Homeopathy (2004) 93, 179–185

doi:10.1016/j.homp.2004.07.006, available online at http://www.sciencedirect.com

ORIGINAL PAPER

Homeopathic proving symptoms: result of a local, non-local, or placebo process? A blinded, placebo-controlled pilot study

H Walach[1,*], J Sherr[2], R Schneider[1], R Shabi[3], A Bond[4], G Rieberer[5]

[1]*Department of Evaluation Research in Complementary Medicine, Samueli Institute—European Office, Institute for Environmental Medicine and Hospital Epidemiology, University Hospital Freiburg, Germany*
[2]*39 Wells Rd, Malvern, Worcester, UK*
[3]*Homoeopathy Department, Michlala Leminhal College, Tel Aviv, Israel*
[4]*North West College of Homoeopathy, Manchester, UK*
[5]*Quarzweg 9, Hamburg, Germany*

***Background*: Homeopathic pathogenetic trials (HPTs) (provings) are the pillar of homeopathy. Symptoms experienced by healthy volunteers are used to find the correct medicine for therapy. It is unclear whether these symptoms are specific or due to placebo noise. Furthermore, it is uncertain whether proving effects, if present at all, are due to a local or non-local process**

***Objectives*: To develop a test model which allows for testing if homeopathic proving symptoms are caused by placebo or causative mechanisms, and if these symptoms are due to local or non-local processes.**

***Design*: Randomised, blinded, placebo-controlled, parallel-group study, with 1-week baseline and 2-weeks proving period.**

***Subjects*: 11 healthy volunteers from two different homeopathic schools.**

***Proving substance*: An homeopathic medicine (*Cantharis* 30c), blindly chosen from 12 potential medicines, compared to placebo.**

***Outcome measure*: Number of symptoms typical for the medicine in the experimental and control group during baseline and proving period.**

***Results*: During baseline there was no difference in the number of typical or atypical symptoms in either group. During the proving period, both more typical symptoms for *Cantharis* ($P = 0.03$) and more atypical symptoms ($P = 0.02$) were observed compared to baseline. Between-group differences were not significant. Effect sizes for the difference between the proving and control group for typical symptoms was $d = 0.4$, and for atypical symptoms $d = 0.6$.**

***Discussion*: This proving model could be valuable in studying the validity of proving symptoms of homeopathic substances in healthy volunteers.**

***Conclusion*: Homeopathic proving symptoms appear to be specific to the medicine and do not seem to be due to a local process. Since this was a pilot study using a small number of provers, rival hypotheses cannot be ruled out and the study needs replication.** *Homeopathy* (2004) **93,** 179–185.

Keywords: proving; homeopathic pathogenetic trial; local effect; non-local effect; Cantharis

*Correspondence: Harald Walach, Institute for Environmental Medicine and Hospital Epidemiology, European Section of the Samueli Institute for Information Biology, University Hospital Freiburg, Hugstetter Str. 55, 79106 Freiburg, Germany.
E-mail: walach@ukl.uni-freiburg.de
Received 26 March 2004; revised 10 June 2004; accepted 7 July 2004

Introduction

Homeopathic provings or homeopathic pathogenetic trials (HPT) are the pillar of homeopathy.[1–3] Ever since Hahnemann's ground-breaking experiment with Cinchona bark, homeopaths have used symptoms experienced by healthy volunteers after ingestion of remedial substances to determine the therapeutic effects of homeopathic medicines.

While, in the beginning, Hahnemann used mainly mother tinctures and low potencies for HPTs, he later switched to centesimal dilutions (30c), and many of his followers did the same. Most recent provings have been conducted with ultramolecular dilutions (>12c).[4] It is highly unlikely that any original molecule is present in such medicines.[5] This raises the question whether the symptoms experienced by volunteers in HPTs are mere placebo background noise (ie responses associated with placebo intake, such as expectation), or if they are different from placebo which, in turn, raises the question of the possible mechanism producing specific symptoms.

Some have advanced what we would classify as 'local hypotheses' for homeopathic medicine effects.[6,7] Although there are different types of local hypotheses, they all converge on one general principle: some pharmacologically active principle is conveyed from the remedial substance to the solvent (alcohol, water, or lactose), and conserved therein. On ingestion of a homeopathic medicine, the organism is stimulated by this 'information' and reacts accordingly.

But ever since Hahnemann, who himself called the active principle of his medicines 'spirit-like', another class of explanation for the effects of homeopathy, non-local ''hypotheses'' have been suggested, most recently by Milgrom ,[8–11] Walach,[12,13] and Weingärtner.[5] According to these authors, the action of homeopathic medicines has to be sought in something more general and cannot be experimentally isolated and pinned down to the consequence of the medicine alone. This could be the whole context of homeopathic treatment, including the patient, the practitioner and the remedy, with the remedy alone having a clear-cut effect only in this context, as proposed by Milgrom. Or it could be the remedy and the whole of its preparation history, together with the homeopathic experience that contains and produces the effect, and not only the remedy as such, as proposed by Walach and Weingärtner.

The difference between local and non-local models and their consequences for experimental testing are important. While local models predict that homeopathic medicines should be active and specific independent of their context of use, non-local models predict that homeopathic effects might develop in a more field-like manner and thus cannot be easily isolated experimentally.

Thus there are two questions at stake:

(1) Are the symptoms experienced by healthy volunteers in provings mere placebo background noise? Previous studies directed at this question yielded inconclusive results and were criticised.[14–16] Two main criticisms of these studies were that the dosage regime was fixed and that the collection of symptoms was not precise enough.[17] The dosage regime in these studies was the same for all participants and did not follow the rules set by Hahnemann and used in some recent HPTs, that dosing should be continued until symptoms develop, then stopped. This was not possible in 'proving-by-mail' studies. And the collection of symptoms used a structured diary and thus tried to reduce variability of response already at the data-collection stage, which might have missed some important information.

(2) Is there evidence that homeopathic medicines produce specific symptoms different from placebo response? If so, is a 'local', or 'non-local' model more likely to fit the data? This question is not just of academic interest. If a non-local model is pertinent, experimental procedures trying to isolate the causative principle in homeopathic medicines are likely to be doomed to failure.[18,19] For it is a consequence of non-local types of explanation of homeopathy that the effects are, by definition, not confined to a certain element of the system, eg to the medicine, but reside in the whole therapeutic procedure.

To investigate these questions, we devised a new pilot model for a HPT which could also be an experimental human volunteer model to test for the specificity and locality assumptions at the same time. We report the results of this pilot study. The main aim of the study was to improve detection symptoms and gain a finer resolution of observed symptoms by implementing the proving method developed by Hahnemann and expounded by Sherr.[17] We combined this phenomenologically accurate method with rigorous methodological standards.

Method

Subjects and procedures

We chose healthy volunteers who were students of two homeopathic schools (see below). Exclusion criteria were used for any homeopathic medicine or participation in an HPT during the 10 weeks preceding the trial, or the presence of organic pathology particularly if under allopathic medication. Compliance with these criteria was verified by the proving supervisor (JS) during a preliminary diagnostic case taking interview.

Medicines were prepared from Helios pharmacy from a tincture made according to the French Homeopathic Pharmacopoeia. Placebos were identical sugar globules, impregnated with the same amount and concentration of unsuccussed alcohol as was used for preparing the verum. Volunteers were recruited from

homeopathic schools studying homeopathic proving as part of their curriculum. Provers gave their informed consent and noted their baseline symptoms in a proving diary. The ethics committee of Freiburg University Hospital was notified and saw no necessity of a formal decision, since only ultra high dilutions were used.

The volunteers recorded their symptoms in the proving diary as they were noticed, at least once a day, and were contacted daily by the proving supervisor. After 1 week of baseline documentation, the provers were randomised to receive either a preselected homeopathic proving medication or an indistinguishable placebo. Provers were advised to take a maximum of six doses over 2 days and to stop intake of medication as soon as symptoms appeared. The proving supervisor ensured that all symptoms observed were noted in the diary, and that any symptoms due to an individual's constitution or recent and existing pathology were eliminated. To determine which symptoms were omitted and which were categorised as new, unusual or different from the normal status, as specified by Sherr's criteria, supervisors asked provers and used the knowledge gained from the previously conducted extensive case-taking interview with the volunteers. Provers were asked whether the symptom observed was new, unusual, or in any sense different from the normal status.[17]

After the proving, the symptoms were saved in an electronic database. Symptoms were coded individually with onset, duration, location, sensation, extension and modality. The symptom database was arranged according to the head-to-foot-scheme, randomly shuffled so that entries were not discernible by volunteer or day, and then given to an external materia medica expert who was not otherwise involved in the proving. The materia medica expert (GR) is an accredited homeopathic physician deemed by the proving supervisor (JS) to be competent for the task since she has been practising in a full time homoeopathy for 13 years.

Design

The design was a double-blind, randomised, placebo-controlled experimental study with two phases with recording baseline symptoms for 1 week before the proving and new symptoms for 2 weeks during the proving. Twelve potential medicines were selected by JS as comparatively well known with a long history of use and mostly acute actions. After a 7 day period of baseline recording by the provers, the pharmacist (JM) chose one of the 12 medicines to be used in the trial by applying a list of random numbers. This medicine was used for the proving. Samples of freshly prepared medicated globules with the medicine in 30c, or samples of unmedicated placebo globules, were dispatched to the provers according to another random-number list. The list of random numbers was generated by using the random number generator of the statistical package SPSS.

The proving director and the principal investigator (HW) were blind to the medicine chosen. The provers and the supervisor were all blind to the actual medicine chosen. In addition, they were also blind to the list of potential medicines and to the design of the study. The only person who knew the allocation code and the medicine to be tested was the pharmacist, who was not otherwise involved in the proving and had no contact with any of the provers. All steps of the study were defined in a pre-specified protocol, which was signed by all parties and deposited before the start of the study.

Data evaluation

The materia medica expert received the database of verified symptoms from the study secretary. All symptoms had been processed and partially blinded, so that she did not know which symptoms were experienced by which prover. In a first qualitative step of evaluation, she tried to determine the medicine tested. Next, she was given the selection of 12 medicines and asked to choose the correct one from this set. Finally after these two steps a quantitative evaluation was performed. The expert was given the name of the medicine used. She had then to decide for each symptom whether it was typical for that medicine or not. She could use any accessible repertory or knowledge base. This enabled us to calculate two single, continuous numerical variables: number of medicine-typical symptoms per subject and phase, and number of symptoms not typical for the medicine. We evaluated these data with dependent and independent *t*-tests.

Knowledge and levels of blindness

The proving coordinators were chosen by the proving director (JS) on the basis of their knowledge of the proving process and their trustworthiness and the institutions they represent. The organisation of the proving itself, including educating provers and supervisors and collating and editing the information was undertaken by the proving coordinators. The coordinators selected supervisors to oversee provers. Every supervisor supervised one or rarely two volunteers. Supervisors had daily contact with the provers in person or by phone. The proving director, apart from instructing the coordinators at the beginning, had no contact with the coordinators or provers during the HPT. The principal investigator and the statistician (RS) had no contact with the provers at any time, did not know the randomisation code of the provers and were not aware of which medicine was tested. The materia medica expert did not know the provers and had no contact with them during the HPT. She learned the identity of the medicine from the pharmacist only after she had made her guesses, and was not given the

randomisation code. The pharmacist did not know the provers. Although he received the randomisation code, he did not know which prover corresponded to which code number. Neither the provers, nor the supervisors were told the design of the study, the potential medicines in the list, or the randomisation code. Thus, the study was strictly double blind for everyone involved.

Results

Eleven volunteers entered the study: seven were students from a school of homeopathy in the UK, six students from a school of homeopathy in Israel. Seven subjects completed the baseline data collection in phase one and four provided no data, eg due to lack of symptoms. Out of the list of 12 pre-established potential medicines (Table 1), *Cantharis* was the medicine randomly chosen to be administered in 30c potency.

Qualitative Findings

The materia medica expert was not able to determine the correct medicine, either in step 1 (unrestricted choice), or step 2 (restricted choice).

Quantitative Findings

There was a strong increase from baseline to proving period both in symptoms typical and atypical of *Cantharis*. This increase was significant for both types of symptoms ($P=0.03$ for typical, $P=0.02$ for atypical). There was no significant difference between groups (*Cantharis* vs. placebo) either for typical or atypical symptoms during the proving period. As our study was a pilot trial and not adequately powered to discover small effects, we calculated effect sizes (Cohen's d) for the interesting differences. We used standardised mean differences with the larger standard deviation of the respective groups in the denominator. Effect sizes for the difference between *Cantharis* and placebo during the proving period were $d=0.4$ for typical symptoms, and $d=0.6$ for atypical symptoms. Effect sizes for the difference between placebo and *Cantharis* during the baseline period were even larger: The difference between groups for typical symptoms during baseline was $d=1.0$, and for atypical symptoms $d=0.7$.

Results are given in Table 2 and Fig. 1.

Table 1 Pre-established list of 12 medicines

Aconitum	Agaricus
Apis	Belladonna
Bryonia	Coccus cacti
Cantharis[*]	Chamomilla
Hypericum	Nux vomica
Rhus toxicodendron	Veratrum

[*]Randomly selected.

Table 2 Number of symptoms typical for *Cantharis* and atypical ones during baseline and proving period for both groups (*Cantharis* and placebo); mean (standard deviation)

	Placebo	*Cantharis*
Typical symptoms baseline	7.0 (6.8)	12.0 (8.7)
Atypical symptoms baseline	6.0 (2.3)	10.0 (9.2)
Typical symptoms proving	28.6 (23.2)	35.7 (14.3)
Atypical symptoms proving	16.2 (5.6)	22.0 (12.2)

Significant difference baseline to proving for all symptoms (typical: P=0.03, atypical: P=0.02). Not significant: difference between placebo and Cantharis during proving.

Discussion

This study was the first using a thorough phenomenological approach, as used by Sherr, in combination with strict methodological standards and quantitative evaluation. The method enabled us to calculate a straightforward outcome variable: number of symptoms typical for a tested medicine in comparison with number of symptoms atypical for the medicine tested. This study was a first pilot study, so its results are not definitive. It used a small number of provers. The lack of subjects, however, was partially compensated by the thorough documentation of symptoms, as witnessed by the relatively high average number of symptoms. Some provers failed to complete the baseline due to absence of symptoms. Therefore, the baseline data are not as reliable as we would have wished. Thus, the strong and significant difference for both typical and atypical symptoms between baseline and proving period is partially artefactual.

Nevertheless, this large difference allows us to conclude that what was experienced during the proving period was different from mere background noise. If this were not the case, one would not observe such a sharp increase of symptoms. What can also be seen is the trend for *Cantharis* to show more typical symptoms than atypical ones. This also contradicts the notion that during the proving atypical and placebo symptoms are experienced. One would of course expect that if symptoms experienced in a proving are placebo symptoms, the majority of symptoms reported should not be typical for a certain medicine, but rather randomly scattered across all remedy pictures. The fact that we see, at least as a trend, more symptoms typical for *Cantharis* and less atypical ones speaks against the pure placebo-hypothesis of symptom pictures.

If the local hypothesis of homeopathy were correct, typical symptoms would mainly be expected in the *Cantharis* group, and atypical symptoms—or at least less typical symptoms—in the placebo group (as indicated in Fig. 1). The finding that more typical symptoms were also experienced in the placebo group should not occur under a local hypothesis. This is especially true since the materia medica expert, who defined which symptom was typical and which was not, was blind as to which symptoms occurred during

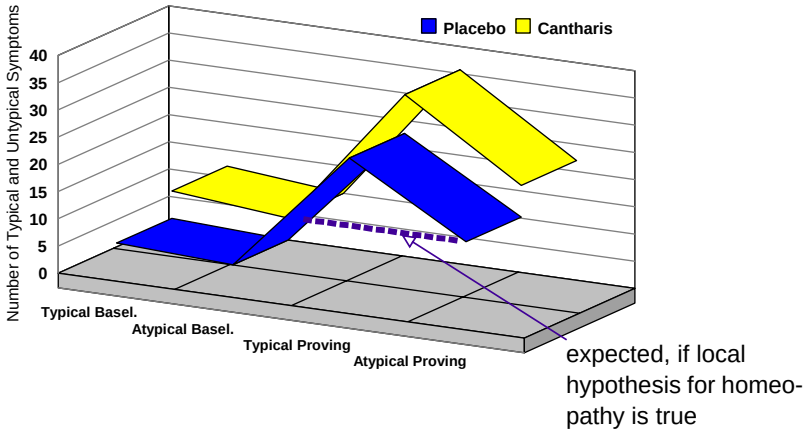

Figure 1 Mean number of symptoms typical for *Cantharis* and atypical ones during baseline and proving period for both groups.

baseline and proving, as well as to group assignment. Using the largest standard deviation, the effect size for the rise in typical symptoms from baseline to proving in the placebo group is $d=0.9$. This is a hint that whatever happens in the placebo group is unlikely to be just a rise in generic placebo activity, but instead is also due to a non-local or 'field effect'. Otherwise, we cannot explain the finding that typical *Cantharis* symptoms are experienced in the placebo group more frequently than atypical ones. Thus, there appears to be a 'mirroring' of effects in the placebo group, illustrated by a highly significant correlation with the *Cantharis* group for typical symptoms ($r=0.69$; $P<0.01$). Thus, more typical symptoms found for *Cantharis* intake were indeed accompanied by more symptoms typical for *Cantharis* in those provers who took placebo. Assuming a local hypothesis, neither this specificity of symptoms in the placebo group nor the high correlation with the medicine group would be expected.

This is a potential experimental model which allows testing for the specificity of homeopathic provings because symptoms typical for a medicine, known from earlier provings and clinical practice can be discerned from atypical symptoms. The effect size between placebo and experimental group for typical symptoms during the proving period is small ($d=0.4$), but not zero. It would take about 80 subjects per group to provide adequate statistical power to confirm this effect. This would be a relatively large study by the normal standards for HPTs. Another option is to use a better known medicine with more typical symptoms.

It is also interesting that so many atypical symptoms were observed, and that the rise in atypical symptoms from baseline to placebo was so sharp. This could be due to the fact that *Cantharis* is not well proven. If a medicine is not well known and earlier provings are not complete, the likelihood of a new proving discovering new symptoms is larger than for a better known one. In this case, new symptoms not hitherto recorded and thus not contained in repertories and materia medicas, would by definition be recorded as symptoms not typical for the medicine. Other possible factors involved in the relatively large number of atypical symptoms are individual prover sensitivity: in any given HPT, provers will produce some symptoms not reported in previous provings of this medicine. The 200 year interval between the original proving of *Cantharis* and the present experiment accounts for a considerable difference, especially because of the increased emphasis on reporting mental symptoms in modern provings.

We think that the results do not reflect a guessing effect of provers and supervisors. An indication in this direction was the experience of the materia medica expert who was unable to determine which medicine was used, even after she was given the 12 options of the potential medicines. She reported that the symptoms observed in this proving were so different from what she had known of *Cantharis* from a clinical perspective that it simply did not occur to her to use *Cantharis* as an internal standard of comparison. This indicates that 'smaller' medicines in the materia medica, or even older medicines which have been used quite frequently (like *Cantharis*), might have provings that are not complete, especially if materia medica knowledge stems largely from clinical experience. One should also keep in mind that concentrated 'essences' and 'images' of medicines can distort and obscure the original proving.

The presupposition we make is that the materia medica contains specific information. Had this not been the case we would not have detected any difference between symptoms typical and atypical for a medicine. It is clear, however, that the accuracy of the information contained in the materia medica is the limit of resolution of our methodology. It is wise, therefore, for future studies along those lines to choose medicines considered to be well-proven or whose limit of accuracy in proving is at least well known.

While our HPT design has some drawbacks, as it was the first of this type, it incorporates some methodological precautions which make it valuable and its results noteworthy. The study was strictly double-blind, and for provers and supervisors triple-

blind. None of the provers, the supervisors, the proving director, nor the principal investigator knew the identity of the medicine used until the very end. The proving director knew the potential range of 12 medicines but was not in contact with the provers nor with the supervisors during the proving, until the database was complete, and could not communicate to them any anticipation about which symptoms were to be expected. It is therefore noteworthy that we saw such a rise in symptoms typical for *Cantharis* and a difference between typical and atypical symptoms for the whole proving. This could not be due to bias. The materia medica expert was unable to determine the correct medicine from the randomly organised experimental data she was given, she is knowledgeable about materia medica and a well-trained homeopathic doctor. It is therefore unlikely that the failure of our study to discern clear-cut effects between placebo and experimental group was due to a lack of knowledge on the part of the materia medica expert.

One might question whether our procedure of isolating single symptoms as the entity of study is a valid. It could be argued that only the whole symptom picture of one single person or even the totality of all provers' symptoms could be the unit of study. While this is true, we do not see how such an approach could be experimentally implemented. We tried to include this factor by adding the medicine guessing stages. However, these guesses may have been confounded by the inclusion of baseline and placebo symptoms. In future trials, separate materia medica experts may be required for each phase, so that the experts are given a pure sample of experimental symptoms on which to base their guesses.

We believe that we have chosen a reliable option of converting the phenomenological richness of an HPT into an experimental test, amenable to quantitative analysis. We have at least shown the possibility of a way of proceeding in further studies. Although our pilot data are preliminary and need to be replicated both with the same and with other medicines, this study indicates that homeopathic medicine provings seem to produce data typical for a given medicine and not just placebo background noise. Thus, symptoms of a HPT appear to be specific. However, this specificity is not linked to the experimental group but is 'smeared' across both groups: verum and placebo, alike. There are indications that a local hypothesis of homeopathy restricting a specific effect of homeopathic medicines merely to the treated group does not capture the whole reality. We have demonstrated that this model is potentially able to study such questions in more detail. We hope that our data stimulate further research among those interested. We believe that our model, with modifications, can be used as a generic test both of the specificity hypothesis of homeopathy and of the locality hypothesis. While the first hypothesis—the one of specificity—has received some support by our data, the locality hypothesis has been, if not refuted, at least weakened.

Acknowledgements

This study was supported by the Institut für Grenzgebiete der Psychologie und Psychohygiene (Institute for Frontier Areas of Psychology and Psychohygiene), Freiburg, Germany. The medicines were produced by Helios pharmacy. We thank John Morgan for his help and support in conducting the study. The Samueli Institute, USA, supported further research and is supporting HW and RS at present.

References

1 Walach H. Provings: the method and its future. *Br Hom J* 1994;**83**:129–131.
2 Walach H. Proving methodology. Letter. *Br Hom J* 1996;**85**:123–125.
3 Walach H. The pillar of homoeopathy: remedy provings in a scientific framework. *Br Hom J* 1997;**86**:219–224.
4 Dantas F *et al*. Homoeopathic remedy provings. *An International Review*, in preparation.
5 Weingärtner O. Über die wissenschaftliche Bearbeitbarkeit der Identifikation eines "arzneilichen Gehalts" von Hochpotenzen. *Forsch Komplementärmed Klass Nat* 2002;**9**:229–233.
6 Berezin AA. Isotopical positional correlations as a possible model for benveniste experiments. *Med Hypotheses* 1990;**31**:43–45.
7 Anagnostatos GS. Small water clusters (clathrates) in the homoeopathic preparation process, In: ed. Endler P C, Schulte J. (eds). *Ultra High Dilution—Physiology and Physics*. Dordrecht: Kluwer, 1994; pp 121–128.
8 Milgrom LR. Patient-practitioner-remedy (PPR) entanglement: a qualitative, non-local metaphor for homeopathy based on quantum theory. *Homeopathy* 2002;**91**:239–248.
9 Milgrom LR. Patient-practitioner-remedy (PPR) entanglement. Part 3. Refining the quantum metaphor for homeopathy. *Homeopathy* 2003;**92**:152–156.
10 Milgrom LR. Patient-practitioner-remedy (PPR) entanglement. Part 4. Towards classification and unification of the different quantum models for homeopathy. Homeopathy 2004;**93**:34-42.
11 Milgrom LR. Patient-practitioner-remedy (PPR) entanglement. Part 7. A gyroscopic metaphor for the vital force and its use to illustrate some of the empirical laws of homeopathy. *Forsch Komplementämed Klass Nat* 2004, in press.
12 Walach H. Entanglement model of homeopathy as an example of generalizsed entanglement predicted by weak quantum theory. *Forsch Komplementärmed Klass Nat* 2003;**10**:192–200.
13 Walach H. Magic of signs: a non-local interpretation of homeopathy. *Br Hom J* 2000;**89**:127–140.
14 Walach H. Does a highly diluted homoeopathic drug act as a placebo in healthy volunteers? experimental study of Belladonna C30. *J Psychosom Res* 1993;**37**:851–860.
15 Walach H, Hieber S, Ernst-Hieber E. The effects of homeopathic belladonna 30CH in healthy volunteers—a randomized, double-blind experiment. J Psychosom Res 2001;**50**:155–160.
16 Walach H, Hieber S, Ernst-Hieber E. Effects of Belladonna 12 CH and 30 CH in healthy volunteers. A multiple, single-case

experiment in randomization design, In: Bastide M. (ed.). *Sings and Images. Selected Papers from the 7th and 8th GIRI Meeting, Held in Montpellier, France, November 20–21, 1993, and Jersualem, Israel, December 10–11, 1994, GIRI-Yearbook*. Dordrecht, Boston, London: Kluwer, 1997;pp 215–226.

17 Sherr J. *The Dynamics and Methodology of Homoeopathic Provings*. West Malvern: Dynamis Books, 1994.

18 von Lucadou W. The model of pragmatic information (MPI). *Eur J Parapsychol* 1995;**11**:58–75.

19 von Lucadou W. Kommentar zu Taylor, *et al*. (2000) in Journal club, *Forsch Komplementärm Klass Nat* 2001;**8**:43–46.