FINAL REPORT

TO

FUNDAÇÃO BIAL

***GRANT* 66/04**

EXTRASENSORY PERCEPTION AND IMPLICIT SEQUENCE LEARNING IN A COMPUTER GUESSING TASK (NORMAL PARTICIPANTS)

John Palmer, Ph.D. and Peter Brugger, Ph.D.

Department of Neurology

University Hospital Zürich

TABLE OF CONTENTS

EXTENDED ABSTRACT

The purpose of this experiment was to explore implicit sequence learning (ISL), response biases, and extrasensory perception (ESP) in a computer guessing task. The participants (Ps) were 64 normal volunteers, 32 of whom described themselves as strong believers in the paranormal and 32 of whom described themselves as strong skeptics. The task consisted of guessing which of the numbers 1-4 would be selected by the computer for each trial. Ps made their responses by calling each guess out loud and simultaneously clicking the mouse to register the response. Each P completed 4 runs. During the 1st 3 runs, Ps received no feedback of the correctness of their guesses. In the 1st 2 of these runs, the target sequence (N = 81) reflected either pure repetition avoidance (RA) or pure counting (CO), which means a sequence of, say, 2,3,4,1,2,3,4,1,2…. The 101 targets for the 3rd run were random, except that every time P clicked the mouse when a computer address contained a 1 rather than a 0 (“1-state”), which occurred randomly 20% of the time, they would receive a target for the next trial that matched their personal response bias as calculated from the preceding 2 runs. In the last run, Ps received feedback that was presented supraliminally to half of the sample and subliminally to the other half. Within each of these groups, for half of the Ps, after the 1st 10 of the 100 trials the targets were their own previous responses with a lag of 10 (pro-bias targets). The other Ps received a target sequence diametrically opposite to their response bias, calculate from the 1st 2 runs (counter-bias targets). In between Runs 2 and 3, Ps completed a drawing task, a tolerance of ambiguity questionnaire, and a questionnaire measuring signs of temporal lobe dysfunction in the brain (LIMBEX). Meanwhile, the experimenter (E) calculated their response bias in the preceding 2 runs, considering the frequency both of individual targets and of pairs (or shifts). Following Run 4, Ps completed a rating scale concerning their orientation to the task and their estimates of the structure of the target sequences. One P who recognized the target bias in Run 4 during the run itself had to be replaced, because for her the learning was not implicit. Finally, Ps completed the Australian Sheep Goat Scale (ASGS) to verify their paranormal belief status.

Four sets of hypotheses were tested, 2 involving ESP, 1 involving ISL, and 1 involving response bias. The first ESP hypothesis was derived from May’s decision augmentation theory. It was expected that in Run 3, Ps would click the mouse more frequently than chance when the computer was in the 1-state, and that this result would be more prevalent for believers than skeptics. The overall mean was suggestively positive. The mean for believers was significant and significantly higher than that of the skeptics.

The second ESP hypothesis concerned the anomalous anticipation effect (AAE) found in 2 previous ISL experiments. Based on these data, it was predicted that skeptics would correctly anticipate the nature of the bias (pro or counter) in Run 4 by manifesting the corresponding bias in Run 3, and that believers would anticipate incorrectly. The hypothesis was confirmed, but only if the Run 4 targets were subliminal.

Regarding ISL, it was predicted that Ps would show a greater increase in correct guesses from the 1st to the 2nd half of Run 4 if the targets were pro-bias rather than counter-bias, and that this ISL effect would be stronger for believers than skeptics and with subliminal than with supraliminal feedback. Contrary to expectation, ISL was significant for skeptics only, with the nature of the bias and the subliminality of the feedback having no effect. An interaction was also found revealing that the skeptics did not evidence ISL in the pro-bias / supraliminal condition, perhaps because of a ceiling effect: scoring was also quite high in the 1st half of Run 4 in this condition. Additional analyses indicated that the superiority of ISL for skeptics was mediated by the LIMBEX scale, on which believers scored dramatically higher than skeptics. These analyses indicated that ISL occurred regardless of condition among low to mid-range scorers on LIMBEX, but that a reversal of ISL (a decline) was manifested by high LIMBEX Ps. In other words, temporal lobe instability seems to prevent ISL.

It was predicted that Ps would score above chance in the runs with RA and CO targets, because both target sequences contained no repeats, and this was expected to match a corresponding response bias of repetition avoidance in the Ps. Scoring was in fact above chance in both of these runs but was significant only in the RA run. Scoring in the random trials of Run 3 was very slightly below chance. Although all Ps demonstrated marked RA, contrary to expectation the deficit was only slightly greater for believers than for skeptics. In the RA run, the degree of RA was correlated negatively and significantly with hits for believers, but surprisingly this correlation reversed slightly for skeptics.

The fact that in several parts of the experiment skeptics outscored believers, in contrast to the standard sheep-goat effect, was speculatively attributed to discomfort on the part of believers because E was perceived to be a skeptic. However, data from a handful of Ps asked to estimate the belief status of E did not support this conjecture.

Finally, exploratory analyses of the personality and rating scale data revealed that believers predicted higher scores than skeptics, as well as greater bias in the targets sequences. Both groups grossly underestimated the number of trials in each run. Estimates of success significantly predicted actual hit rates only in the control condition (counter-bias targets) of Run 4. Ps high on tolerance for ambiguity scored significantly higher in the random and total trials of Run 3 than did low tolerance Ps.

INTRODUCTION

This experiment, undertaken with support from the Bial Foundation, was intended to increase our understanding of two cognitive processes, implicit sequence learning (ISL) and extrasensory perception (ESP). As will become evident during the course of this background section, these processes are closely related both conceptually and operationally, which allows us to test them simultaneously in a single experiment. The two processes are also related historically in the genesis of the proposed research, in that the ESP hypotheses and design features grew out of 2 ISL experiments conducted in applicant Brugger's laboratory. As we believe this time sequence should be reflected in this background section, we will begin with a discussion of past research and theory of ISL. We will then discuss in some detail the particular ISL experiments that inspired one ESP aspect of the proposal. Following some further comments on the relationship between ISL and ESP, we will conclude with a discussion of the ESP research and theory of specific relevance to the proposed experiment.

IMPLICIT SEQUENCE LEARNING

A common methodological paradigm in psychological studies of learning is for experimental participants (Ps), both humans and animals, to be presented with discriminable stimuli to which they make behavioral responses. Following each response, they are presented with a positive reinforcement if the response is correct and/or a negative reinforcement if the response is incorrect. P uses the information component of the reinforcement, which can be called feedback, to learn the required discrimination. Over the course of repeated trials, the percentage of correct responses increases, indicating learning.

At least for humans, it is traditionally assumed that Ps become consciously aware of the contingency they have learned and that guides their responding. However, such awareness is not logically necessary for learning to take place and indeed is not always present. Research on implicit sequence learning (ISL) has demonstrated empirically the ability of Ps to accurately discriminate among different types of stimuli without being able to report the principle involved. Typically they are given a finite number of symbol alternatives (for example, the numbers 1, 2, 3, and 4) and asked to guess which of these symbols will be shown to them next. After each guess, the correct symbol is revealed to them. The sequence of target symbols is not random, but rather biased in a systematic fashion. For example, there might be an excess of 2s in the sequence. ISL would be indicated by a tendency of P's responses during the course of the session to conform to the nature of the target bias. In our example, this would be to guess 2 more frequently than the other alternatives.

As early as 50 years ago, Hake and Hyman (1953) convincingly showed that the patterning in a binary series of alternatives can be exploited by human Ps for predictive guesses as to which of two alternatives will be presented next. Later experiments on "probability learning" (see Jones, 1971, for an overview) have confirmed this finding and demonstrated that learning from feedback occurs in the absence of P's verbal awareness. Such unconscious learning of the statistical properties of event sequences can also be found in paradigms of "artificial grammar learning" (Reber, 1967). Here, P is confronted with brief strings of symbols that are arranged according to some probabilistic rule ("grammar"). After a prolonged training phase P will be able to differentiate between test strings that do and do not contain the critical regularity. Importantly, despite this above-chance recognition, the nature of the regularity cannot be overtly, i.e. verbally, described. The processes of ISL as revealed by probability and guessing tasks and in paradigms of artificial grammar learning are highly relevant for the understanding of language acquisition (Saffran et al., 1996; Seidenberg, 1997).

ISL and ESP

As noted above, ISL and ESP are related both conceptually and operationally. Both concern the acquisition of information in ways P is not fully cognizant of. Having Ps guess a sequence of symbols with immediate feedback of the accuracy of each guess is a common procedure for testing both ESP and ISL. The only important difference between these two procedures is that in ESP research the target sequences are (or are supposed to be) random and in ISL research they are biased. Belief in the paranormal, which was discussed above as a likely correlate of ISL, is also one of the most reliable correlates of success in ESP tasks (Lawrence, 1993; Palmer, 1971; Schmeidler & McConnell, 1958/1973).

Each of these processes could be proposed as an explanation of some data gathered in the other paradigm and originally explained by the other process. In particular, applicant Brugger and Taylor (2003) have proposed that certain ESP experiments in which target randomisation was suboptimal could be better explained by ISL. That suboptimal randomisation can lead to pseudo-ESP effects was demonstrated in a study where Ps were asked to indicate if they were being stared at by a remote observer, a popular and successful ESP testing paradigm (Schmidt, Schneider, Utts, & Walach, 2004). Colwell, Schroder, & Sladen (2000) found significant results only when their Ps were given trial-by-trial feedback of their success. However, it was subsequently discovered that the target sequence had too many alternations between staring and non-staring trials, a common feature of non-random calling patterns generated by normal humans. When the target sequences were made adequately random, the effect vanished. It should be noted, however, that a meta-analysis indicates that whatever randomisation deficiencies there may be in some of the staring studies cannot account for the overall significance of the effect (Schmidt et al., 2004).

ISL and Belief in the Paranormal

People differ in their capacity to demonstrate ISL, and one determinant of this difference is belief in the paranormal. With the exception of an experiment to be described later (Krummenacher, 2003), the evidence for the facilitating role of belief on ISL is indirect. Several experiments have shown that believers are more likely than skeptics to "detect" patterns in stimulus conglomerations that are in fact purely random, such as photographs totally degraded by electronic noise (Blackmore & Moore, 1994; Brugger et al., 1993), and they are also more likely to attribute meaning to artificial coincidences (Brugger & Graves, 1997; Mohr, Graves, Gianotti, Pizzagalli, & Brugger, 2001). The relevance of such data to ISL arises from the plausible although not demonstrated assumption that *attribution* of patterns to stimuli is positively associated with *detection* of real patterns in stimuli. Both of these mechanisms imply a *sensitivity* to patterns, as well as a motivation to seek them out. One would expect such sensitivity and motivation to also be applied to the biases in linear sequences of symbols that must be identified for success in ISL tasks.

RESPONSE BIAS

Closely related to ISL is the topic of response bias. It is well known that human experimental Ps are unable to generate or call out truly random sequences of symbols even when explicitly asked to do so (Brugger, 1997; Tune, 1964). Instead they manifest non-random response biases, a common example of which, at least among normal adults, is repetition avoidance. Response biases are relevant to ISL because what Ps are asked to learn in an ISL task is a comparable kind of bias applied to a target sequence . It is also true that Ps bring to an ISL task response biases of a particular type, and these serve as the background against which the learning takes place. Thus, matching response and target biases should lead to some enhanced success even at the beginning of the ISL task. However, despite this higher baseline we would expect (barring a ceiling effect) those Ps whose response biases match the target biases to show the greatest increase in manifestation of the bias during the course of the run. In other words, ISL should be enhanced if the target biases match the response biases Ps bring with them to the session. The reason, we expect, is that Ps will find it easier to detect biases with which they are already familiar. We plan to test this hypothesis in our proposed experiment.

ESP and Target / Response Bias: Global Responding

It is reasonable to hypothesize that in ISL and ESP experiments Ps respond primarily to global patterns in the target sequences rather than to individual targets. This makes a great deal of sense from the parapsychological perspective. In a forced-choice ESP task (for example, card guessing), the assessment of patterns requires less transfer of information than does assessment of individual targets, even if one assumes that ESP is operative only on a minority of the trials. Thus the ability to succeed in the task may be restricted to such global perceptions in all but the most talented Ps. If such individuals implicitly recognize this fact, they might well choose to attend to, or look for global biases for practical reasons.

Of course, in most ESP experiments the patterning in the target sequences is severely limited by the fact that they are selected to be random. However, there have been two ESP studies we are aware of in which target biases were deliberately introduced into a card-guessing task. This use of so-called "unbalanced decks" was pioneered by Child and Kelly (1973). The decks of cards consisted of fixed but unequal numbers of the five standard ESP symbols: star, circle cross, star, and waves. Specifically, the frequencies in each deck of 25 cards were 9, 7, 5, 3, and 1, with the number of times each symbol appeared counterbalanced across decks. P was a psychic by the name of Lalsingh Harribance (L. H.), who had shown success in previous ESP tests (e.g., Roll & Klein, 1972). Overall, L. H. obtained 21.1% hits (20% expected by chance) over 9,000 trials ($p < .01$). His hit rates on the five individual targets were about the same. However, in line with the global perception hypothesis, he called symbols appearing 7 times in the decks significantly more often than those appearing 3 times, and this was true even when only the misses were analysed. However, the 9-1 split yielded no such response preference. About 25 years later, applicant Palmer attempted to replicate this finding during a visit by L. H. to the Rhine Research Center (Palmer, 1998). However, no significant results were obtained. L.H. had not been involved in controlled testing of this type for a number of years, and during his visit he put a great deal of psychological pressure on himself to perform well. This may explain the fact that during other experiments conducted during this visit, evidence of psi-missing was obtained, which Palmer has long considered to be a symptom of such stress.

Palmer (1996) also found evidence of global responding in a reanalysis of the raw data from two highly successful ESP experiments using a random number generator (RNG) with trial-by-trial feedback (Schmidt, 1969). Ps, who were selected by preliminary testing as having some ability with the task, had to guess which of 4 numbers the RNG would choose as the target for the next trial (precognition). Feedback was presented as the lighting of 1 of 4 colored lamps on the machine console. It was found that when excesses of single targets or pairs of targets were present in the target sequences, Ps increased their frequency of responding with those particular target alternatives or pairs. In other words they seemed to adjust their response biases to biases in the target sequences. This patterning accounted for some, but not all, of the ESP success. The effect occurred despite the fact that the target sequences proved to adequately random, both when the target sequences were evaluated in mass and when summed over blocks of 25 to 200 trials each to check for possible localized nonrandomness. This paradoxical finding makes sense if one considers that even random sequences contain subsequences that are "biased". Their frequency is indicated by the type-1 error rate; for example, 5% of subsequences will be non-random at the $p = .05$ level. It would appear that Ps were using ESP to predict when these biases would occur and adjust their responding accordingly.

Palmer (1997) later found that 4 of the 5 Ps had a strong response bias to give as their guess for trial t+1 the target for trial t if the latter was a hit. Although this bias might account for some of the bias-matching effect discussed in the preceding paragraph, it was not responsible for the significant ESP scores, at least not by itself.

Response Bias and Belief in the Paranormal

People differ among themselves in both the degree and type of the biases they exhibit. Among variables shown to positively affect one or both of these parameters are age, schizophrenia, and perceived meaningfulness of everyday coincidences (Brugger, 1997). However, we will focus here, once again, on belief in the paranormal.

Several experiments by applicant Brugger have shown that believers in the paranormal are more likely to exhibit repetition avoidance than skeptics (Brugger & Baumann, 1994; Brugger, Landis, & Regard, 1990; Brugger, Regard, Landis, & Graves, 1995; Brugger, Regard, Landis, Krebs, & Niederberger, 1994). Successful independent replications have been reported by Bressan (2002) and by Musch and Ehrenberg (2002). (Three other independent confirmations have also been reported, but not in published form: Blackmore and Kahn; Wilson and Sturrock; Schienle, based on reanalysis of Schienle, Vaitl, and Stark, 1996.) On the other hand, failures to find the effect have been reported by Blackmore & Troscianko (1985), Broughton (1994), and Houtkooper and Haraldsson (1997), although the latter study showed a strong trend in favor of the hypothesis (p = .072, one-tailed). Finally, Lawrence (1990/91) found a significant reversal of the effect, with believers showing more repetitions (rather than repetition avoidance), but the fact that he explicitly told his Ps that one should expect repetitions may have caused his believers to overcompensate (Brugger, 1997). In conclusion, there is a strong trend in support of the hypothesis that believers exhibit more repetition avoidance than skeptics.

DECISION AUGMENTATION THEORY

Edwin May and colleagues have proposed a model called Decision Augmentation Theory (DAT) that is intended to explain micro-psychokinesis (PK) effects as in fact due to ESP (May, Spottiswoode, Utts, & James, 1995a; May, Utts, & Spottiswoode, 1995b). Although controversial (Dobyns & Nelson, 1998), DAT appears to account successfully for at least some of the relevant data. Most of the data to which the theory has addressed itself is of the RNG type. What typically occurs in such experiments is that a "hardware" RNG converts electronic noise into binary numbers, the distribution of which should follow the stochastic laws of chance. Most notably, the numbers of each digit generated should be exactly equal in an infinite sequence and approximately equal in finite sequences, the approximation improving as the length of the finite sequences increases. The traditional interpretation of data from these experiments is that P uses PK to bias the noise source such that it yields a significantly unequal distribution of the two binary numbers. DAT rejects these "force" models. It maintains instead that P intersects the target stream being produced by the RNG at that point at which a "biased" subsequence is about to appear, i.e., a sequence with an excess of one of the binary digits. Recall from the discussion above that such subsequences will occasionally occur even in a true random sequence. The anomalous mechanism P uses to detect (or predict) the point at which the target stream should be intercepted is what we call ESP, or more precisely, precognition. In DAT experiments, the decision as to when to intersect the target stream is made by P pressing a button to initiate the test or run. In an important sense the button press *is* the ESP response, and the theory is best tested by situations in which the number of these button presses is relatively large. DAT and the existing force models make different predictions about the relationship between the length of the sequences and the ESP scores resulting therefrom, and it is on the basis of tests of these predictions that May and colleagues claim confirmation of the model (May et al., 1995b).

DAT can be applied to ESP as well as PK data, and it is especially well suited to ESP experiments of the RNG type. One way it could work would be for P to use the DAT mechanism to enter the target stream at the time it was about to produce a "biased" subsequence that is consistent with a naturally occurring response bias of P. This would then create a bias-matching situation analogous to what was found with Schmidt's data, and thereby yield an increase in ESP hits.

The Influence of Belief on the DAT Effect

As noted above, there is evidence in the parapsychological literature that believers in ESP score more positively than skeptics in ESP tests, and most of this evidence comes from ESP tests of the forced-choice type, which we intend to use in our proposed research. Since DAT is postulated to be the mechanism by which ESP operates, we would expect it to have comparable correlates, including belief.

SUBLIMINAL PERCEPTION

There is evidence suggesting that subliminal stimulation influences behavior on a variety of tasks more than or differently than does stimulation above the awareness threshold, and some of these changes (e.g., those involving memory) fall in the category of performance enhancement (Bornstein & Pittman, 1992). We are not aware of any ISL studies that have used subliminal feedback, but the success of such stimulation in other contexts makes it reasonable to explore whether it might enhance ISL as well. Furthermore, subliminal stimuli are thought to interact more directly with unconscious mental processes than do supraliminal stimuli, and ISL is by definition unconscious. For this reason, the subliminal stimuli might register better in Ps' minds than the supraliminal stimuli that have always been used heretofore.

A very topical issue within the neuroscience of "unconscious" learning is the apparent paradox that the harder one tries, the less effective implicit learning will be (Fletcher et al., 2005). A similar paradox has repeated surfaced in the parapsychological literature, where conscious effort on the part of the guessing subject has diminished scoring levels in psi tasks (Schmeidler, 1986).

THE KRUMMENACHER EXPERIMENT

An important ESP aspect of the proposed experiment follows first from a reanalysis by applicant Palmer of data from a previous ISL experiment conducted at the University Hospital Zürich by Peter Krummenacher (2003) in order to fulfill the requirements for a Masters degree. We will summarize the key elements of the experiment and reanalysis here.

Method

The sample consisted of 18 strong believers and 17 strong skeptics in the paranormal. As close as possible to half of each group received a pill containing the neurotransmitter dopamine (l-dopa) before the session. All Ps completed 3 100-trial runs in a 4-choice computer guessing task. The targets were 4 arrows pointing up, down, left, and right in a circular display, and Ps had to guess which arrow the computer had selected for that trial. Immediate feedback of the correct target was displayed after each trial. Targets for the 1st run were random, whereas targets for the other 2 runs were biased. In 1 of these, the target in trial t+1 was displaced 90 degrees clockwise (CW) from the target in trial t on 46% of the trials, and displaced 90 degrees counterclockwise (CCW) on only 10% of the trials (P = .25). In the other biased run the direction of the bias was reversed, and the order of these 2 runs was counterbalanced.

Results and Discussion

Implicit Sequence Learning. ISL was defined as the difference between the number of pro-bias and counter-bias responses within the run. This measure was labeled relative pro-bias responding (rPBR). Krummenacher found significant positive rPBR score among believers who had received dopamine, a result that supported a hypothesis of the study.

Palmer thought that a better measure of ISL would be the *increase* in pro-bias responding across the run, and thus he introduced the half-run as a new variable in the reanalysis. In accordance with the hypothesis, believers who had taken dopamine showed a modest increase in rPBR from the first to the second halves of the combined biased runs. The mean was nonsignificant in the 1st half-runs and significant in the 2nd half-runs. However, the increase itself was not significant. Moreover, the interaction underlying the effect was only suggestive statistically ($p < .1$) and was due primarily to a significant decline effect from the 1st to 2nd half-runs among dopamine skeptics.

Anomalous Anticipation Effect. When Palmer began his analyses of the Krummenacher experiment, he neither attempted to nor expected to find any significant ESP effects, primarily because the experiment was conducted by a skeptical research team in a typically sterile academic environment. However, an unexpected finding emerged from an analysis of the random run that was comparable to the analysis applied to the biased runs. Statistically significant results were found in support of the interpretation that Ps manifested response biases in the random run that mimicked, or anticipated, the target biases they would confront in the 1st subsequent biased run, without knowing if this bias would be CW or CCW. Specifically, believers increased rPBR (defined with respect to the target bias in the next run) from the 1st to the 2nd half of the random run, whereas skeptics showed a decline in rPBR from the 1st to 2nd half of the run. This pattern will hereafter be referred to as an "anomalous anticipation effect" (AAE). It is illustrated in Figure 1.

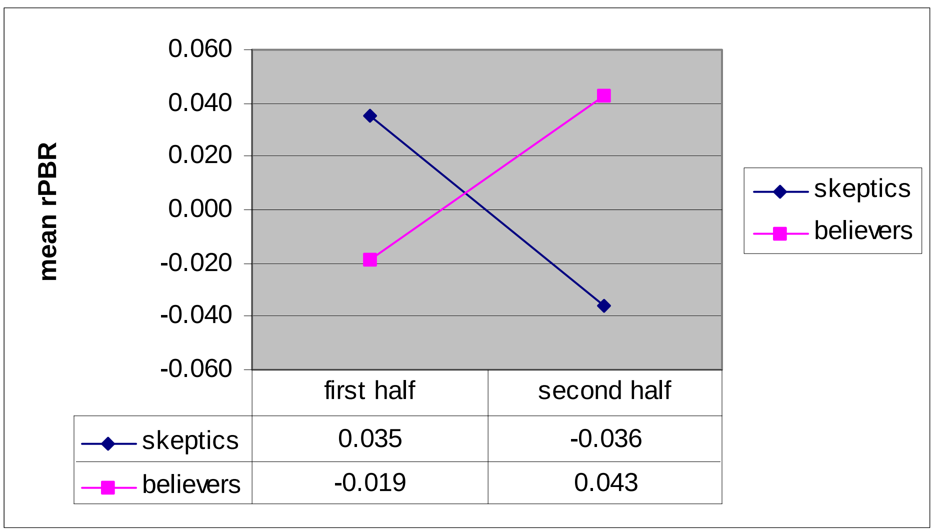

Figure 1. Mean relative pro-bias responding in the 1st and 2nd halves of the random run of the Krummenacher experiment as a function of belief in ESP.

It is most reasonable to assume, based on the significant difference in rPBR between believers and skeptics in the 1st half of the random run, that prior to or early in the random run Ps somehow received, unconsciously, correct information about the *nature* of the target bias, but either incorrect information or, more likely, no information at all about *in which run* the bias would manifest. If they received no information about timing, it would be natural for them to assume that the bias applied to the run they were then engaged in -- the random run. Skeptics processed this information about the nature of the target bias in 1st biased run correctly, whereas believers distorted the information, identifying the bias as opposite to what it really was – exactly the reverse of the traditional sheep-goat effect (Schmeidler & McConnell, 1958/1977). As the random run progressed, both groups realized from the feedback that their "strategy" wasn't working, so they tended to undertake the exact opposite strategy for the remainder of the run. The fact that they chose a reverse strategy rather than an entirely different strategy suggests that they may still have attached some credibility to the original information, thinking perhaps that they had misinterpreted the direction. Finally, it is assumed that all this reasoning was implicit, that is, unconscious.

The AAE is conceptually similar to the presentiment effect, which has appeared in a number of recent parapsychological experiments, both as a precognitive electrodermal response (e.g., Bierman & Radin, 1997; Radin, 1977) and a precognitive behavioral response (Bem, 2003) to emotionally evocative stimuli that P views subsequently. In both the AAE and the presentiment effect, Ps make responses that seem to reflect foreknowledge of stimuli to be presented to their normal senses shortly thereafter. The main difference is that in presentiment studies, the visual stimuli follow the precognitive responses by just a few seconds, whereas for the AAE the lag is much longer.

The Schuler Experiment

This was the first ISL experiment conducted after Palmer arrived in Zürich. The experimenter was a University of Zürich student named Matthias Schuler. The experiment was undertaken simultaneously with Palmer’s reanalysis of the Krummenacher experiment and thus the hypotheses were not influenced by the results of that reanalysis. Although the designs of the two experiments were similar, the Schuler experiment was not envisioned as a replication of the Krummenacher experiment.

Method

The methodology of the Schuler experiment differed from that of the Krummenacher experiment in the following respects:

First, whereas Krummenacher used an exclusively male sample, the present experiment used an exclusively female sample. The primary reason for this decision was to balance the invasion of our limited pool of potential Ps. Had we again used all males, for example, an insufficient number of males might be available for subsequent experiments. We generally have more difficulty recruiting males than females for our experiments.

Second, we decided to maintain the same kind of target bias (CW or CCW) for both biased runs. We became concerned that shifting to the opposite bias after the 1st biased run, as was done in the Krummenacher experiment, might confuse Ps by introducing a negative transfer effect and thus inhibit learning. For similar reasons, we compressed the two 100-trial runs of the Krummenacher study into a single 200-trial run, thereby maintaining the continuity of the task. We also decided, for purely exploratory purposes, to add an additional 200-trial run, to see if learning would stabilize or perhaps advance as a result of more extensive testing with feedback.

Third, the present experiment did not include a dopamine condition. Although the ISL results in the Krummenacher experiment occurred exclusively among dopamine Ps, new departmental regulations on the administration of drugs to research Ps rendered the reintroduction of this condition impractical.

Results

Implicit Sequence Learning. The results provided no evidence of ISL, which is not particularly surprising in view of the fact that no ISL was demonstrated among the non-dopamine Ps in the Krummenacher experiment. The changes in the design clearly were not sufficient to cause ISL to take place.

However, there was highly significant rPBR in both of the biased runs. As this effect was evident from the beginning of the runs, it cannot be attributed to ISL. On the other hand, we are reluctant to interpret it as ESP because the targets were grossly non-random. Strong pro-bias responding at the beginning of the 1st biased run had also been found in the Krummenacher experiment among dopamine Ps who were responding to a CW biased target sequence.

Anomalous Anticipation Effect. Results in the random run were consistent with those found in the Krummenacher experiment, with believers showing an increase in “rPBR” (as defined by the target bias in the immediately following biased run) from the 1st to the 2nd half of the run and the skeptics a decrease. However, this pattern was not statistically significant. Instead, there was a significant difference in rPBR between believers and skeptics throughout the random run. Contrary to what one might expect from the sheep-goat effect, believers showed significantly negative rPBR (akin to psi-missing) and skeptics showed nonsignificantly positive rPBR. The results are illustrated in Figure 2.

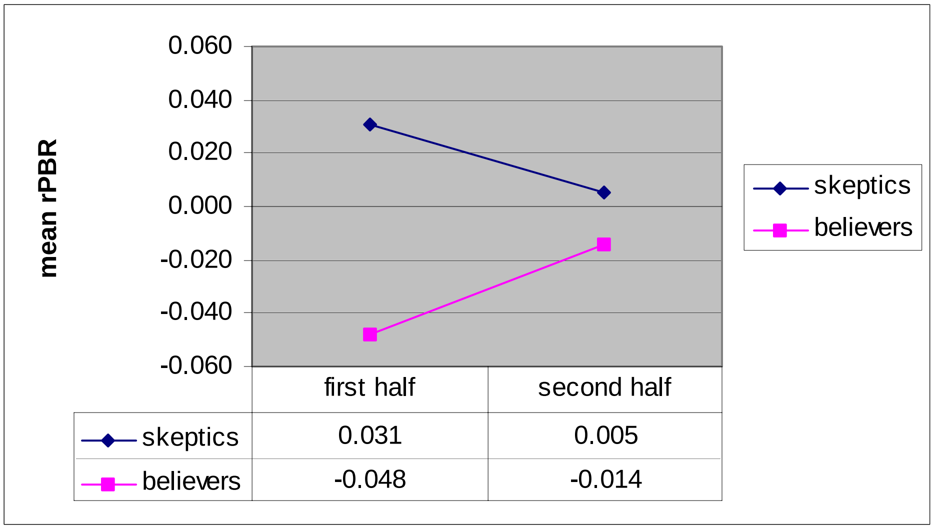

Figure 2. Mean relative pro-bias responding in the 1st and 2nd halves of the random run of the Schuler experiment as a function of belief in ESP.

COMBINED AAE RESULTS OF THE TWO EXPERIMENTS

As the procedure was essentially identical in the random runs of the 2 experiments, it is possible to combine the results from these runs, introducing experiment as an additional between-P factor in the ANOVA. As the only meaningful consistent difference across the 2 experiments in this regard was the sex of Ps (male in the 1st experiment, female in the 2nd), this factor is best interpreted as a gender factor. The only other difference of note is that about half Ps in the 1st experiment received levodopa, but as this manipulation had no significant effect on the results from the relevant ANOVA in the random run of Experiment 1 it can be considered inconsequential for the comparison between the experiments.

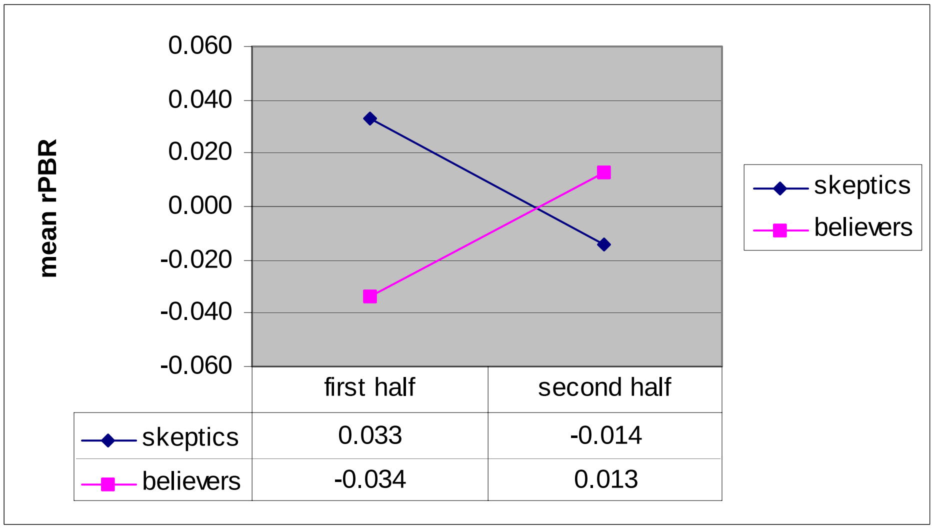

Figure 3. Mean relative pro-bias responding in the 1st and 2nd halves of the random run of the combined experiments as a function of belief in ESP.

As illustrated in Figure 3, the ANOVA yielded a significant belief by half interaction in support of the original AAE hypothesis. Believers scored significantly below chance in the 1st half of the random run and nonsignificantly above chance in the 2nd half. The increase was suggestively significant. Skeptics scored above chance to a suggestive degree in the 1st half of the random run and nonsignificantly below chance in the 2nd half. The decrease is significant. Finally, the difference in rPBR between believers and skeptics in the 1st half of the run is significant.

The Bial Experiment

There were four major respects in which the design of this experiment differed from the previous experiments: (1) the new task involved a 4-choice linear display as opposed to Krummenacher and Schuler's 4-choice circular display: (2) The random runs (N = 3) preceding the 1 biased run in this study, in contrast to Krummenacher and Schuler, did not include trial-by-trial feedback to P. The non-feedback runs were used to test the hypotheses related to global responding and the matching of target and response biases: (3) For half the Ps, the targets in the feedback run were fed back to them subliminally: (4) In the experimental condition, after the first 10 trials in the feedback run, the feedback matched Ps' own previous responses in this run. As the targets thus strongly reflected Ps' own response bias tendencies (pro-bias targets), we expected the opportunities for ISL to be maximized. In the control condition, targets were selected so as to diametrically oppose Ps' response tendencies, as determined from their biases in the preceding non-random runs (counter-bias targets).

Because of the nature of the feedback manipulation in the experimental condition, it was not practical to define the ISL in terms of relative pro-bias responding (rPBR), as in the Krummenacher and Schuler experiments. Instead, guessing accuracy (hitting) was the dependent variable. ISL so defined was not evident in the Krummenacher experiment, but this analysis has yet to be performed for the Schuler experiment.

We tested the DAT-inspired model in this experiment by assigning Ps a hit in the lst non-feedback run whenever they made a mouse click registering their guess at the same time a hidden computer address was in a certain state, determined randomly with a 1/5th probability. Moreover, Ps were rewarded for these hits by being given in the next trial a target that matched their response bias, as determined from the preceding runs.

Although the AAE was also looked for in this experiment, its primary purpose was to test the other hypotheses discussed above. Thus the context for the AAE was different from that in Krummenacher and Schuler experiments, which means a confirmation of the effect would represent a conceptual rather than a strict replication. We believe that effects are more interesting and important if they appear in a variety of different circumstances. Specifically, the AAE was tested by noting if Ps' response tendencies in the last non-feedback run would mimic the target bias they would receive in the immediately following feedback run. As in our previous ISL experiments, Ps did not know at the time of the non-feedback run which target bias they would receive in the feedback run.

Finally, with the exception of the AAE, we expected the hypothesized effects to be stronger (more positive) for believers than for skeptics.

HYPOTHESES

ESP

(1) Participants will register their responses more often than expected by chance when the computer is in a state leading to a favourable target in the next trial (DAT).

> (1a) Hypothesis (1) will be confirmed more strongly for believers than for skeptics.
>
> (1b) Ps will score significantly above chance on those trials with biased targets resulting from DAT.

(2) Skeptics will demonstrate a positive AAE and believers a negative AAE in the non-feedback run immediately preceding the feedback run.

(2a) The AAE will be stronger in the subliminal condition than in the supraliminal condition.

ISL

(3) With trial-by-trial feedback, target sequences reflecting the target biases of P will be implicitly learned better than target sequences reflecting target biases opposite to P’s.

(3a) Believers will learn both kinds of sequences better than skeptics.

(3b) Target sequences reflecting the target biases of P will be learned better in the subliminal condition than in the supraliminal condition.

Response Bias

(4) In the non-feedback runs, guessing accuracy will be higher when target sequences are biased than when they are random, because of matching target and response biases in the runs with biased targets.

(4a) Hypothesis (4) will be confirmed more strongly for believers than for skeptics.

METHOD

Participants

Sixty-four normal volunteers were recruited from the University of Zürich community and the city of Zürich. Written informed consent was obtained at the beginning of the test session.

As an additional requirement, these Ps had to indicate either that they have a strong belief in ESP and have had previous psychic experiences, or that they have a strong disbelief in ESP and no previous psychic experiences. This variable will be referred to as "belief". This specification was included in the recruitment poster.

Midway through the experiment we developed a concern that we would not be able to obtain a sufficient number of Ps before Dr. Palmer had to leave Zürich. It was thus decided to offer a prize of 500 Swiss Francs (approximately $400) to the P who achieved the highest score in the experiment[1]. To which runs this applied was left undefined so Ps would be equally motivated for all the runs in the experiment. In fact, the prize applied

[1] We gratefully acknowledge the Parapsychology Foundation for providing the funds for the prize.

only to the first 3 runs, as the procedure for Run 4 was not the same for all Ps. Although all Ps were eligible to win the prize, only those 38 Ps who were tested after the prize was decided upon knew about it before their test session. The winner was a skeptic.

Questionnaires

During the course of the experimental session, Ps completed the following questionnaires.

(1) *The Australian Sheep-Goat (ASG) Scale* (Thalbourne & Delin, 1993) was used as a check on the status of Ps who assigned themselves to the believer and skeptic groups. The ASG Scale consists of 18 items reflecting both belief in and experiences of various types of psychic phenomena. The items are presented in a visual analogue format, with scores on each item ranging from 0 to 13.

(2) *The Post-Test Assessment (PTA) Scale* was developed by applicant Brugger to assess how Ps react to the test procedure in implicit learning experiments of the type conducted by himself and his associates. The most important question asks Ps whether and, if yes, at what point in the testing, they came to expect that a target sequence was biased, and the nature of that bias. Ps who could correctly identify the bias were classified as "detectors". Their data was not included in the formal analyses, and they were replaced by new Ps with the same mental status and belief in ESP. In past research of this type conducted by Brugger and associates, about 15% of the original sample have proven to be detectors. A second set of questions asks Ps to estimate how many trials were included in each run and (in the present case) the number of such runs in the feedback condition. Third, Ps are asked if they responded intuitively, adopted a logical strategy, or a combination of the two. They are asked to describe any strategies they used and when they used them.

(3) *Drawing Task*. This test was developed as a measure of cerebral lateralization with respect to perceptuo-motor organization (Alter, 1989: Alter, Rein, & Toro, 1989). Ps are asked to rapidly draw on separate sheets of paper 6 familiar objects: bicycle, walking dog, bus, facial profile, airplane, and pitcher (ewer). The score is the number of drawings in which the object is facing right minus the number in which it is facing left, the difference divided by the total number of drawings, and has a range from -1 to +1. Drawings in which the object faces neither right nor left are not counted. Right–handers tend to produce drawings facing left, and left-handers tend to produce drawings facing right, but the discrimination is not absolute (Alter, 1989).

(4) *LIMBEX Scale*. The Limbex is intended to measure signs of temporal lobe dysfunction, or what is referred to more specifically as signs of complex partial epilepsy. It was developed by Brugger, who chose the 13 items from a longer scale by Makarec and Persinger (1990) that had the highest point biserial correlations with the total score in a sample of 40 volunteers. Each item of the LIMBEX is a 6-point scale, resulting in a

theoretical range of scores from 0 to 65. Although persons with complex partial seizures have been shown to score high on the original scale, some others also obtain high scores, and a high score by itself is not diagnostic of a seizure disorder.

(5) *Ambiguity Tolerance Scale (AT-20)*. This 20-item true-false scale is a revision of the 16-item Rydell-Rosen Ambiguity Tolerance Scale (McDonald, 1970). MacDonald defines a high scorer on the scale as a person who seeks out ambiguity, enjoys ambiguity, and excels in the performance of ambiguous tasks. The task in this experiment clearly could be described as ambiguous. The AT-20 correlates in the .4 range with Rokeach's Dogmatism Scale and Gough and Sanford's Rigidity Scale (MacDonald, 1970).

The ASGS was given primarily to confirm the self-reports of Ps as believers or skeptics. The other scales were included for strictly exploratory purposes and in some cases to give Ps something to do while E calculated their response biases from the 1st 2 runs (see below).

Equipment

Testing was performed on a Compaq Deskpro EXM/P800 computer with a 17" LCD Touch Monitor 1725L. Random target sequences were generated using Visual Basic, whereas the on-screen presentation was programmed with Java-Script.

The ESP/ISL Task

P was seated in front of the computer monitor, which displayed squares containing the digits 1, 2, 3, 4, arranged in a vertical column in either increasing or decreasing numerical order from the top to the bottom of the screen. The reason for the vertical display was to eliminate the effect of left/right response biases potentially confounding P's choice. This would correspond to "pseudoneglect" in the case of normal Ps (Luh, 1995). Superimposed over the column of digit was a box containing the word "start". P clicked on this box to begin the run, at which time the box disappeared. P's task was then to guess which digit the computer would select for the ensuing trial. Ps indicated their choice by saying the digit out loud and simultaneously clicking the mouse. E, who was seated next to P, immediately entered Ps response on the keyboard. P's oral responses were tape recorded, and after the session E checked the typed responses against these oral responses to check for possible entry errors.

If the protocol called for feedback, after 1 sec the column of digits was replaced in the center of the screen by a feedback stimulus consisting of the single digit selected by the computer as the target for the trial. If P's choice matched the target (a "hit"), the feedback square was colored green. For some Ps, the feedback square, which was exposed for 30 ms, was both forward and backward masked with random line drawings[2]. In this condition, the green background for hits was removed. After 3 sec, the feedback stimulus was replaced by the array of die faces, in anticipation of the next trial. If there was no feedback, the array reappeared 1 sec after P's response.

The number of milliseconds between the appearance of the array and P's mouse click to indicate their guess was recorded by the computer as a measure of reaction time. The computer also recorded and stored the target sequence type (see below), the run number, the targets, and P's responses.

Procedure and Testing Protocol

Each P completed 2 sets of 2 runs. The 1st 2 runs each consisted of 81 trials and were administered without feedback. For one of the runs, targets were assigned by an algorithm that, after the randomly selected 1st target, produced the extreme form of the "counting" bias characteristic of responses produced by Alzheimer's patients (Brugger, Monsch, Salmon, & Butters, 1996). If the 1st target was 2, the sequence was 2, 3, 4, 1, 2, 3, 4, 1, 2, 3 …". The 2nd run drew exclusively biased targets generated by an algorithm created to mimic a kind of response bias often demonstrated by normal Ps, namely repetition avoidance. In this run the targets never repeated, but after the 1st target in the sequence each target appeared an equal number of times (i.e., 20). Otherwise the sequence was random. The order of these 2 run types was counterbalanced across Ps.

The 1st run was preceded by as many practice trials as necessary to assure that P understood the procedure and that P and E were "in synch" regarding their respective mouse and keyboard entries. If 2 successive mouse clicks or keyboard entries occurred without an intervening input of the opposite type, the computer indicated the error by 1 (successive keyboard entries) or a series (successive mouse clicks) of beeps. E then said "repeat" or "next", thereby instructing P what to do to correct the error, and repeated the keyboard entry if necessary. This problem arose quite rarely in the formal testing.

Target randomization employed an algorithm developed by Marsaglia and Zaman (1987) and thoroughly tested to assure passage of numerous tests of nonrandomness. The 1st pair of seed numbers for the formal experiment were 1 and 2[3], and every time in the experiment that a new sequence was called for, the seeds were advanced to the next pair. This procedure provided each P with unique target sequences (no stacking effect).

Following the 2nd run, P moved to a chair facing away from E and the computer screen, and then completed, in order, the drawing task, the AT-20, and the LIMBEX scale. At the same time, E moved to the chair in front of the computer and determined P's most marked

[2] Pilot testing indicated that this double masking was necessary to render the feedback digit consciously undetectable at 30 ms, which was the fastest exposure time the software would allow.
[3] The Marsaglia algorithm requires input of 2 seed numbers.

response bias across the previous 2 runs. The computer records from these runs were merged and the resulting file submitted to analysis using software developed by Towse and Neil (1998). The frequency of each *single* target (1, 2, 3, and 4) and the relationship of each target to its predecessor (*shifts* of 0, +1, +2, or +3 units) was recorded from the Towse output[4]. The chance probability for each of these 8 alternatives is .25. The summed frequencies for the 2nd and 3rd most frequent single and shift responses were then computed and recorded. A table had been developed which indicated the chance likelihood for each of these frequencies and ranked them, with the least likely alternatives getting the highest ranks. The table provided ranks for each of 24 possible response biases, i.e., the sum of the most frequently called 1, 2, and 3 choices for singles and shifts respectively. For example, conformance to the counting bias would yield a high rank for +1 shifts, whereas repetition avoidance would be reflected by a high rank for the sum of +1, +2, and +3 shifts (equivalent to a low frequency of 0 shifts, or repetitions). The bias that received the highest rank, and the value of that rank, were then recorded by E.

Following completion of their respective tasks, P and E resumed the seating arrangements in effect for the 1st 2 runs. Following a few practice trials, Run 3 (N = 101 trials) was initiated. From P's point of view the procedure was the same as for the 1st 2 runs, except that a 2 sec delay was introduced before each trial, during which the computer screen was blank. Ps were instructed to blank their minds during the 2 sec interval and only formulate their guesses when the column of digits returned to the screen. This modification of procedure was introduced in an effort to increase the variability of reaction times by attempting to break up the rhythm Ps often got into during the 1st 2 runs. Pilot testing had indicated this modification would have the desired effect.

Run 3 was chosen to test for DAT. During this run, an address inside the computer randomly alternated its content between 0 and 1, such that it (or, we could say, the computer) was in the "1-state" 20% of the time during the run. This outcome was programmed as follows. Thirty repetitions of the digits 2 through 6 (N = 150) were randomly permutated, separately for each P. Each digit represented a .2 sec interval, during which the computer would be in the "0-state". Following this time span, the computer would be in the 1-state for .2 sec. Thus, it would be in the 0-state anywhere from .4 to 1.2 sec (the sequence of these intervals being random) before the next 1-state, and there were never 2 1-states in a row. The subroutine was activated at the time P clicked the "start" box on the screen, and the sequence simply recycled after it was exhausted (every 2.5 min).

Each time P clicked the mouse while the computer was in the 1-state, the next target was guaranteed to conform to Ps most likely response bias, as defined by the calculation (described above) of Ps most extreme response bias during the 1st 2 runs. For example, Ps who called an excess of 4s in the 1st 2 runs would be guaranteed to receive a 4 as the

[4] This required that +1 and –3, +2 and –2, and –1 and +3 each be summed from the Towse table.

target for the next trial following any trial in which they clicked the mouse while the computer was in the 1-state. Likewise, if Ps had demonstrated repetition avoidance previously, their target following a 1-state mouse click would never duplicate their immediately preceding response. The effect of this procedure was to increase Ps' chances of a hit on the manipulated trials, insofar as they maintained the response bias they demonstrated in the 1st 2 runs.

Run 4 (N = 100 trials) was the feedback run. For half Ps (experimental condition), the 1st 10 trials were random, after which the targets repeated P's own responses with a lag of 10 trials. For example, the target for Trial 11 was P's response on Trial 1, the target on Trial 12 was P’s response on Trial 2, and so forth. For the other half of the sample (control condition), a counter-bias target situation comparable to the pro-bias target situation created by the lag procedure described above was created, based on the response bias inferred from that P's 1st 2 runs. The response tendency (singlet or shift) having the proportion of choices most deviant from .25 was shifted two steps. For example, if a P demonstrating repetition avoidance had his or her most deviant proportion of choices on 0 shifts (say, .05), the target sequence in Run 4 would exhibit +2 shifts between adjacent targets on only .05 of the 100 trials, with each of the other 3 shift alternatives having the proportion (1-.05)/3 = .317. If in the 1st 2 runs P had called an excess of 1s and 3s, for a combined proportion of .70, the target sequence for Run 4 would contain a proportion of .35 of targets 2 and 4 respectively, and a proportion of (1-.7)/2 = .15 of targets 1 and 3 respectively.

After Run 4, P was administered the PAT and ASG scales, in that order. During this period, E returned to his office and printed out the results of the 4 guessing runs and entered the data on the Participant Feedback Form that also explained the rationale of the experiment. When E returned to the testing room, and after P had completed the scales, E gave P the feedback form, which P read over. E then showed P the data sheets and answered any questions P had about the experiment or their results. Finally, P was asked not to discuss the details of the experiment with anyone who might participate in the experiment at a later time.

Summary of Design

Five between-Ps variables were counterbalanced: (1) *belief* in the paranormal (believer vs skeptic), (2) order of the 4 digits on the *screen* (ascending vs descending), (3) target *bias* in the 1st 2 runs (repetition avoidance vs counting), target bias *condition* in the feedback run (experimental vs control), and (5) *speed* of presentation visibility of the feedback digits in the feedback run (supraliminal vs subliminal). The 4 runs served as the single within-Ps variable.

RESULTS

ELIMINATION OF FLAWED DATA

Eight Ps were replaced during the course of the experiment. Five were replaced because of recording errors of either targets or responses in one or more runs. This came about because of errors in the sequence of oral calls and mouse clicks by P or E that could not be resolved by listening to the tapes of P's calls. There were 5 other cases involving Run 1 or 2 in which such errors involved the final 5 or fewer trials in the run. In these cases, the suspect trials were eliminated from the calculations of the run scores. One P was replaced because she had been defined as a skeptic but scored above the midpoint in the ASGS, i.e., in the believing direction. One believer was replaced because in Runs 3 and 4 she called the same number many times in succession, creating extreme response bias scores. Finally, 1 believer was replaced because she correctly detected during Run 4 that the targets were related to her own responses. This caused her to obtain an extremely high number of hits.

After completion of testing it was found that for 1 skeptic in the control condition of Run 4, the protocol for defining the target bias for this run was grossly violated, such that the targets reflected the Ps response bias in Runs 1 and 2 positively rather than negatively. There was not sufficient time to replace this P, so her Run 4 guessing data were eliminated from the analyses.

TESTS OF HYPOTHESES

ESP: Decision Augmentation Theory

Hypothesis 1 was tested by examining how frequently Ps clicked the mouse when the computer was in the "1-state" in Run 3 -- 20% of the time by chance. The actual percentage of such clicks was 20.80 (*SD* = 3.34), *t*(63) = 1.94, *p* = .057.[5] As this result does not quite reach significance, Hypothesis 1 is suggestively supported. However, the percentage for believers was significantly high (*M* = 21.7; *SD* = 3.23), *t*(31) = 3.06, *p* = .006 and significantly higher than the nonsignificant percentage for skeptics (*M* = 19.85; *SD* = 3.21), *t*(62) = 2.36, *p* = .022. Thus hypothesis 1a was strongly supported.

On trials in which Ps received targets consistent with their response biases in the preceding runs – trials in which the computer was in the 1-state for the preceding trial – the percentage of hits (mean chance expectation = 25%) was quite high (30.92; *SD* = 11.18) and strongly significant, *t*(63) = 4.23, *p* = .0001. Thus, Hypothesis 1b was strongly supported. This result confirms that 1-state trials produced the intended positive reinforcement. However, this advantage was not enough to produce significant positive

[5] All *p*-values in this report are two-tailed.

scoring in the entire Run 3 for either believers (*M* = 26.06; *SD* = 4.56), *t*(31) = 1.32, or skeptics (*M* = 25.84; *SD* = 3.07), *t*(31) = 1.55. However, due to greater power the mean for the whole sample just missed significance (*M* = 25.97; *SD* = 3.86), *t*(63) = 1.97, *p* = .053.

ESP: The Anomalous Anticipation Effect

To test for the AAE, we applied the same bias analysis to the targets of Run 4 as we had applied to the responses in Runs 1 and 2 during the session break (see above). This allowed us to determine for which target or target pair (shift) the bias was greatest. We then calculated the relative pro-bias responding in Run 3 by determining the number of responses representing the target bias in Run 4 minus the number representing the opposite bias. For example, if the target bias in Run 4 was an excess of 3s, we would calculate for Run 3 the number of 3 calls minus the number of 1 calls. Because the number of trials considered in these analyses differed widely among Ps, these difference scores were converted to *z*-scores, with an expected value of 0 by chance.

The two AAE hypotheses predicted that skeptics would show a negative AAE and believers a positive AAE, and that the effect would be stronger with subliminal feedback of targets in Run 4 than with supraliminal targets.

An ANOVA of the of the AAE *z*-scores was performed using belief, speed (supraliminal vs subliminal), and condition (experimental vs control) as independent variables. Hypothesis 2 was tested by the belief main effect. Although, as predicted, the z-score mean was positive for skeptics (*M* = 0.38; *SD* = 2.59) and negative for believers (*M* = -0.30; *SD* = 3.17), the difference was not significant, *F*(1,55) = 2.27, *p* = .137. Neither mean differed significantly from chance. Thus, Hypothesis 2 was not supported. Hypothesis 2a was tested by the speed by belief interaction, which was significant, *F*(1,55) = 5.66, *p* = .021. The means confirmed that the AAE was, as predicted, stronger in the subliminal than the supraliminal condition. Thus, hypothesis 2a was supported. In the subliminal condition, skeptics showed a positive AAE (*M* = 1.08; *SD* = 3.17) and believers a negative AAE (*M* = -0.72; *SD* = 3.17), *t*(29) = 2.08, *p* = .046. Neither mean differed significantly from chance. The AAE reversed slightly in the supraliminal condition (skeptics: *M* = -0.27; *SD* = 2.55, believers: *M* = 0.11; *SD* = 3.91), *t*(25.8) = 0.32, *ns*. The variance was significantly greater for believers than skeptics, (*F* = 7.94, *p* = .008) by Levene's test.

The above interaction supporting Hypothesis 2a was superseded by a significant belief by speed by condition interaction, *F*(1,55) = 5.34, *p* = .025. Inspection of the cell means revealed that condition had no effect in the subliminal condition, but a strong effect in the supraliminal condition, as evidenced by a significant belief by speed interaction in this condition, *F*(1,28) = 9.43, *p* = .005. The predicted AAE occurred in the supraliminal control condition, with skeptics scoring more positively (*M* = -1.73; *SD* =2.41) than believers (*M* = -3.39; *SD* = 1.05), *t*(14) = 1.78, *p* = .097[6]. However, in the supraliminal experimental condition, the AAE significantly reversed, with believers revealing a more positive AAE (*M* = 3.60; *SD* = 1.96) than skeptics (*M* = 1.19; *SD* = 1.79), *t*(14) = 2.56, *p* = .023. This result from the supraliminal condition is not an AAE. It indicates that believers in this cell maintained their response bias more consistently from Run 3 to Run 4 than did skeptics.

Implicit Sequence Learning: Feedback Run

An ANOVA was performed with belief (believer vs skeptic), condition (experimental vs control), screen (targets ascending vs targets descending), and speed (subliminal vs supraliminal) as between-P variables and half-run as the sole within-P variable. The dependent variable was the percentage of hits in Run 4. There was no significant overall learning effect, *F*(1,47) = 1.61, *p* = .211 for the halves main effect, with a mean of 25.11 (*SD* = 7.57) in the 1st half run and a mean of 26.84 (*SD* = 8.15) in the 2nd half-run. However, there was an almost significant half x belief interaction, *F*(1,47) = 3.86, *p* = .055. A further breakdown revealed significant evidence of ISL for skeptics. Ignoring the first 10 trials of the run (because matching of targets and previous responses did not begin until Trial 11 in the experimental condition), skeptics scored slightly below chance in the 1st half of the rest of the run (*M* = 24.08; *SD* = 7.64) and significantly above chance in the 2nd half (*M* = 28.53; *SD* = 7.70), *t*(31) = 2.55, *p* = .016. The increase from the 1st to the 2nd half was also significant, *t*(30) = 2.42, *p* = .022. Believers actually declined slightly from the 1st half-run (*M* = 26.11; *SD* = 7.49) to the 2nd half-run (*M* = 25.21; *SD* = 8.36), but neither mean nor the difference between them was significant.

A suggestive half x belief x condition x speed interaction, *F*(1,47) = 3.19, *p* = .080, indicated that the significant ISL effect for goats might need additional qualification. This interaction is illustrated in Figure 4. It reveals that the ISL effect for skeptics did not apply to the experimental condition when the targets were supraliminal. In this condition, the percentage of hits was quite high and significant among goats for both halves combined (*M* = 30.14: *SD* = 2.94), *t*(7) = 4.94, *p* = .002, and differed hardly at all between halves. Comparable results were obtained by believers in this condition (*M* = 29.72: *SD* = 3.40), *t*(7) = 3.93, *p* = .006. This high scoring rate also appeared in the 1st 10 trials for believers and skeptics combined (*M* = 30.00; *SD* = 21.60), although with such a small number of trials and high variability it was not significant, *t*(15) = 0.92.

---

[6] Note that the *z*-scores are substantially more negative in the control condition than in the experimental condition, because in the experimental condition the targets in Run 4 reflected Ps' response biases, whereas in the control condition they reflected the opposite. For this reason, the conditions main effect in the overall ANOVA was very large, *F*(1,55) = 100.87, *p* << .001.

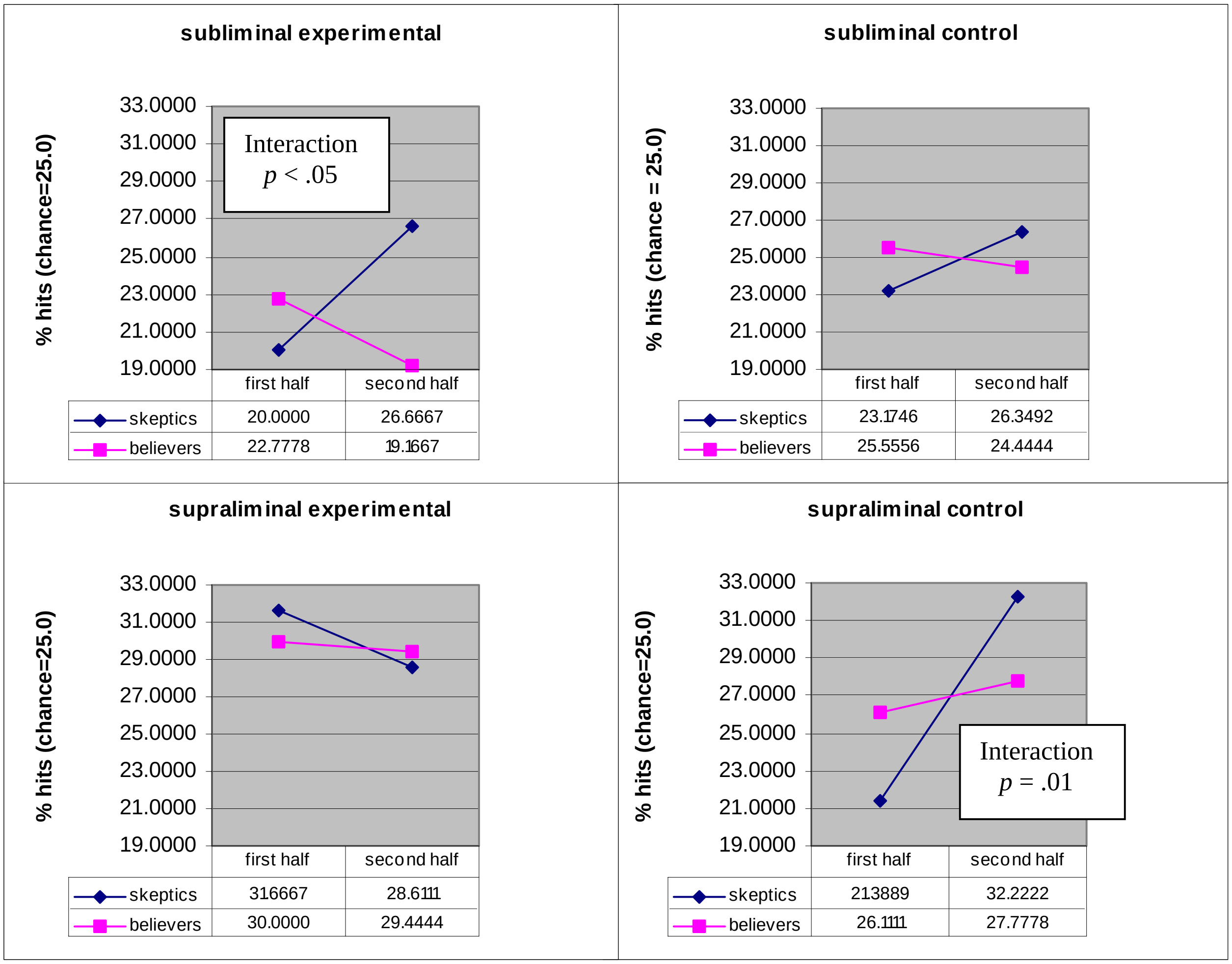

Figure 4. Differences in ISL between believers and skeptics in each of the 4 experimental conditions defined by nature of the target feedback (experimental vs control) and speed of presentation (supraliminal vs subliminal).

Response Bias: Non-Feedback Runs

These analyses compared the results from the 1st 2 runs, each having biased target sequences, with those trials from Run 3 having random targets (not determined by the DAT manipulaton). In the run where the targets demonstrated repetition avoidance (RA), our Ps had a hit percentage of 26.40 (*SD* = 5.51), which was significantly higher than the chance expectation of 25, *t*(63) = 2.04, *p* = . 046. P's scored almost as well on the run with targets reflecting the counting (CO) bias (*M* = 26.0; *SD* = 6.22), but this result was not significant, t(63) = 1.36, *p* = .179. Finally, with random targets, the hit rate was very close to chance (*M* = 24.77; *SD* = 4.69) and nonsignificant. However, the hit rate for the combined biased runs did not differ significantly from that of the random run, *t*(63) =

1.64, p = .107, nor did the RA run differ significantly from the random run, $t(63) = 1.67$, p = .100. However, both differences can be considered statistically suggestive. It can be concluded that Hypothesis 4 received modest support for repetition avoidance but not for counting.

Believers scored significantly above chance in the RA run (M = 27.15; SD = 5.40), $t(31) = 2.25$, p = .032, and higher than skeptics (M = 25.66; SD = 5.60), but the difference was not significant, $t(62) = 1.08$. The performances of believers and skeptics were virtually identical in the CO run (Believers: M = 26.09; SD = 6.60: Skeptics: M = 26.02; SD = 5.94) and the random run (Believers: M = 24.88; SD = 5.57: Skeptics: M = 24.67; SD = 3.70). Thus, Hypothesis 4a was not supported. In the random run, the variance was significantly higher for believers than for skeptics, F = 4.51, p = .038, by Levene's Test.

In line with our expectation we found a significant negative correlation between how many repeats Ps called in the RA run and their guessing success, $r(62) = -.259$, p = .038. In other words, the Ps who called the fewest repeats got the highest scores in the run in which the target sequence had no repeats but was otherwise random. In a similar vein, we found a significant positive correlation between how frequently Ps "counted" in the ascending direction (+1 shifts: e.g., 1-2, 2-3) and their success in the CO run, in which the targets also demonstrated ascending CO, $r(62) = .289$, p = .020. There were no such significant relationships in the random trials of Run 3. The general conclusion to be drawn from this pattern of findings is that Ps can elevate their guessing success when they exhibit a response bias that matches an extreme target bias of the same kind, and provides support for the rationale underlying Hypothesis 4 for both RA and CO.

A breakdown of these correlations by belief indicated that the correlation for RA is attributable exclusively to the believers, $r(30) = -.524$, p = .002, and it actually reversed for skeptics, $r(30) = .079$, *ns*. These correlations are significantly different from each other, z = 2.51, p = .012. On the other hand, the correlations for CO were quite comparable for believers, r(30) = .252, and skeptics, r(30) = .331. This result provides indirect support for Hypothesis 4a for RA, that an RA target sequence is more likely to be taken advantage of by believers than skeptics, provided that the believers avoid repeats in their calls.

It would seem to follow from the above interpretation that believers would call meaningfully fewer repeats than skeptics, but surprisingly that was not the case. The average number of repeats called by believers (M = 10.28; SD = 7.04) and skeptics (M = 11.00; SD = 5.41) were quite similar, $t(62) = 0.46$, *ns*. One might also have expected the avoidance of repeat calls to correlate with success in the CO run, as the targets in this run also contained no repeats. This correlation was, however, very close to 0.

EXPLORATORY ANALYSES

Post-Test Assessment Scale

In the final 3 runs, Ps on average predicted a hit rate greater than the 25% expected by chance, although the magnitude of the effect was not large (range of 28.2% to 29.9%). As one would expect, in all 4 runs the predicted success was greater for believers than for skeptics. The difference was significant in Runs 2, $t(44.1) = 2.25$, $p = .029$ and 3, $t(45.6) = 2.32$, $p = .025$ and suggestive in Run 4, $t(48.5) = 1.74$, $p = .089$. In all 3 runs, the variance was significantly higher for believers and skeptics. Although there was high variability among Ps in the number of trials they estimated for each run, the averages in the 4 runs ranged from 58.5 to 62.3. These are marked underestimates, and they occurred despite the fact that the written instructions mentioned that the number of trials per run would vary between 80 and 120.

Although estimates by Ps of the presence of particular kinds of target biases in Run 4 were generally quite low on the 6-point scale, believers estimated the likelihood of 3 of these to be significantly or suggestively greater than did skeptics. These were non-randomness in the frequency of individual targets (Believers: $M = 2.50$; $SD = 1.72$, Skeptics: $M = 1.16$; $SD = 1.61$), $t(62) = 3.23$, $p = .002$, non-randomness in the frequency of target pairs, or shifts (Believers: $M = 1.44$; $SD = 1.72$, Skeptics: $M = 0.81$; $SD = 1.61$), $t(62) = 1.77$, $p = .081$, and repetition of own calls (Believers: $M = 1.59$; $SD = 1.72$, Skeptics: $M = 0.91$; $SD = 1.61$), $t(62) = 2.10$, $p = .040$. As noted previously, only 1 P, a believer, had to be eliminated because she detected the target bias (repetition of own responses during the feedback run) in the experimental condition.

For Ps in the control condition, there was a significant positive correlation indicating that the higher proportion of hits they expected in Run 4, the more hits they actually obtained, r_s ($N = 31$) = .367, $p = .042$.The corresponding correlation in the experimental condition was negative, r_s ($N = 28$) = -.237. For Run 3, the prediction of success correlated negatively with success in the random trials of this run to a significant degree among all Ps, r_s ($N = 61$) = -.346, $p = .006$.

Psychological Tests: Drawing, AT-20, and LIMBEX

LIMBEX and ISL. There were 3 significant correlations of the above psychological tests and guessing accuracy. The first was a significant negative correlation between the LIMBEX scale and the increase in hits from the 1st to the 2nd half of Run 4 (discounting Trials 1-10), $r(62) = -.362$, $p = .003$. In other words, ISL was greatest among Ps who showed the least tendency toward temporal lobe dysfunction on the LIMBEX. Because believers averaged much higher on the LIMBEX ($M = 38.72$; $SD = 10.73$) than skeptics ($M = 16.59$; $SD = 7.91$), $t(62) = 9.39$, $p << .001$, the above correlation needs to be computed separately for believers and skeptics. The correlation is attributable primarily to the believers, $r(31) = -.366$, $p = .039$, with the skeptics making a minimal contribution, $r(31) = - .095$, *ns*. This result may be partly attributable to the fact that the variance of LIMBEX scores was higher for believers than skeptics, $F = 3.30$, $p = .074$, by Levene's test, allowing more range for a significant correlation to manifest.

Combined with the overall superior ISL among goats, this pattern of result suggests that it is a relatively high level of temporal lobe dysfunction, exhibited primarily by believers, that inhibits ISL. Skeptics as a group were able to demonstrate ISL because they included very few Ps with high LIMBEX scores. To verify this inference, an ANOVA was computed on Run 4 with half-run (minus trials 1-10) as the within-P factor and LIMBEX the between-P factor. LIMBEX was converted to a categorical variable, with those whose average item score was 3 or above on the 0-5 scale defined as high as the other Ps as low. This translated into a cutoff point of 39 on the total scale (theoretical range 0 to 65). There were 51 Ps with low LIMBEX scores (<39) and 13 with high LIMBEX scores (39+). The ANOVA yielded a significant half x LIMBEX interaction, $F(1,61) = 6.20$, $p = .016$ (see Figure 5). Among low LIMBEX Ps there was a significant ISL effect, with a nonsignificant mean percentage of hits in the 1st half of Run 4 ($M = 24.71$; $SD = 7.14$) and a more positive and statistically significant mean percentage in the 2nd half ($M = 28.04$; $SD = 7.50$), $t(49) = 3.82$, $p = .0004$. This ISL effect was balanced among high LIMBEX Ps by a decline effect of about the same magnitude as the ISL effect among low LIMBEX Ps, but because of the lower *n* (and power), it was not significant, $t(12) = 1.69$, $p = .116$. Their mean percentage hits in the 1st half-run was positive ($M = 26.67$; $SD = 9.21$) and in the 2nd half-run was negative ($M = 22.22$; $SD = 9.21$). Neither mean differed significantly from the chance expectation of 25.

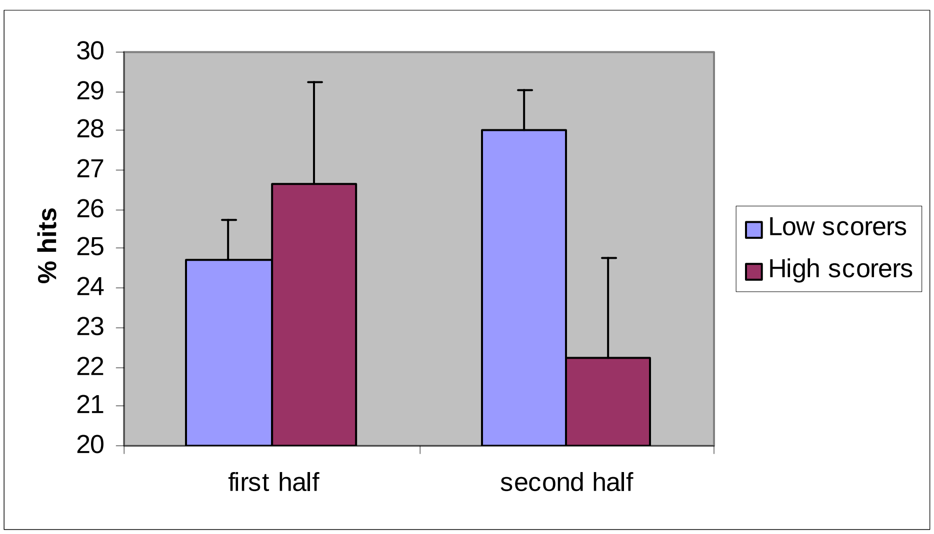

Figure 5. Implicit sequence learning as a function of scores on LIMBEX

LIMBEX and DAT Hits. There was a significant positive correlation between LIMBEX and the proportion of 1-state mouse clicks, $r(62) = .280$, $p = .025$. However, because of the large mean difference between believers and skeptics on the LIMBEX, once again we must look at the above relationship for believers and skeptics separately. Both correlations were quite low: $r(30) = .089$ for believers and $r(30) = .113$ for skeptics. Thus the significant overall correlation is best interpreted as simply a reflection of the fact, noted above, that believers obtained significantly more 1-state hits than did skeptics. It is still possible that LIMBEX represents a mediator of the believer / skeptic difference on DAT hits, but no evidence can be claimed for this mediation.

AT-20 and Run 3 Hits. Finally, there was a significant positive correlation between ambiguity tolerance and the proportion of hits in Run 3, both for all trials, $r(61) = .320$, $p = .010$, and for the random trials, $r(61) = .266$, $p = .035$. In contrast to the LIMBEX, the means on the ambiguity scale were quite similar for believers ($M = 11.09$; $SD = 3.16$) and skeptics ($M = 10.82$; $SD = 3.18$), $t(61) = 0.34$.

DISCUSSION

ESP: Decision Augmentation Theory

DAT predicts that in RNG "PK" experiments positive scoring is achieved by P intersecting a random target stream at times that produce targets that tend to match the calls that P will make. We operationalized this principle in the present experiment by allowing Ps to create targets matching their response biases, as estimated by their response sequences in previous runs. Ps could create these targets by making their responses at an opportune time, namely at a time in which the computer was randomly in the "1-state". Doing so would allow them to improve their score, and we thus predicted that Ps would generate more 1-state trials than predicted by chance. This hypothesis was suggestively confirmed with $p < .10$. As believers are more likely than skeptics to be motivated to attain a high score, it is not surprising that, as we hypothesized, a significant excess of 1-state trials was achieved only by believers. The fact that trials determined by the manipulation yielded a high percentage of hits (30.92%) demonstrated that the manipulation had the intended effect, although it was not strong enough to yield overall significant positive scoring in Run 3 for either believers or skeptics.

These DAT results represent what we call "implicit psi", that is, psi occurring without awareness by Ps that psi is being tested. Although Ps probably realized that ESP was tested in Run 3, they were not informed that the timing of their mouse clicks had any influence on their results. It is noteworthy that in the other test of implicit psi in this experiment, the AAE, the positive evidence of psi was produced by skeptics rather than believers. We will discuss this paradoxical finding below.

ESP: Anomalous Anticipation Effect

The AAE was only a minor consideration in the design of this experiment, and one consequence of this fact is that the design with respect to the AAE differed from that of the previous ISL experiment in 3 key respects. First, the random run in the present experiment (Run 3) did not include feedback of the targets. For this reason, we could not define the response bias as the relationship between the response and the previous target, as in the previous experiments. We thus defined it as the relationship between each response and the immediately succeeding response. Second, the time interval between the random run and the biased run was longer than in the previous research, because in the present experiment this was the period in which E calculated P's response bias and P was filling out questionnaires. Third, whereas in the previous experiments the two opposite target biases in the feedback run always involved a relatively large and fixed number of either CW or CCW relationships between successive targets, the opposite biases in this experiment were geared to each P's individual response bias.

Despite these differences, we obtained a significant AAE, but only if the target feedback in Run 4 was subliminal. Although a stronger AAE in the subliminal condition was predicted, the failure to also find any AAE in the supraliminal condition is disappointing, because the feedback was supraliminal in the two previous experiments in which the AAE was demonstrated. So in this sense, the current experiment failed to replicate the earlier ones. This pattern has occurred in other ESP experiments where a manipulation was added to increase the strength of a previously found effect, only to discover that the added manipulation, rather than building on the previous effect, replaced it (e.g., Bem & Honorton, 1994). It is as if there is a ceiling that these anomalous effects cannot exceed.

The tendency for believers to maintain their response bias from Run 3 to Run 4 in the supraliminal condition more than skeptics likely reflects Ps' "natural" tendency in the absence of AAE influence. Although there was no difference between believers and skeptics in their reported adherence to the instruction to respond spontaneously in Run 4, one might still expect that skeptics would explore, perhaps unconsciously, different logical strategies based on their perceptions of the feedback, whereas believers would approach Run 4 the same way as Run 3.

ISL: Feedback Run

This is the first of the 3 ISL experiments to provide a clear-cut ISL effect, but, contrary to expectation, it occurred with skeptics rather than believers. In the 1st half of the feedback run, skeptics scored close to chance, but their performance significantly increased in the 2nd half of the run, and this mean was significantly positive. It is noteworthy that the ISL effect occurred even when the targets were fed back to Ps subliminally, although ISL was not statistically superior with subliminal stimuli. Nonetheless, these results, along with the AAE results discussed above, indicate that the subliminal stimuli were being apprehended and processed, albeit unconsciously.

A higher level interaction indicated that the ISL effect did not apply to the experimental condition with supraliminal stimuli. Because the 1st half mean was so high in this condition, there may have been little additional room for learning to take place (that is, a ceiling effect). Although Ps had not at this time had sufficient time to benefit from the feedback, this result is not surprising because the targets matched Ps' response biases: results from the non-feedback runs illustrated that even without feedback such concordance can lead to increased hitting. However, this interpretation cannot explain why the same increased hit rate applied to the 1st 10 trials of the run, for which the targets were random. This hitting at the beginning of the run also was found in the Schuler experiment, regardless of whether the target bias was CW or CCW. Such results are puzzling, because the high scoring is evident before Ps had the opportunity to benefit from the feedback. Although it looks like ESP, we did not want to claim the Schuler result as evidence for ESP, because the targets were non-random. We don't want to claim the results from the 1st 10 trials of the feedback run in the present experiment as evidence for ESP, both because the effect was not significant, and also because the target sequence was the same for all Ps (stacking effect). With this small a number of trials, the consequences of stacking are not trivial.

A likely reason for the superior ISL by skeptics is provided by the post-hoc analyses of ISL in relation to scores on the LIMBEX, a measure of tendencies toward temporal lobe dysfunction. These analyses indicated a strong ISL across all experimental conditions by Ps who scored at or below the theoretical midpoint of the ISL scale, with high LIMBEX Ps actually exhibiting a suggestive decline effect from the 1st to 2nd half of Run 4. In other words, tendencies toward temporal lobe dysfunction appear to destroy ISL ability. The reason skeptics revealed better ISL than skeptics is that, as it turned out, none of them scored above the cutoff point on LIMBEX, although 13 of the 32 believers did.

This finding is not consistent with the results of the Krummenacher experiment, in which ISL was demonstrated only by believers who had ingested dopamine prior to the test session. However, the ISL in this group was quite weak, in particular because the increase in relative pro-bias responding from the 1st to 2nd half of the runs was small and nonsignificant. Thus, the results of the current experiment are more likely to reflect the true state of affairs. On the other hand, it must be kept in mind that the learning task and nature of the target bias differed between the two experiments. Also, the dependent variable in the present experiment was hitting rather than relative pro-bias responding. L-dopa believers showed a slight decline in hits from the 1st to the 2nd half runs in the Krummenacher experiment. It may be that the learning task used in the present experiment is better able to allow ISL than the learning task used in the previous 2 experiments. The bias to be learned in the present experiment (mostly the presence or absence of repeats) is less complex and probably easier to recognize than the CW or CCW relationship between adjacent trials in the previous experiments.

RESPONSE BIAS: NON-FEEDBACK RUNS

In general, the results from Runs 1 and 2 supported the experimental hypotheses. The overall hit rates for these runs were elevated, but significantly so only for RA. This is not surprising, because RA matched the response bias tendencies of most Ps better than CO did. Indeed, RA among our Ps was quite prevalent. For only 7 of the 64 Ps did the dominant response bias defined for Runs 1 and 2 not include RA, an outcome consistent with previous research (Wiegersma, 1982). The average number of repeats in the RA run was 10.64, well below the chance expected value of 20.

Although believers did not score higher than skeptics on the RA run, unlike skeptics the number of repeats they called correlated positively and significantly with their success in this run. This was true despite the fact that they averaged only slightly fewer repeats in this run than did skeptics. This pattern implies that if believers had called substantially fewer repeats than skeptics, their advantage in hits over the skeptics would have been greater and unequivocally significant. Conversely, the reversal of the repeats/hits correlation for skeptics adds weight to the conclusion that the excess hits for skeptics did not constitute a real effect. Although the mean hits for both groups combined was significantly positive and the difference between believers and skeptics was not significant, the mean hits for skeptics did not differ significantly from chance.

In short, this line of reasoning leads to the conclusion that only believers with a relatively small number of repeats obtained an excess of true hits in the RA run, which supports our prediction. On the other hand, the reason for the support is not what we expected. We expected the superiority of the believers to be mediated by their calling fewer repeats than skeptics. Although this was the case, the difference was miniscule and unlikely to have been of any practical consequence. The superior hitting of believers in the RA run seems, on the other hand, to be due to the fact that their dearth of repeats produced an increase in hits. Why skeptics' almost equal dearth of repeats did not translate into more hits for them remains a mystery.

A weakness in the design of this experiment is that the random run was always the last nonfeedback run. This ordering was necessary to properly test the DAT hypothesis. Because of this design feature, the dropoff of scoring from the 1st 2 runs to the 3rd, such as it was, could be at least partly attributable to a decline effect. We think this factor is unlikely to influence our interpretation of the results in this section. First, there was an increase in scoring from the 1st (M = 25.91; SD = 5.94) to the 2nd (M = 26.54; SD = 5.80) run. A decline effect interpretation would assume an ESP process; otherwise, the theoretical chance value would be an appropriate baseline against which to compare the results of the 1st 2 runs. We have yet to find any significant evidence of ESP defined by explicit guessing accuracy in any of our ISL experiments. Also, the correlations between response bias and hit rate favour response bias rather than ESP as the proper explanation of the above-chance scoring in Runs 1 and 2.

BELIEF IN THE PARANORMAL

This is the third experiment in which the hit rates of believers and skeptics were virtually identical with random targets, thus providing no support for the traditional sheep-goat effect. However, in the present experiment sheep did produce significantly higher variance in their scores than did skeptics. Palmer (1972) has previously summarized data indicating high between-subject variance among extreme sheep in standard forced-choice ESP tests. Note that in the present experiment our sheep sample was selected to consist of strong believers.

This is also the third straight experiment in which skeptics demonstrated the AAE positively and believers negatively. We have no good explanation for the success of the skeptics, except for the possibility that, unlike in most experiments on which the evidence for the sheep-goat effect is based, the Ps did not know that E was a believer. Given the context in which the experiment took place, they might well have assumed he was a skeptic. A belief that E is a skeptic could have increased the comfort level of the skeptics and decreased the comfort level of the believers, thus causing the scoring reversal. The last few Ps in the present experiment were asked to predict at the end of the session the attitude of E toward the paranormal; they tended to assume that his attitude was the same as theirs. Although this sample is too small to draw any firm conclusions, this result does not support the above speculation.

We still need to explain why this reversal of the sheep-goat effect did not apply to the DAT effect and the results of the RA run. Believers produced significantly better results than skeptics in both these cases. The discomfort hypothesis can make at least some sense of this. Although both the AAE and the DAT effect fall under the rubric of implicit psi, the DAT effect is more strongly implicit. According to our interpretation of the AAE, Ps were responding to the targets in Run 3 just as they would in an explicit psi task. They simply were getting information about the Run 4 targets and mistakenly assumed that this information applied to the Run 3 targets. The DAT test was completely implicit in that Ps had no awareness that that the timing of their mouse clicks was of any relevance whatsoever to their task. It may be that implicit psi in the strong sense of the term is less susceptible to discomfort than explicit psi; in other words, it is only the conscious mentation that the discomfort affects. This conclusion at least seems intuitively plausible. Although it is not entirely clear what mechanisms were operative in the RA Run for skeptics, for believers the mechanism seems to involve calling a small number of repeats, which is simply a maintenance of their natural bias. Discomfort would be expected to actually stabilize this bias; it is the spontaneity necessary to break the bias that discomfort would more likely compromise.

Finally, the current results directly contradict our original hypothesis that believers should be better at ISL than skeptics. The results with the LIMBEX scale suggest moreover that the cognitive deficits characteristic of many believers actually inhibit ISL. The temporal lobe scale from which the LIMBEX was derived (Makarec & Persinger, 1990) has been shown to correlate highly with the Magical Ideation Scale (MIS), which is closely associated with belief in the paranormal (Brugger & Graves, 1997). Our finding of a strong difference in LIMBEX scores between believers and skeptics is quite consistent with these previous findings. The MIS is also reflective of schizotypy (Eckblad & Chapman, 1983), which in turn reflects the kind of global, right-hemisphere cognitive processing that we had assumed would facilitate ISL. We found exactly the opposite to be the case, and these data must be considered as evidence against our proposed explanation for ISL success.

REFERENCES

Alter, I. (1989). A cerebral origin for „directionality“. *Neuropsychologia*, **27**, 563-573.

Alter, I., Rein, S., & Toro, A. (1989). A directional bias for studies of laterality. *Neuropsychologia*, **27**, 251-257.

Bem, D. J. (2003). Precognitive habituation: Replicable evidence for a process of anomalous cognition. *Proceedings of Presented Papers: The Parapsychological Association 46th Annual Convention*, 6-20.

Bem, D. J., & Honorton, C. (1994), ‘Does psi exist? Replicable evidence for an anomalous process of information transfer’, *Psychological Bulletin*, **115**, 4-18.

Bierman, D. J., & Radin, D. I. (1997). Anomalous anticipatory response on randomized future conditions. *Perceptual and Motor Skills*, **84**, 689.

Blackmore, S., & Moore, R. (1994). Seeing things: Visual recognition and belief in the paranormal. *European Journal of Parapsychology*, **10,** 91-103.

Blackmore, S. J., & Troscianko, T. (1985). Belief in the paranormal: Probability judgements, illusory control, and the "chance baseline shift." *British Journal of Psychology*, **76,** 459-468.

Bornstein, R. F., & Pittman, T. S. (1992). *Perception without awareness*. London: Guilford.

Bressan, P. (2002). The connection between random sequences, everyday coincidences, and belief in the paranormal. *Applied Cognitive Psychology*, **16,** 17-34.

Broughton, R. (1994). Repetition avoidance in ESP tests: do sheep and goats really differ? [abstract]. In E. W. Cook & D. L. Delanoy (Eds.), *Research in parapsychology 1991* (pp. 147-150). Metuchen, NJ: Scarecrow Press.

Brugger, P. (1997). Variables that influence the generation of random sequences: An update. *Perceptual and Motor Skills*, **84**, 627-661.

Brugger, P., & Baumann, A. T. (1994). Repetition avoidance in responses to imaginary questions: The effect of respondents' belief in ESP. *Perceptual and Motor Skills*, **75**, 883-893.

Brugger, P., & Graves, R. E. (1997). Testing vs. believing hypotheses: Magical ideation in the judgment of contingencies. *Cognitive Neuropsychiatry*, **2**, 251-72.

Brugger, P., Landis, T., & Regard, M. (1990). A "sheep-goat effect" in repetition avoidance: Extra-sensory perception as an effect of subjective probability? *British Journal of Psychology*, **81**, 455-468.

Brugger, P, Monsch, A. U., Salmon, D. P., & Butters, N. (1996). Random number generation in dementia of the Alzheimer type: A test of frontal executive functions. *Neuropsychologia*, **34**, 97-103.

Brugger, P., Regard, M., Landis, T., Cook, N., Krebs, D., & Niederberger, J. (1993). "Meaningful" patterns in visual noise: Effects of lateral stimulation and the observer's belief in ESP. *Psychopathology*, **26**, 261-265.

Brugger, P., Regard, M., Landis, T., & Graves, R. E. (1995). The roots of meaningful coincidence. *Lancet*, **345**, 1306-1307.

Brugger, P., & Taylor, K. I. (2003). ESP: Extrasensory perception or effect of subjective probability? *Journal of Consciousness Studies*, **10**(6-7), 221-246.

Colwell, J., Schroder, S., & Sladen, D. (2000). The ability to detect unseen staring: A literature review and empirical tests. *British Journal of Psychology*, **91 (Pt 1),** 71-85.

Dobyns, Y. H., & Nelson, R. D. (1998). Empirical evidence against decision augmentation theory. *Journal of Scientific Exploration*, **12**, 231-257.

Eckblad, M., & Chapman, L. J. (1983). Magical ideation as an indicator of

Schizotypy. *Journal of Consulting and Clinical Psychology*, **51**, 215-225.

Fletcher, P. C., Zafiris, O, Frith, C. D.. Honey, R. A. E., Corlett, P. R., Zilles, K., & Fink, G. R. (2004). On the benefits of not trying: Brain activity and connectivity reflecting the interactions of explicit and implicit sequence learning. *Cerebral Cortex*, **15**, 1002-1015.

Hake, H. W., & Hyman, R. (1953). Perception of the statistical structure of a random series of binary symbols. *Journal of Experimental Psychology*, **45**, 64-74.

Houtkooper, J. M., & Haraldsson, E. (1997). Reliabilities and psychological correlates of guessing and scoring behavior in a forced choice ESP task. *Journal of Parapsychology*, **61**, 119-134.

Krummenacher, P. (2003). Predicting the future from the past: Paranormal belief, dopamine, and implicit sequence learning. Unpublished Masters Thesis, University of Zürich, Switzerland.

Jones, M. R. (1971). From probability learning to sequential processing: A critical review. *Psychological Bulletin*, **76**, 153-185.

Lawrence, T. R. (1990/91). Subjective random generations and the reversed sheep-goat effect: A failure to replicate. *European Journal of Parapsychology*, **8**, 131-144.

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

Luh, K. E. (1995). Line bisection and perceptual assymetries in normal individuals: What you see is not what you get. *Neuropsychology*, **9**, 435-448.

MacDonald, A. P., Jr. (1970). Revised scale for ambiguity tolerance: Reliability and validity. *Psychological Reports*, **26**, 791-798.

Makarec, K., & Persinger, M. A. (1990). Electroencephalographic validation of a temporal lobe signs inventory in a normal population. *Journal of Research in Personality*, **24**, 323-337.

Marsaglia, G., & Zaman, A. (1987). Toward a universal random number generator (FSU-SCRI-87-50). Gainesville, FL: Florida State University.

May, E. C., Spottiswoode, S. J. P., Utts, J. M., & James, C. L. (1995a). Applications of decision augmentation theory. *Journal of Parapsychology*, **59**, 221-250.

May, E. C., Utts, J. M., & Spottiswoode, S. J. P. (1995b). Decision augmentation theory: Applications to the random number generator database. *Journal of Scientific Exploration*, **9**, 453-488.

Mohr, C., Graves, R. E., Gianotti, L. R. R., Pizzagalli, D., & Brugger, P. (2001). Loose but normal: A semantic association study. *Journal of Psycholinguistic Research*, **30**, 475-483.

Musch, J., & Ehrenberg, K. (2002). Probability misjudgment, cognitive ability, and belief in the paranormal. *British Journal of Psychology*, **93**, 169-177.

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*, **66**, 1-26.

Palmer, J. (1971). 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*, **65**, 373-408.

Palmer, J. (1996). Evaluation of a conventional interpretation of Helmut Schmidt's automated precognition experiments. *Journal of Parapsychology*, **60**, 149-170.

Palmer, J. (1997). Hit-contingent response biases in Helmut Schmidt's automated precognition experiments. *Journal of Parapsychology*, **61**, 135-141.

Palmer, J. (1998). ESP and RNG PK with Sean Harribance: Three new studies. *Proceedings of Presented Papers: The Parapsychological Association 41st Annual Convention*, 124-134.

Radin, D. I. (1997). Unconscious perception of future emotions: An experiment in presentiment. *Journal of Scientific Exploration*, **11**, 163-180.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, **6**, 855-863.

Roll, W. G., & Klein, J. (1972). Further forced-choice ESP experiments with Lalsingh Harribance. *Journal of the America n Society for Psychical Research*, **66**, 103-112.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, **274**, 1926-1928.

Schienle, A., Vaitl, D., Stark, R. (1996). Covariation bias and paranormal belief. *Psychological Reports*, **78**, 291-305.

Schmeidler, G. R. (1986). Subliminal perception and ESP: Order in diversity? *Journal of the American Society for Psychical Research*, **80**, 241-264.

Schmeidler, G. R., & McConnell, R. A. (1973). *ESP and personality patterns*. Westport, CT: Greenwood. (Original work published 1958)

Schmidt, H. (1969). Precognition of a quantum process. *Journal of Parapsychology*, **33**, 99-108.

Schmidt, S., Schneider, R., Utts, J., & Walach, H. (2004). Distant intentionality and the feeling of being stared at: Two meta-analyses. *British Journal of Psychology*, **95**, 235-247.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science*, **275**, 1599-1603.

Thalbourne, M. A., & Delin, P. S. (1993). A new instrument for measuring the sheep-goat variable: Its psychometric properties and factor structure. *Journal of the Society for Psychical Research*, **59**, 172-186.

Towse, J. N., & Neil, D. (1998). Analyzing human random generation behaviour: A review of methods used and a computer program for describing performance. *Behaviour Research Methods, Instruments, and Computers*, **30**, 583-591.

Tune, G. S. (1964). Response preferences: A review of the relevant literature. *Psychological Bulletin*, **61**, 286-302.

Wiegersma, S. (1982). Can repetition avoidance in randomisation be explained by randomness concepts? *Psychological Research*, **44**, 189-198.

PROJECT STAFF

John Palmer, Ph.D. (Co-principal Investigator): Recruited and tested participants, analyzed the data. Major role in designing the experiment and writing the report. Contributed to software development.

Peter Brugger, Ph.D. (Co-principal Investigator): Contributed to design of the experiment and writing of the report.

Enrique Wintsch, Liz. Phil.: Major role in software development.