Paper submitted for the Parapsychological 43rd Annual Convention

Backward Causation and the Hausdorff-Dimension of Singular Events

Walter v. Lucadou, WGFP, Freiburg i. Br.

Abstract:

The theoretical paper deals with the finding that the effectsize of nearly all psi-experiments shows a tendency to decline over time. Thus one gets the impression that experimental psi is not only a very small effect but also restricted in relation to repeatability. In spontaneous cases, however, large effect seem to occur. This seemingly paradoxical situation can be explained by applying the Model of Pragmatic Information (MPI) to backward causation and introducing a fractal dimension of time. For some special examples a precise mathematical definition is provided. Some experimental data are supportive to this approach.

1. Introduction

In a recent meta-analysis Dick Bierman (2000) shows that nearly all psi-experiments which had been performed since the days of J.B. Rhine exhibit a significant inter-experimental decline-effect. For such a large database including a large number of different experimental designs and settings it seems highly improbable that this result could be explained by psychological factors such as loss of motivation, exhaustion or experimenter expectation. Even though this psychological interpretation is in principle unfalsifyable it is reasonable and legitimate to assume that the inter-experimental and possibly also the intra-experimental decline-effect exhibits an essential characteristics of psi-phenomena.

Moreover, certain theoretical models (the so-called observational theories, OTs) predict such intrinsic decline effects. Within this class of models for psi it is assumed that (human) observation of the experimental results (i.e. feedback) is responsible for deviations from certain expectation-values of the random process (target). There are two possible intrinsic mechanisms which lead to decline effects. The first one, which will not be discussed here in detail, is the so-called divergence effect. It is assumed that future observers (such as readers of the published results) might blur out the "psi-source" of the initial operator (subject or experimenter). This model needs further assumptions about future observers. The second mechanism for decline is more fundamental, because it has to do with the notion of time and timely order. In this paper we discuss the problem only in the context of one specific candidate of the OTs, namely the Model of Pragmatic Information (MPI) (see Lucadou 1995a). It is true, that there is still no definite conclusion whether metaanalyses exhibit inter-experimental decline effects for all kinds of psi-experiments. But from the point of view of the MPI this is comprehensible, because in the MPI each experiment is considered as a *unity* which reflects the meaning of the experimental situation and which cannot be divided in parts or accumulated with other experiments without essential loss - at least as far as replications are not identical. This is in agreement with Z. Vassy's finding that the psi-effect is distributed in a holistic way over the whole run (see Vassy 1990).

2. Backward causation

Decline over many experimental studies means that these studies cannot be regarded as independent. This is of course not jet a model but just a more general description. It is normally assumed that the mechanism underlying time dependent series is that previous events influence later events. In parapsychological experiments, however, such an influence cannot be assumed prima-facie because under the null-hypothesis any influence on the random target sequence is ruled out by experimental conditions. In the OTs an anomalous influence is assumed, which, however, comes from future events, because it results from the observation of the feedback. This means that the timely order is reversed in this case. Some have introduced the term "backward causation" for this situation. There is also some direct empirical evidence that backward causation really exists, because pk-experiments with prerecorded targets (PRT) turned out to show effect-sizes of the same magnitude as "normal" pk-experiments. The concept of backward causation could also be useful to describe precognition (just as the opposite of the coin).

The MPI assumes that psi is a non-local process (or more precisely a non-local correlation) and thus intrinsically includes backward causation. The question we have to ask now is: What is a physical effect which depends from future instead of from the past.

3. What is a physical effect?

Systemtheoretically speaking any physical effect (E) can be defined as an information about a physical system gained by a measurement. Information is defined as a signal (S) which enables us to decide at least two alternatives, if a certain context (C) is given. S could be for instance a random sequence including a part with a deviation from randomness (e.g. noisy signal, or a fluctuation). C gives a criterion how the deviation has to be interpreted (e.g. a pk-effect is present). Note that S alone would not yield any information. Thus a physical effect always needs not only a signal S but also a context or a criterion C for its existence. The two alternatives are defined by the decision whether an expected effect is present or not present.

In this sense, a correlation K between the variable A and B is no simple "physical effect", because the measurement of A and B alone does not yet allow to decide whether the correlation K is present or not. To do this, both signal S(A) and S(B) have to be "combined" or "compared" to establish the correlation K. This is of crucial importance for our problem of backward causation. Formally we can write the situation as follows:

In the case of a normal causal link we get the timely order:

S -> C -> E   (-> indicates the timely order)

Signal S (and to some extend criterion C) can be considered as "causa efficiens" (in the Aristotelian sense) for the effect E

In the case of backward causation the time order is reversed at least concerning the criterion C to establish an effect E

S <- C -> E

Here C can be considered as "causa finalis" (dto) of E

Finally we have the situation of (non-local) correlations K. Here, a timely order cannot be defined anymore:

[S(A), S(B)] -> K   ([ ] indicates "comparison")

Nevertheless K is sometimes interpreted as an effect E.

4. Physical effects and backward causation

If backward causation would lead to a real physical effect this would enable us to build an "oracle" which could be used to create an intervention paradox:

```
          ----------------------> time
              Preinspector
                   C
 ‖ Target S    ‖---------‖ Storagesystem ‖----‖ Operator ‖--->-
       |           |                                         |
       ------------V-----------<-----------------------------
  "oracle" E
```

The "oracle" (E) would be a significant deviation of an random sequence from the null-hypothesis in an PRT-experiment (S) which operationalizes backward causation. If the criterion C (C: $Z > Z_{crit}$, Z means Z-score) is fulfilled it is decided by the "preinspector" (for instance by a computer) that the random sequence (S) will not be used for the subject. This is of course paradoxical because the operator will not be able to exert an influence E on the sequence, which however, was the reason for the selection.

The MPI starts from the basic assumption that nature does not allow (intervention) paradoxes. In the MPI this is formulated as the "two fundamental laws of parapsychology":

I. Psi-phenomena are non-local correlations in psycho-physical systems which are induced by the pragmatic information which creates the system (organizational closure, OC).

II. Any attempt to use a non-local correlation as a signal transfer (= physical effect in the sense of "causa efficiens") makes the non-local correlation vanish or change.

One way how backward causation could fulfil law II. (or avoids the intervention paradox) is that "nature" discriminates the signal S such that the criterion C cannot be reached. In parapsychgology the criterion C is usually the Z-score:

$$Z = (T - n*p)/s$$

T means "hits", n = number of trials, p = probability for a hit, s is the standard deviation:

$s = \sqrt{(n*p*(1-p))}$, $\sqrt{}$ means square root

If one defines the effectsize E as: $E = (T-n*p)/n$, one obtains a critical effectsize E_{crit} which cannot be surmounted. If n increases in a single experiment or in a series of identical experiments E must decline with n:

$$E < E_{crit} = const/\sqrt{(n)}$$

The value of const may depend from the experimental setting and/or psychological conditions and is not specified here. For n = 1 we get the maximal effect. We call such a situation a "Singular Event" (SE). The term "singular" refers to the term "singular point" in mathematics, where it describes a singularity of a function. For larger n we obtain an increasing decline effect, especially if we combine identical experimental runs (see Lucadou 1995a). We will see later that singular events are somewhat different from single events. In statistical experiments singular events do not play a role because the power of any statistical test reaches its minimum for n = 1. In spontaneous cases, however, the situation is quite contrary. Here mostly singular events are observed.

Usually it is assumed that the validity of singular events does not depend on the fact, that it occurs only once, but on the quality of its documentation (QD). It is clear that in this situation E_{crit} is no good criterion for the second law (II.). In this case the MPI predicts that the quality of documentation QD is restricted (QD_{crit}) by the following formula:

$QD_{crit} * E < OC$ (OC means organizational closure of the System)

In recent experiments other criteria than Z-scores have been used to validate psi-effects (see Lucadou 1986, Pallikari 1999, Atmannpacher 2000). In these cases E_{crit} will probably show a different functional dependency of n, however, it must be a function which decreases with n. Further research is needed here.

It is important to mention that the MPI assumes that the second law does not depend on the subjectivistic view, whether one actually uses the criterion C to create an intervention paradox or not. Whenever it is operationally possible to use the criterion C the decline occurs. If, however, the experimental design is such, that this is operationally not possible, no decline effect will occurs. For instance a randomized matching could prevent the "preinspector" to use the criterion C.

5. What is an event?

In science we mainly have to do with experimental events which are in most cases statistical, which means that many measurements of a prepared system are taken. "Prepared" means, that the system is not in an natural context but a given experimental setting. The events are prepared to be operational, which means observational and documentable. We call such events: *stochastic events*.

Sometimes, but not very often, the signal to noise ratio is such that repeated measurement is not necessary. We call such events: *experimental events*.

Most events that occurs in everyday life, however, are not experimental, they are not prepared, but very often they are preparable. Otherwise we could not plan or control anything. They can also be observed and even if they are single events we have no principle problem with documentation. We call such events: *regular events*.

There is sometimes a problem, if such regular events occurs very seldom, like eclipses, meteorites etc. In this case, public and science is inclined to neglect them if they are not generally accepted, because they occur so spontaneously that it is difficult to be prepared for a proper observation and documentation. We call such events: *rare events*.

Many persons believe that spontaneous paranormal events (SPE) are rare events, however this is not true as many representative inquiries have shown. It is also not true that they are not accepted because they are theoretically not explained. This may be true for the small group of scientists, but today most ordinary people believe that paranormal phenomena exist. However, there is a problem with observation. Paranormal phenomena (SPE) seem do avoid observation (see Lucadou 1995b). We call them *elusive events*.

Finally *singular events*, as defined above, may be part of all sorts of events defined beforehand such as stochastic, experimental, regular, rare or elusive events. The property which defines them is that they are part of a series of events which belong together or, to be more precise, which correlates within a group of events. Thus one can define a transition-probability from the singular event to any other members of the group. If psi-experiments are not independent in time - as we have assumed above - each trial is a singular event. A good example from mathematical statistics for such a group of events are Markov-chains. They are described by transition matrices. One should mention here, that not all time series with transition probabilities from one event to others are necessarily Markov-chains, for instance, if the transition-probabilities itself depend on time. In this case one could, however, as an approximation use combinations of Markov-chains with different transition-matrixes and time-scales. For the purpose of discussion we start with simple binary Markov-chains.

Normal random sequences can be regarded as degeneration of Markov-chains. They show always the same transition probability regardless which event just occurred. For such sequences the events E_i and E_j at different point of time i,j are independent:

$$p(E_i,E_j) = p(E_i) * p(E_j)$$

6. How to produce randomness?

In statistics it is assumed that we can enlarge effects by increasing the number n of stochastic events. Especially in parapsychology the so-called Rhinean paradigm started with the assumption that psi-effects can be accumulated statistically. However, the inter- and intra-experimental decline effect, discussed above, threatens this paradigm. Nevertheless there are also some empirical hints that psi may be enhanced under certain conditions. Psychological conditions like psi-conducive state, however, are not meant here, they cannot be easily accumulated in a statistical manner. Here, we deal with the procedure how random events are generated in psi-experiments. Normally physical random event generators REGs are used to produce the target sequence of psi-experiments. In most cases these REGs produce pure chance results with a given target-probability e.g. $p = 1/2$. There are only a few studies which deal with the question, whether the probability p or the method of random generation has an influence on the effect-size of psi. Schmidt found that the complexity of the REG does not influence the psi-effect. The target-probability seems not to have a dramatic effect too. Most researchers have concentrated on the requirement that the REG used in the experiment should be free of bias and other "deficiencies" in order to rule out statistical artefacts, because psi is only defined by exclusion. (Here we do not consider studies with pseudo-REGs because in this case it not clear, whether the usual statistical evaluation techniques are valid anymore and how the effect-size has to be interpreted, see Krengel 1979). There exists only one study (at least to my knowledge) where a different type of REG namely a Markov-REG was compared with the usual REGs (see Lucadou 1986). The difference between the Markov-REG and the usual one was extremely significant and at that time totally counter-intuitive.

The "normal" REG was a binary random-sequence with the target-probability $p = 1/2$. The Markov sequences was produced in the following way: The pulse rate R_i of a radioactive SR90 decay is measured after a fixed interval at a certain instance i. R_i is compared with the pulse rate R_{i-1} of the previous instance i-1.

If $R_i < R_{i-1}$ then a miss "0" is generated
If $R_i > R_{i-1}$ then a hit "1" is generated
If $R_i = R_{i-1}$ then the target generation will be repeated, which means that this case will be ignored.

Since the variance of R_i is large enough the last case only occurs rarely.

It can be shown (see Lucadou 1986) that the resulting random sequence is a Markov-chain of first order which is specified by the following transition matrix:

$$\mathbf{M}_{ij} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix}$$

p_{00} describes the probability to get a "0" after a previous "0" and so on. It is remarkable that these values hold exactly and do not depend on the form of the distribution of the initial random process. The only requirement which must be fulfilled is that the single events of the source are stochasticly independent. The transition matrix completely describes the Markov-chain. Especially, it follows that the probabilities of a hit and a miss in the whole sequence are equal:

$p_0 = 1/2 \quad p_1 = 1/2$

The distribution p(n,T) of hits T in such a Markov-chain of a given length n can be calculated from an algorithm (see Lucadou 1986). For n = 10 this distribution is very similar to a normal distribution (Gaussian) with p = 1/2 and $s = \sqrt{(n/12+1/6)}$. For large n one can even use $s = \sqrt{(n/12)}$. The advantage of the Markov-sequence compared with a normal random sequence is, that a "hit" is directly linked with a physical variable (e.g. decay-rate). A sequence of hits means a momentarily increasing decay-rate.

Since the variance (standard deviation, $s = \sqrt{(n/12)}$) for the Markov REG is smaller than for normal REGs ($s = \sqrt{(n/4)}$) it was assumed that the psi-effect would be smaller. In parapsychology and especially in the OTs it seems plausible that divergent processes can more easily be affected by pk. However, the opposite result was obtained: The Markov-REG turned out to be more than twice as effective (or sensitive) as the normal one. (The sum of all significant correlation coefficients between psychological and physical variables was 2.2 times larger for the Markov-REG than for the normal one, details see Lucadou 1986).

7. Haussdorff dimension

In mathematics there exist a generalisation of the term dimension known from geometry. In Euclidean geometry we know one -, two -, and three-dimensional objects such as line, plane, and cube. If we define a as the number of parts we need to produce the same enlarged object, and m as the magnifying- or scaling factor from the initial to the enlarged object, the Hausdorff-dimension D of the object is defined as:

$$a = m^D \quad \text{or} \quad D = \log a / \log m$$

As examples may serve:

Straight line:	a = 3,	m = 3	: D = 1
Square:	a = 9,	m = 3	: D = 2
Cube:	a = 27,	m = 3	: D = 3

For fractal objects like e.g. the Koch-curve (_/_) we get rational numbers for D:

Koch-curve:	a = 4,	m = 3	: D = 1,262

To obtain the Hausdorff-Dimension of an object empirically, one uses a lattice of the width ε which contains the object and counts the number N(ε) for smaller and smaller ε. Than the negative slope:

$$D = - \log N(\varepsilon) / \log \varepsilon$$

gives the Hausdorff-dimension of the object.

8. Scaling events

Normally natural timely ordered events (for instance a random sequence) cannot be "enlarged" like a film in slow motion. However this becomes possible to a some extend if one does not consider singular events themselves but their transition matrices.

The transition probability of a Markov-chain starting with the singular event i to the event j is given by the transition matrix $\mathbf{M}_{i,j}$. It can be calculated from $\mathbf{M}_{i,i+1}$ by the following rule:

$$\mathbf{M}_{i,j} = \mathbf{M}_{i,i+1}^{(j - i)}$$

The power is defined by the usual matrix multiplication applied j-i times.

For a normal binary random sequences $\mathbf{M}_{i,i+1} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \mathbf{M}_0$.

In this case $\mathbf{M}_{i,j}$ always remains $\mathbf{M}_0$. This means that it is the same for all singular events and does not change with increasing j. From this point of view a normal random sequence is a very static object; it has no "history", no "extension" and; no "internal connectivity".

Markov-chains, however, show such an "extension" or "internal connectivity" or "history". The transition matrix $\mathbf{M}_{i,j}$ changes from step to step with increasing j. But for certain values of p_{00}, p_{01}, ... it converges to $\mathbf{M}_0$.

As an example the transition matrix of the Markov-chain used in the experiment described above is given for 8 subsequent steps (only the first row of the Matrix: p_{00} p_{01}; is given, the second one is symmetric):

1/3 2/3; 0.556 0.444; 0.481 0.519; 0.506 0.494;
0.498 0.592; 0.501 0.499; 0.500 0.500; 0.500 0.500; ...

One can see that the elements of the matrices converge rapidly (exponentially), after 6 steps the difference to $\mathbf{M}_0$ can be neglected.

In general we can now ask the question how many steps (singular events of a Markov-chain) are needed for a given matrix $\mathbf{M}_{i,j}$ to approach $\mathbf{M}_0 \pm \alpha$ (a given error α). This sequence of d subsequent steps is equivalent to a normal random event. It can be interpreted as a virtual unity of interrelated events. We call such a sequence a "closed sequence of scaling length d (CS)". In a given run of n trials several of such CS can exist. Thus the "effective length" l of the run is only l = n/d. We define the Hausdorff-dimension D of the run in the following way:

$$D = \mathrm{Min}_{i=1,n}\, d\ (\mathbf{M}_{i,i+d} = \mathbf{M}_{i,i+1}^{\,d} = \mathbf{M}_0 \pm \alpha) \text{ with } \alpha \in \alpha,\ \alpha = 1/n$$

The Hausdorff-dimension D of a sequence is defined as the smallest number of subsequent steps d for all n singular events of the sequence, such that each element of $\mathbf{M}_{i,i+d}$ lies in the α-interval of the corresponding element of $\mathbf{M}_0$.

The value of α represents the whole run, and takes into account that in longer runs internal correlations have a longer reach, so D increases slightly with n. As result one can write:

$$\mathbf{M}_0 = \mathbf{M}_{i,i+1}^{\,D} \quad \text{for Markov-chains.}$$

Similar to the geometrical case the Hausdorf-dimension for singular events tells us, how many elements are needed to create a new "enlarged" unity.

This means, that the transition matrix $\mathbf{M}_{i,j}$ of the sequence is "compared" with the transition matrix $\mathbf{M}_0$ of a random sequence. A possible interpretation of this definition is that every singular event is not an independent event which counts for its own value, but is only a "partial" event. For a normal binary random sequence D = 1, and each "singular event" (in the limit) is independent. For the Markov-chain in our experiments given above with n = 600, $D \approx 6$. This means that each 6 subsequent trial form a CS. Thus one could also say that a singular event in the sequence is only "a sixth of an event". If such a sequence is target of a psi-effect, obviously such "partial events" do not fully contribute to the limitations which are induced by the second law (see paragraph 4.). Therefore we can reformulate the limiting formula: $E < E_{crit} = 1/\sqrt{(n)}$ by the following expression:

$$\boxed{E < E_{crit} = \mathrm{const}/\sqrt{(n/D)}}$$

For the Markov-chain given above this means that the psi effect could be larger than for a normal binary random sequence of the same length n by a factor of √(6) = 2.4. One could also say that dependent singular events are better targets for psi-effects. Further it is to be expected that the first singular event shows the highest effect-size. This could give a natural explanation for the fact, that spontaneous psi-events (SPE) seem to have a much higher effect-size than experimental events. Everyday life-events and especially SPE are normally dependent events, which are part of long histories.

Finally it may be an interesting theoretical question for further research to find out whether the distinction between the level of stochastic events, and the level of their description by transition matrixes can be interpreted as "epistemic-" versus "ontic" description in the sense Hans Primas has introduced it (see Primas 1999).

9. Conclusion

If we assume that psi-events are elusive singular events of fluctuating physical systems which are able to "connect" experiments in such a way that they are not independent among each other anymore (as discussed in 1. and 2.). This has of course consequences for the interpretation of spontaneous cases on one side, and the design of experiments on the other side.

From our considerations a natural explanation for the seemingly large effect-size in SPE emerges. SPE are interwoven with (personal) histories such that psi has enough CS "to link with". Further, the limiting laws do not apply because the events are spontaneous, or of short duration, or of poor documentation quality, and mainly elusive (especially RSPK-phenomena, see Lucadou 2000).

In principle the same applies for experiments. First of all, it seems not useful to work with "ideal" REGs anymore. One could speculate whether the decline-effect observed in meta-analyses may partly be a result of using increasingly "better" REGs. Of course one has to avoid statistical artefacts. A possible solution of this problem could be the use of Markov-REGs. However, Markov-REGs must not be build from pseudo-REGs - but this is another story, which will not be discussed here. A further experimental requirement from our consideration is that very long runs are not really helpful because due to the limiting relations (see 4.) the psi-effect would be blurred out. This could also be part of the observed decline, especially in pk-research, where the run length has become abundantly large during the last decade. Finally, it is expected that there is a optimal value for the transition probability p_{00} and/or p_{01} in binary Markov-chains, if used as a psi-target. If the probabilities would be 1/2 we have a normal binary random sequence, which, from our point of view, is not advantageous, because psi has - so to say - no "working surface". On the other hand, if the probabilities would be near 0 or 1, the Markov-chain degenerates to an oscillating sequence which also provides no "working surface" for psi. The optimum may be between this two values. But this is a question of further research.

Literature

Atmanspacher, H., Scheingraber (2000). Investigating Deviations from Dynamical Randomness with Scaling Indices. Journal of Scientific Exploration, 14, 1-18.

Bierman, D.J. (2000): The nature of anomalous phenomena: Another reality between the world of subjective consciousness and the objective world of physics? In: Ph. van Loocke (Ed.) The physical nature of consciousness. Benjamins Publ. New York (in print).

Lucadou, W.v. (1986). Experimentelle Untersuchungen zur Beeinflußbarkeit von stochastischen quantenphysikalischen Systemen durch den Beobachter. H.-A. Herchen Verlag: Frankfurt.

Lucadou, W.v. (1995a). The Model of Pragmatic Information (MPI). European Journal of Parapsychology 11, 58-75.

Lucadou, W.v. (1995b). Psyche und Chaos - Theorien der Parapsychologie. Inselverlag, Frankfurt a.M.

Lucadou, W.v. (2000). Spuk. In: Syliva M. Schomburg-Scherff und Beatrix Heintze (Hrsg.) Die offenen Grenzen der Ethnologie. Lembeck, Frankfurt a.M. S. 219-230.

Krengel, U. (1979): Über mathematische Modelle für statistische ASW-und PK-Experimente. Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 21, 157-172.

Pallikari, F., Boller E. (1999). A Rescale Range Analysis of Random Events. Journal of Scientific Exploration, 13, 25-40.

Primas, H. (1999). Basic Elements and Problems of Probability Theory. Journal of Scientific Exploration, 13, 579-613.

Vassy, Z. (1990). Experimental study of precognitive timing: Indications of a radically noncausal operation. Journal of Parapsychology, 54, 299-320.

Lucadou, W.v. (2000): "Backward Causation and the Hausdorff-Dimension of Singular Events", in: Steinkamp, F., (ed.), Proceedings of Presented Papers, The Parapsychological Association 43rd Annual Convention August 17-20, 2000, Freiburg i.Br. p. 138-147.