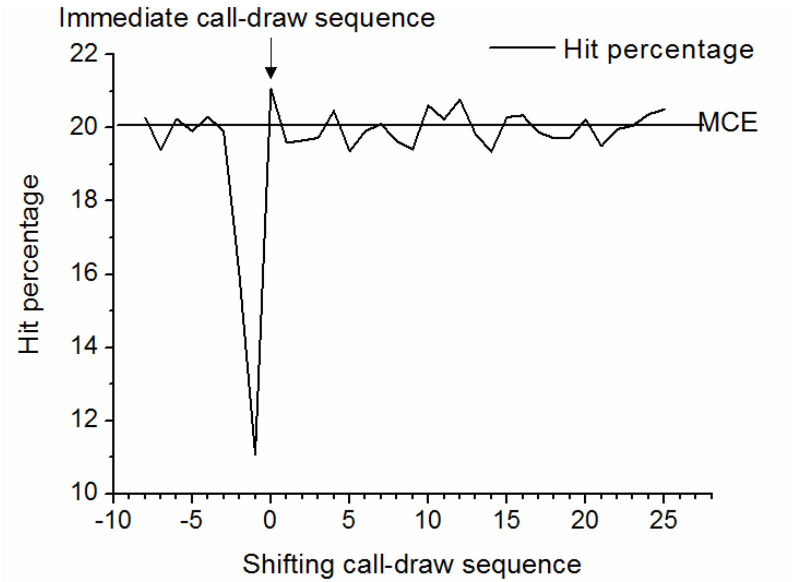

Figure 1. Hit percentages for delay = 0 (selections immediately following calls without delay), as well as for 25 delayed trials (delay = 1, 2, 3, etc.) and 8 “anticipated” trials (“delay” = –1, –2, etc.).

Figure 1 also shows a considerable drop in the curve, reaching its nadir at –1. This is a call avoidance effect—participants tend to avoid calling numbers they have just drawn so that extremely low frequencies of call-draw matching can be expected. The plunge in the curve ranges between –3 and 0, showing that participants avoid calling numbers they have just drawn for roughly three trials. This observation, however, has no bearing on the actual presence (or not) of the primary psi effect (i.e., 'on-trial' hitting).

3. Do participants' responses and the psi process differ in accordance with called ball numbers and selected ball numbers? In Table 1, 'Sum1' of the columns show considerable variance. Participants seem to prefer to call numbers 2 and 3 (number 2 is called 2,830 times; number 3 is called 2,886 times; but number 5 is only called 1,843 times). The row totals A, B, C, D, and E (listed in the 'SumR' column), which represent frequencies of drawn numbers, are much more balanced. This is as might be expected because chance is the main 'determinant' of selection frequencies.

To further evaluate the data in Table 1, we note that one might calculate the distribution of frequencies using the chi-square test. An overall chi-square test is not powerful enough for testing the significance of hit frequencies for which only the diagonal frequencies in Table 1, containing hit frequencies, are relevant. However, an exploratory investigation of 25 chi-square values, based on 5 × 5 frequencies in Table 1, might be of some value. Thus, a total of 25 deviations from chance (i.e., column and row sums) are considered. They need not be summed to obtain a chi-square of the total.

Chi-square values based on deviations of chi-square expectancy are given in Figure 2; deviations are also pictorially represented as grey circles (positive chi-square values) and white circles (negative chi-square values). Note that circle size indicates strength of the chi-square value because the latter determines the radius of the circle.

Figure 2 shows hit surpluses for all numbers 1 to 5 (see the grey circles in the diagonal). An unexpected large negative deviation from expectancy is found with number 5 calls (see the large white circle), following which number 4 is rarely drawn. More experiments are necessary to find out whether this deviation is merely chance or a secondary psi effect.

The same visualization technique has been applied on the sequence of calls. One might wish to know which numbers are often or rarely called after calling a given number. Figure 3 shows that participants tend to avoid calling a number repeatedly (see the large white circles in the diagonal).

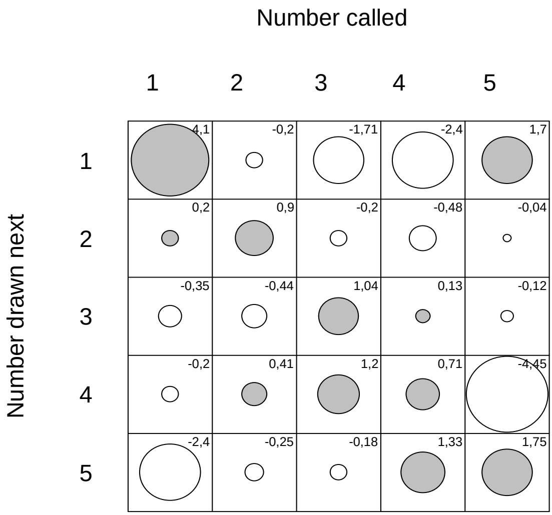

Figure 2. Chi-square values for 25 cells of a matrix formed by calls 1 to 5 as columns and draws 1 to 5 as rows.

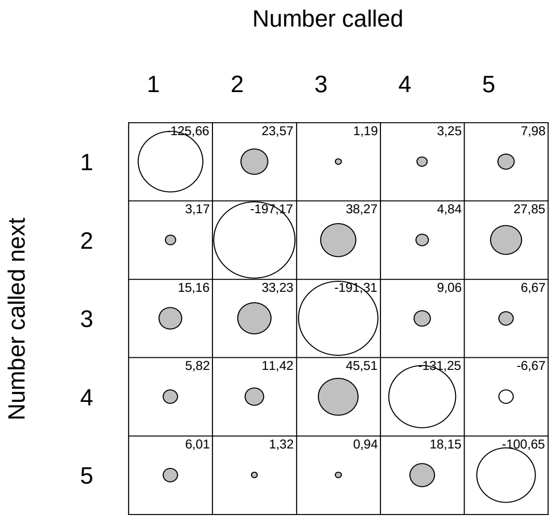

Figure 3. Chi-square values for 25 cells of a matrix formed by calls 1 to 5 as columns and immediately succeeding calls 1 to 5 as rows.

Figure 4 shows another type of avoidance calling. Participants avoid calling a number which they have just *drawn* (i.e., just selected).

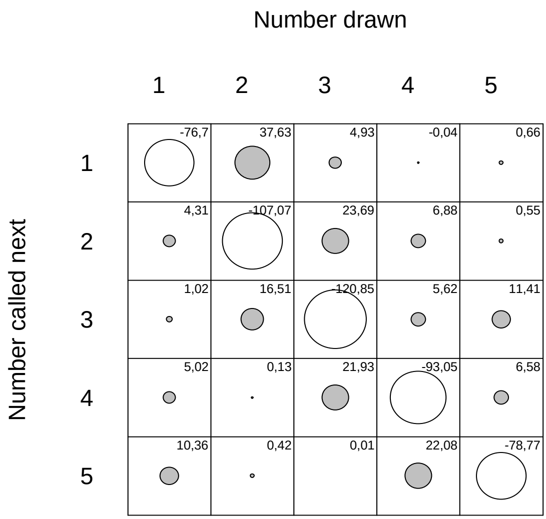

Figure 4. Chi-square values for 25 cells of a matrix formed by drawn 1 to 5 as columns and immediately succeeding calls 1 to 5 as rows.

A comparison of Figures 3 and 4 shows that the avoidance of calling just *drawn* (just selected) numbers is weaker than the avoidance of calling *just-called* numbers.

The last matrix (see Figure 5) shows deviations from expectancy for draw sequences. No regularity is visible and deviations from expectancy are small, which is expected. Deviations from expectancy of subsequent draws are particularly small if number 5 is drawn. Without further tests one cannot tell whether this result is attributable to chance.

Visualized deviations from expectancy for number calls, followed by immediate number selections, which are of prime interest in the Ball Test, might give some researchers confidence in test outcomes, if *p*-value statistics do not suffice. The visualization technique might also reveal, with call-call, draw-call, and draw-draw matrices, unexpected deviations from chance, and these might encourage investigations into alternative forms of psi phenomena.

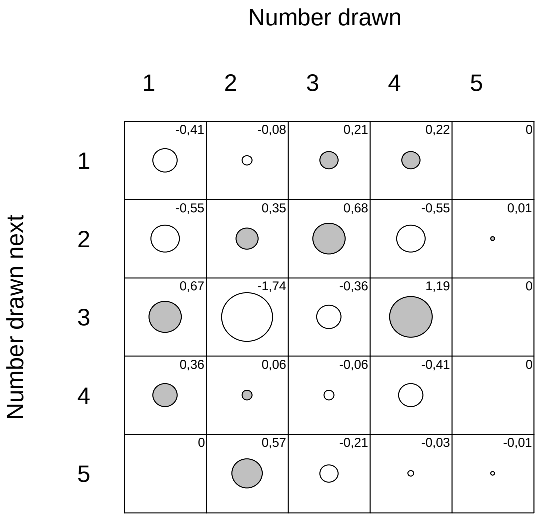

Figure 5. Chi-square values for 25 cells of a matrix formed by draws 1 to 5 as columns and immediately succeeding draws 1 to 5 as rows.

Post Hoc Analysis

We conducted further analyses to determine if the statistical support we obtained for the psi hypothesis may be undermined by alternative hypotheses such as sensory leakage and/or mnemonic aid.

Sensory Leakage. In earlier psi studies, Ertel (2005c) argued that if there was any kind of sensory leakage (i.e., tactile and temperature cues), significant hit rates might be artifacts. However, if such artifacts did exist, then hit rates would not be high on participants' first trials (i.e., in the proximity of the *start* of the first run). A more or less gradually ascending learning curve would be revealed because tactile cueing increases as more and more heat transfers from warm hands to certain numbered balls that have previously been held and drawn out of the bag. However, Ertel did not find such a curve in his standardization sample.

Visual inspection of Figure 6 shows a slightly ascending slope, but this is far from significant (p = .51), thus giving us no indication of a hitting artifact.

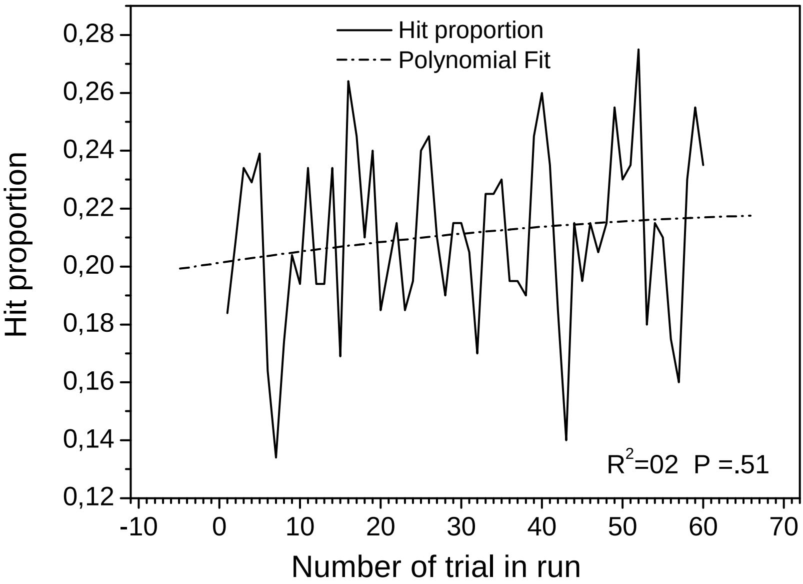

Figure 6. Hit proportions for each trial of one run, from trial 1 to 60, summed across participants and runs.

Just drawn numbers (JDNs) as mnemonic aids. Another test of a possible artifact conducted by Ertel (2005c) was aimed at determining whether participants were more successful with calls of numbers that they had just selected (i.e., just drawn numbers, JDNs). If they called a JDN, and if they then put the ball back in the bag in some corner that they might remember, and if they jumbled the balls poorly, they might have a better opportunity to hit this number compared with numbers that were not-JDNs (Ertel called this hypothetical artifact “memory leakage”). Note that memory leakage (i.e., mnemonic aids) may not necessarily occur intentionally or consciously. Ertel’s test with N = 238 (his standardization sample) gave no indication that hit counts were larger due to calling JDNs compared to calling not-JDNs. Also, preliminary results by Ertel (2005a) with psi-gifted participants shows that hit rates under conditions of free number selection (the ‘standard’ test condition) does not differ, and nor are they larger than hit rates under conditions where targeting is assigned (i.e., participants are told to draw the numbers in prescribed sequences). Three assigned sequences of calls were used: (i) the same number repeated (e.g., 1, 1, 1, 1, 1, etc.); (ii) the assigned number sequence in ascending order (1, 2, 3, 4, 5, repeated), and (iii) the assigned number sequence in descending order (5, 4, 3, 2, 1, repeated). For all three conditions, psi hit rates were slightly lower than in the standard (free-choice) condition. A psi-detrimental effect by an assumed avoidance of equal number repetition does not show up.

In the present study, we did not expect to find a JDN-effect in our data either. As expected, the difference between calls of JDNs and not-JDNs (21.56% vs. 21.00%) is small and not significant (see Table 2). Note too that hit rate for JDNs is not significant, whereas it is highly significant for not-JDNs. Our results show that the suspicion that some number of participants might have used mnemonic aids, evident by a significantly high number of JDNs, is not warranted.

Table 2
Hit Proportions for Calls of JDNs and Not-JDNs

Calls	Trials	Hits	Hit Proportion (%)	*p* (one-tailed)
JDN	1331	287	21.56	.082
Not-JDN	10685	2244	21.00	.005
Calls Total	12016	2531	21.06	.002

Note: JDN = Just Drawn Numbers (where Just Drawn = Just Selected)

DISCUSSION

In our initial study (Storm et al., 2013), we sought insight into the psi performance differences between sheep and goats after a reactance manipulation using Ertel's (2005b,c) Ball Selection Test as a means by which psi effects might be elicited. In conventional parapsychology, sheep are encouraged to do their 'best', but goats are not encouraged to do their 'worst'. Following the principles of reactance theory (Brehm, 1966; Brehm & Brehm, 1981), we manipulated reactance with an opinionated communication. We showed that reactance does have an effect on psi performance, with the treatment clearly showing adverse effects on psi hitting (these effects parallel those found by Silvia, 2005, 2006, though his studies are psychological, not parapsychological). We also endeavoured to show that the reactance treatment would have more of a psi-scoring effect on goats than on sheep, and we found that scoring was in the direction hypothesized. We also found scoring differences between control sheep and reactant goats with reactant goats scoring significantly lower than control sheep.

For the present study, our planned and post hoc analyses were mainly focused on the validity of the Ball Test, but these analyses also revealed some important points about the dynamics of the Ball Test. Two tests revealed there were no secondary psi effects (i.e., ‘good misses’ and ‘delayed target selection’). In the case of good misses, the null effect indicates that there was no non-intentional psi, and there was only intentional psi. In the absence of any evidence of delayed target selection (i.e., delayed hits), we can also say that participants were not displacing to other targets.

We also found post hoc that participants tended to follow what might be termed superstitious behaviour involving outright avoidance of numbers if those numbers had been recently called or were hits. These interesting effects were discovered as patterns in our data that are best observed visually in 5 × 5 grid-forms (Figures 2 to 5). Participants avoid calling numbers if they have just called them or just selected them from the bag—i.e., hit rates did not show significant increases in hit rates across runs, so that the alternative hypothesis that normal sensory leakages (i.e., tactile or temperature cues) were responsible for overall hit surpluses was not supported. Of course, these avoidance effects are psychological, indicating that most individuals do not understand that trials are independent of each other. That is, most participants, if they were to conceptualize it, would have the idea that a number series must be random, but they mistakenly assume that this means that numbers must, for the most part, always change from trial to trial so that, in a random sequence, a just-called or just-selected number is not likely to re-occur immediately. Such biasing effects (in this case, subjective probability) are well known (Esgate & Groome, 2001, Mischo, 1975, Pratt & Soal, 1952; Rhine, Smith, & Woodruff, 1938).

The question arises whether biased calling of numbers based on this misconception has any effect on psi as measured by hit scores. One might surmise that deviations of subjective randomness from true randomness might lower hit scores since not selecting just-called numbers or just-drawn numbers (JDNs) might imply avoidance and a concomitant restriction on psi. In the present study, there was again no evidence of a hitting artifact—if sensory leakage had been effected, a curve with a significant incline would have resulted (see Figure 6). Such an incline was not evident. Ultimately, however, a solution to the enduring problem of selection bias requires further research, preferably with very large samples.

ACKNOWLEDGEMENTS

Preparation of this article was supported by a grant from the Bial Foundation. The authors thank Fiona Permezel for her assistance in recruiting participants.

REFERENCES

Alcock, J. E. (1981). *Parapsychology: Science or magic?* New York: Pergamon.

Beischel, J., & Rock, A. J. (2009). Addressing the survival vs. psi debate through process-focused mediumship research. *Journal of Parapsychology, 73,* 71-90.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social*, *100*, 407-425.

Brehm, J. W. (1966). *A theory of psychological reactance*. New York: Academic Press.

Brehm, S. S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control*. New York: Academic.

Crandall, J. E. (1985). Effects of favourable and unfavourable conditions on the psi-missing displacement effect. *Journal of the American Society for Psychical Research*, *79*, 27-38.

Crandall, J. E., & Hite, D. D. (1983). Psi-missing and displacement: Evidence for improperly focused psi? *Journal of the American Society for Psychical Research*, *77*, 209-228.

Ertel, S. (2005a). Are ESP test results stochastic artifacts? Brugger & Taylor's claims under scrutiny. *Journal of Consciousness Studies*. *12*(3), 61-80.

Ertel, S. (2005b). Psi test feats achieved alone at home: Do they disappear under lab control? *Australian Journal of Parapsychology*, *5*, 149-164.

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

Ertel, S. (2010a). On individual differences in extrasensory perception. In R. Rao (Ed.). *Yoga and parapsychology: Empirical research and theoretical essays* (pp. 331-350). Delhi, India: Motilal Banarasidass.

Ertel, S. (2010b). Psi in a skeptic's lab: A successful replication of Ertel's Ball Selection Test. *Journal of Scientific Exploration*, *14*(3), 461-478.

Ertel, S. (2013). Sensory leakage or psi effect? Probing the ball test. *Journal of Scientific Exploration*, *27*(3), 387-391.

Esgate, A. E., & Groome, D. (2001). Probability and coincidence. In R. Roberts & D. Groome (Eds.), *Parapsychology: The science of unusual experience* (pp. 19-34). London: Arnold.

Hansel, C. E. M. (1980). *ESP and parapsychology: A critical reevaluation.* Buffalo, NY: Prometheus Books.

Houran, J., & Brugger, P. (2000). The need for independent control sites: A methodological suggestion with special reference to haunting and poltergeist field research. *European Journal of Parapsychology*, *15*, 30-45.

Kraus, S. J. (1995). Attitudes and the prediction of behavior: A meta-analysis of the empirical literature. *Personality and Social Psychology Bulletin*, *21*, 58-75.

Lawrence, T. R. (1993). Gathering in the sheep and goats: A meta-analysis of forced-choice sheep-goat ESP studies, 1947-1993. *Proceedings of the 36th Annual Convention of the Parapsychological Association*, pp. 75-86.

Mischo, J. (1975). Parapsychische Phänomene im quantitativ-statistischen Modell: Teil III [Quantitative models of parapsychological phenomena: Part III]. *Zeitschrift für Parapsychologie*, *17*, 201-218.

Palmer, J. (1971). Scoring in ESP tests as a function of belief in ESP: Part I. The sheep-goat effect. *Journal of the American Society for Psychical Research*, *65*, 373-408.

Palmer, J. (1972). Scoring in ESP tests as a function of belief in ESP: Part II. Beyond the sheep-goat effect. *Journal of the American Society for Psychical Research*, *66*, 1-25.

Palmer, J. (1977). Attitudes and personality traits in experimental ESP research. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp. 175-201). New York: Van Nostrand Reinhold.

Pratt, J. G., & Foster, E. B. (1950). Displacement in ESP card tests in relation to hits and misses. *Journal of Parapsychology*, *14*, 37-52.

Pratt, J. G. & Soal, S.G. (1952). Some relations between call sequence and ESP performance , *Journal of Parapsychology*, *16*, 165-186.

Rhine, J. B., Smith, B. M. & Woodruff, J. L. (1938). Experiments bearing on the role of precognition hypothesis: II. The role of ESP in the shuffling of cards. *Journal of Parapsychology*, *2*, 119-131.

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.

Schmeidler, G. R. (1945). Separating the sheep from the goats. *Journal of the American Society for Psychical Research*, *39*, 47-49.

Schmeidler, G. R., & McConnell, R. A. (1973). *ESP and personality patterns.* Westport, CN: Greenwood.

Silvia, P. J. (2005). Deflecting reactance: The role of similarity in increasing compliance and reducing resistance. *Basic and Applied Social Psychology*, *27*, 227-284.

Silvia, P. J. (2006). Reactance and the dynamics of disagreement: Multiple paths from threatened freedom. *European Journal of Social Psychology*, *36*, 673-685.

Steinkamp, F. (2005). Forced-choice ESP experiments: Their past and their future. In M. A. Thalbourne & L. Storm (Eds.), *Parapsychology in the 21st century: Essays on the future of psychical research* (pp.124-163). Jefferson, NC: McFarland.

Storm, L. (2013). A comparative approach to the theory of psychopraxia. In S. Krippner, A. J. Rock, J. Beischel, & H. Friedman (Eds.), *Advances in parapsychological research 9* (pp. 82-96). Jefferson, NC: McFarland.

Storm, L., Ertel, S., & Rock, A. J. (2013). The sheep-goat effect as a matter of compliance vs. noncompliance: The effect of reactance in a forced-choice ball selection test. *Journal of Scientific Exploration*, *27*(3), 393-411.

Storm, L., & Thalbourne, M. A. (2005). The effect of a change in pro attitude on paranormal performance: A pilot study using naïve and sophisticated skeptics. *Journal of Scientific Exploration*, *19*, 11-29.

Storm, L., Tressoldi, P. E., & Di Risio, L. (2012). Meta-analyses of ESP studies 1987-2008: Assessing the success of the forced-choice design in parapsychology. *Journal of Parapsychology*, *76*, 243-273.

Thalbourne, M. A. (2003). *A glossary of terms used in parapsychology*. Charlottesville, VA: Puente.

Wright, R. A. (1986). Attitude change as a function of threat to attitudinal freedom and extent of agreement with a communicator. *European Journal of Social Psychology*, *16*, 43-50.

Brain and Cognition Centre
School of Psychology
University of Adelaide
South Australia 5005
AUSTRALIA

Email: lance.storm@adelaide.edu.au