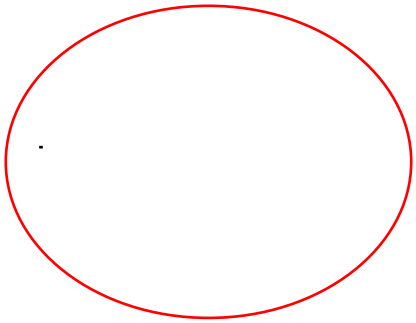
USING INTERNET TO STUDY ANOMALOUS COGNITION:
Getting rid of noise in a noisy environment

Dick Bierman
Rens Wezelman

Universities of Utrecht & Amsterdam

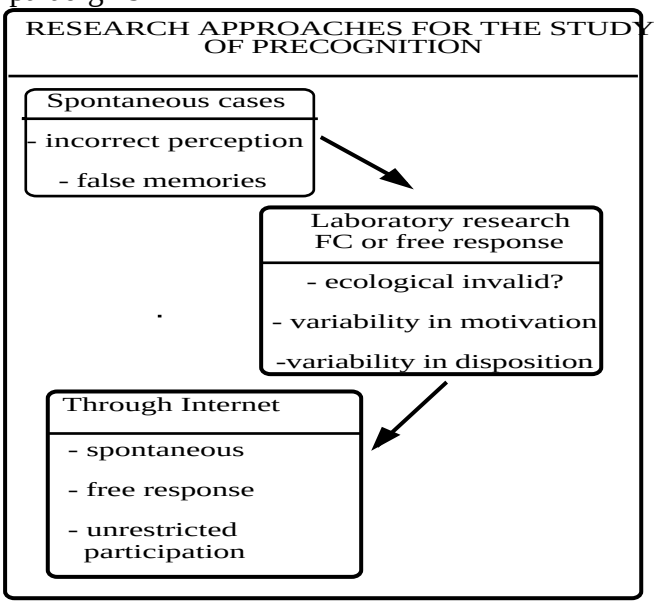


e-mail: bierman@psy.uva.nl
URL: http://www.psy.uva.nl/anomal

the Electronic Journal of Anomalous Phenomena:
http://www.psy.uva.nl/eJAP

The phenomenon of precognition was originally studied by detailed investigations of spontaneous cases. The major research question concerned the veridical nature of the phenomena. In other words were these phenomena real?

SHEET 1: 3 major paradigms

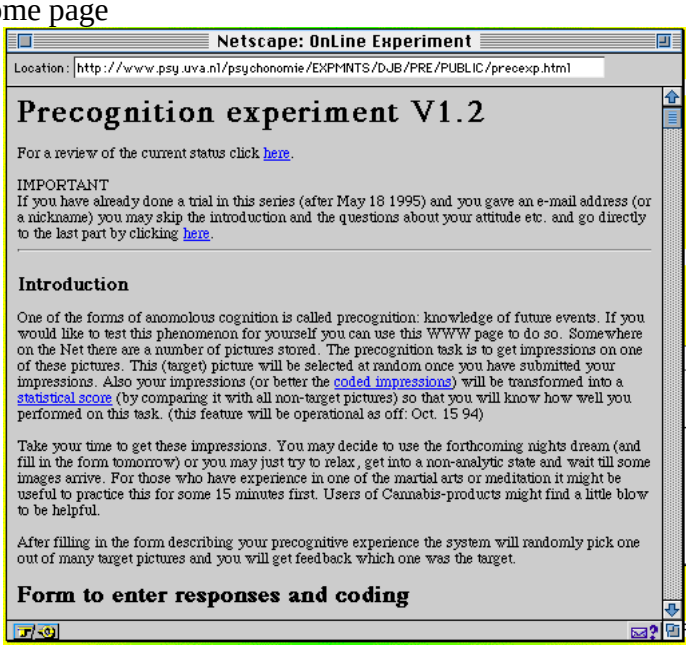


The reported phenomena were at times striking but the problem for the researchers was that it concerned reports based upon human perception and memory both of which are fallible. So the data were intrinsically noisy. For this reason Rhine started to investigate precognition under well controlled conditions in the laboratory. The effects found in that card guessing work were small and even after a research program of 30 years no firm

conclusions about the nature of the phenomena could be drawn. Even the very reality of the phenomena was and is still doubted by mainstream science. It was felt that larger effects were needed to investigate the underlying processes. Therefore free response techniques were introduced in the lab. Still one of the major disadvantages of the lab remained. The subjects were supposed to perform upon request in contrast to the spontaneous nature of the phenomena in the field. Could it be that the artificial context of the lab buried the phenomenon in the noise of varying dispositions and motivations?
In 1994 we have started to explore the possibility to use Internet for doing well controlled precognition experiments outside the lab and in a way that supposedly resembled the conditions under which subjects have a spontaneous experience.
Today I will report on 4 series each covering about 6 months of data. The total data set comprises nearly 5000 sessions.

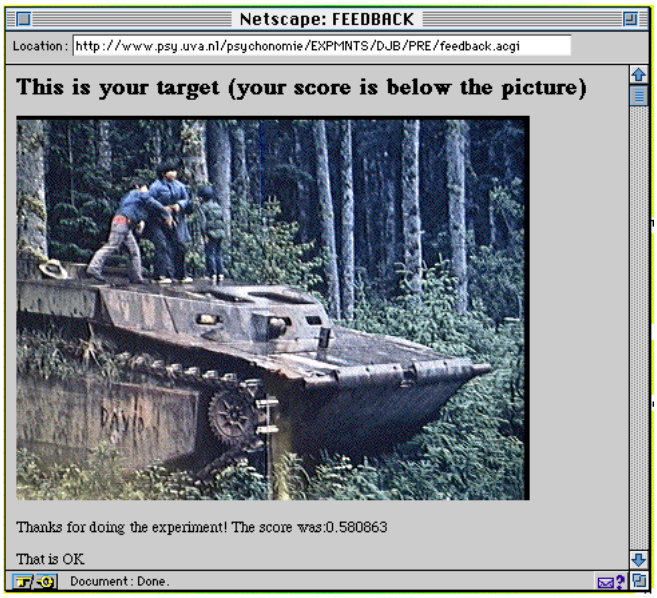
A person who connects through World Wide Web to our site is asked to participate in a precognition experiment.

SHEET 2: The home page

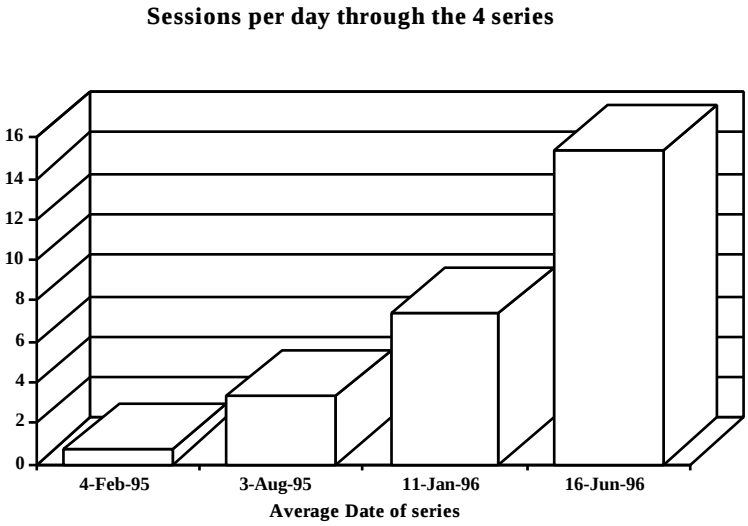


The goal of the subject is to describe a picture, submit this description on line and immediately afterwards a picture will be randomly selected from a pool of 72 diverse pictures. Before the subject can submit the picture he or she has to fill an simple electronic form asking for personality characteristics like belief in precognition, openness, artistic interests and state of consciousness during production of the description. As soon as the server receives the information, a matching score will be calculated upon which the subjects receives feedback of the picture and of this matching score.

SHEET 3: example of feedback
Here is an example of the feedback that a subject receives. It can be seen although probably not read that there is a quantitative score beneath the picture. In order to have this quantitative score the subject has to select a number of binary descriptors, fitting to his or her experience. For instance, one of the descriptors would be "is the experience mainly dark of mainly light". Another one: "Are there human figures in the experience" etc. etc.



It should be stressed at this point that ANYBODY on the Internet can participate and that we have no control whatsoever about the motivation of the participant.



SHEET 4: number of sessions for 4 series

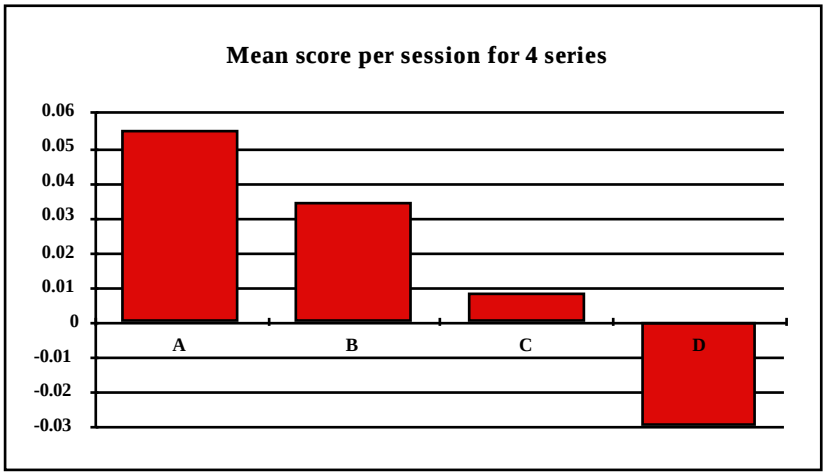
The results show a dramatic incline of no of sessions per day in the course of the 4 series. Increasing from a meagre 0.7 session per day in 1994 to a 15 per day in 1996. The global trend is a doubling every 4 months or so.

An analysis of the first series however revealed that many of these sessions were not to be taken seriously. Sometimes people submitted the same description 10 times in a row within a few minutes. Sometimes there was no verbal description or less than 3 descriptors were marked. So we had to filter these data afterwards which was a tedious procedure. Finally it turned out that 40% of the sessions had to be removed. The remaining 107 sessions gave a mean normalised score of 0.055 which was not significant.

For the subsequent series we changed the software so that if subjects submitted the forms without proper descriptions or whenever they submitted twice or more in a row

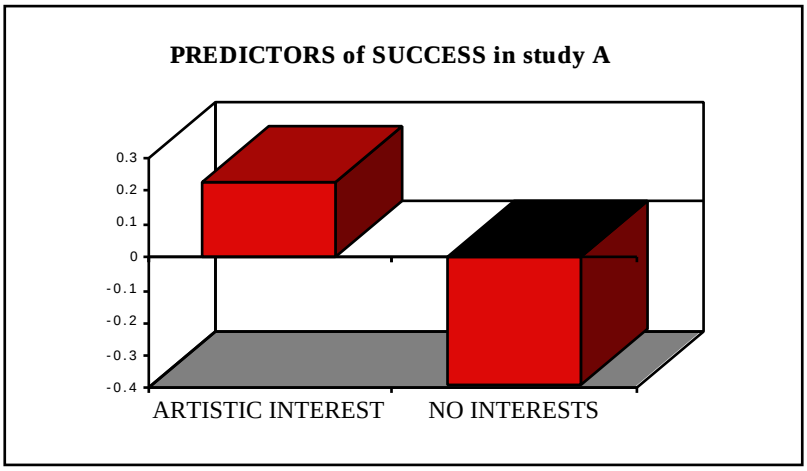
within a short period of time, they got a message that the submission was refused. Thus the tedious post hoc hand filtering was avoided.

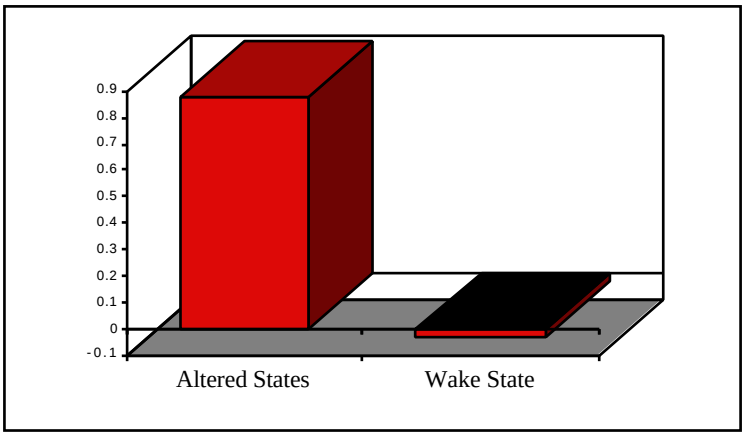
SHEET 5: mean z-score for 4 series.



Although there was a strong incline in sessions per day there was a strong decline in mean score. The last series was even marginally significant in the negative direction.
These results look rather disappointing and certainly do not support the notion that effects would become stronger if the subjects can do these experiments in their home or office environment.
It could be however that these completely uncontrolled conditions attract lots of people that just did it for curiosity reasons even if their submissions looked serious, thereby still introducing lots of noise due to varying motivations and intentions. Therefore we decided to have a closer look at the data and to see if we could find a way to filter them from this potential noise.

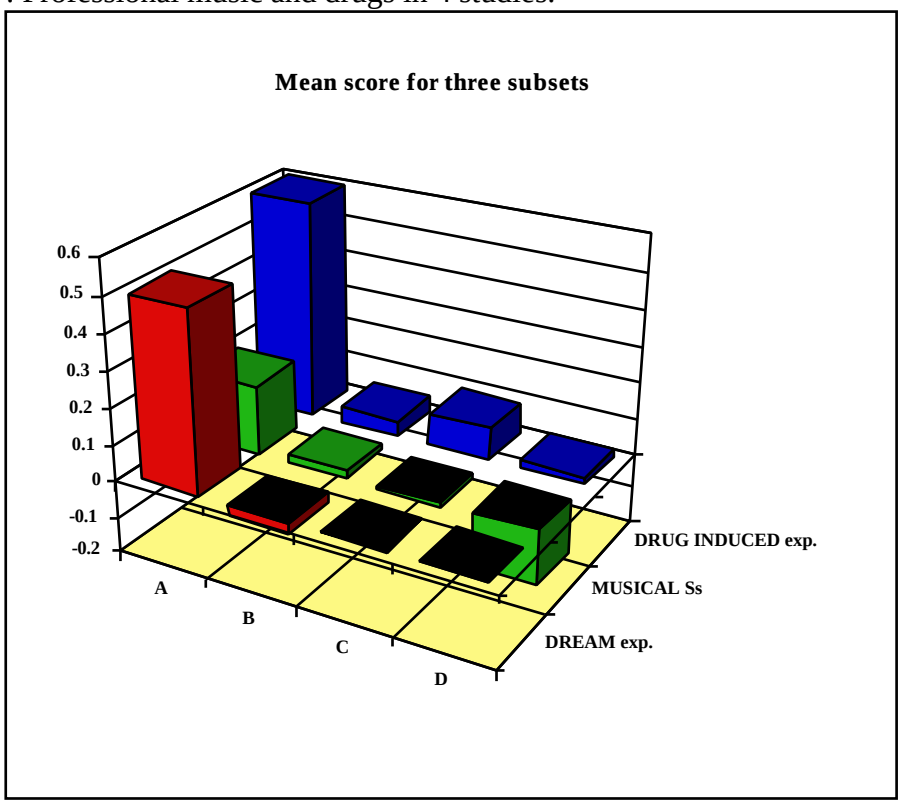
SHEET 6: internal effects of series I
An obvious way to do this was to look at the personality questionnaire. Indeed the first series showed that subjects with no artistic interest did score significantly worse than subjects with such an interest. Notably musical interest was a good predictor of performance. Also in that series subjects who had their experience in an altered state of consciousness like dream or drugs-induced did perform better than people in an ordinary State of Consciousness. Remarkably belief and openness did not predict any of the variance.





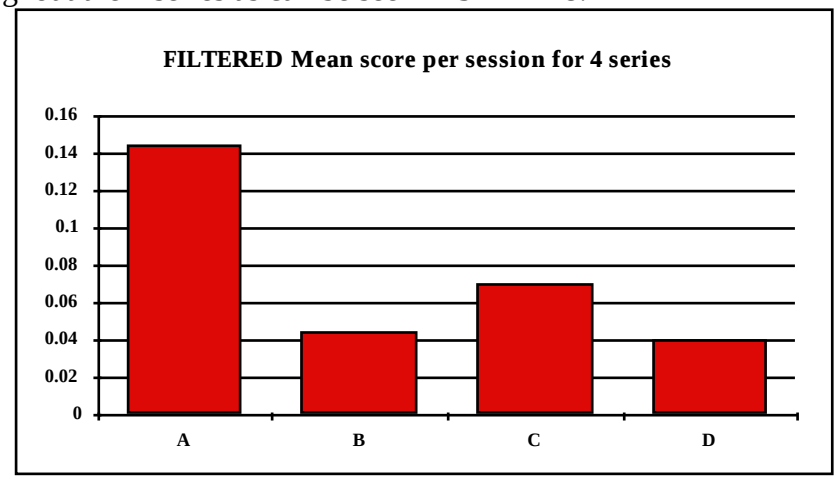
So it looked that we might be able to filter the data on the basis of the questions about artistic interest and state of C. However it turns out that these internal effects generally did not hold up in the subsequent series . For instance the dreamers did not perform at all after the first series. The effect of musical interest disappeared apart from the fact that professional musicians did well. The only predictor that gave consistent but weaker effects was the drugs induced state.

SHEET 7: Professional music and drugs in 4 studies.



Since it appeared that filtering based upon the questionnaire did not work too well i.e. did not yield consistent results over the 4 series, we decided to explore if it was possible to filter based upon purely statistical considerations. If one has as many data as we had, the data will theoretically follow closely a normal distribution with the mean of about 0 (as we have seen) and of course the distribution should be completely symmetrical under the assumption that we have to do with a pure chance distribution. Standardised scores smaller than 1.5 are statistically generally seen as non significant and therefore we removed all sessions which had an absolute z-score of 1.5 or less. The remaining distribution should still be symmetrical and have an average close to 0. However it turns out that removal of the middle part of the distribution gives us a more or less consistent

result throughout the 4 series as can be seen in SHEET 8.



The implication is that at the positive side of the distribution there are a few more sessions. Their contribution gets buried in the noise of the mediocre sessions. But when one removes these average sessions the few outstanding sessions appear out of the noise. The average effect size thus obtained is around 0.05. This is nearly identical with the effect sizes found in card guessing work. So either there is still lots of noise around or we have to give up the idea that subjects who participate more or less spontaneously will do better than in the artificial lab environment.

Should we now quit this work because it did not live up to its expectations? I think the answer should be a clear NO. Although about 9 out of 10 sessions are filtered out in the latter approach, we see that the Internet is growing so fast that we expect to have more than 100 sessions per day in the near future.

With so many contributions that come in without any effort by experimenters we can get a large data pool enabling us to do the analysis of the personality factors on a selected, less noisy, data set.

Conclusions

1. Only consistent results after statistical filtering

2. Effect size around 0.05 (= card guessing) after removal of 90% of the data.

3. 10000 sessions needed for process oriented analyses.

4. With current Internet growth 10000 sessions reached in 1997.

Thus we may get some insight in the factors that are relevant in the psi-process. It should be remarked that from experience in the until recently classified remote viewing research we now know that a selection of subjects was done which resulted in 1 useful subject out of 100 pre-tested subjects. So it wouldn't surprise us if in subsequent analysis of the Internet data we will go further and refine the filtering procedure till we reject 99 out of 100 sessions. But still then we will have a large data set to do theory

driven analysis with.
The current results do suggest that we add questions which would help us to further filter the data, like "are you doing this session from home or from your office". Finally it is our intention to apply neural network modelling on the future data after the preliminary filtering has removed part of the noise.