PUBLICATION POLICY

Twice a year the Parapsychology Laboratory of the University of Utrecht publishes the European Journal of Parapsychology. The object of the European Journal of Parapsychology is to stimulate and enhance the activity in this field, especially in our corner of the world, by communicating research results and issues related to professional parapsychology. Although there will be an emphasis on experimental work, theoretical articles are also welcome. Contributions from all over the world will appear in the journal.

A hallmark of the European Journal of Parapsychology is the attempt to avoid selective reporting, that is, the tendency to bury 'negative' results and only to publish studies that 'turn out'. To avoid turning the journal into a graveyard for all 'unsuccessful' studies, we require that the acceptance or rejection of a manuscript should take place prior to the phase when the experimental data are collected. The quality of the design and methodology and the rationale of the study are considered more important than the level of significance of the outcome of the study. As a practical rule, we advise the potential contributor of an article to submit the design of his planned study before the work is actually carried out. The rationale of the study should be stated, as well as all the hypotheses related to it. Furthermore one should try to specify the number of subjects, the number of trials, etc., plus the type of statistical methods one plans to use for evaluation.

Priority will be given to the publication of studies which fulfil the above-stated publication policy.

The final manuscript with presentation of results must reach us two months in advance of the official publication dates, which are May 1st and November 1st.

THE EXTENT OF SELECTIVE REPORTING OF ESP GANZFELD STUDIES

Susan Blackmore
Parapsychology Laboratory
University of Utrecht

The repeatable psi experiment has proved an elusive goal for parapsychology. However, certain recent methods for eliciting psi, notably the use of ganzfeld stimulation, seem to have provided some degree of replicability. In 1977 Honorton reviewed all ganzfeld studies then published and showed that half, eight out of sixteen, obtained significant results. In a second review (Honorton, 1978) he gave 14 out of 26, or 54%, as significant. Sargent (1979) suggested that the ganzfeld technique has given 58% replicability both across laboratories and studies. My own survey of the literature published to date shows that of 31 studies 18, or 58%, obtained significant results. These figures are shown in table 1.

These estimates appear encouraging. However, various factors have to be taken into consideration before accepting them as valid. These include the methodological weaknesses in many of the studies, the use of inappropriate statistics or overanalysis, and selective reporting. Sargent (1980a) separated studies into those he considered adequate and inadequate in methodology. Of 11 deemed adequate 6, or 54%, obtained significance. In my own survey I considered 12 published studies to be adequate. Others were ruled out for such reasons as the

---

Note: This paper fulfils the publication policy of this journal.

TABLE 1
Estimates of the repeatability of ESP in the ganzfeld (published studies)

	Total	Number of significant studies	Percentage of significant studies
Honorton 1977	16	8	50%
Honorton 1978	26	14	54%
Sargent 1979	24	14	58%
Blackmore 1980	31	18	58%

use of picture targets without a duplicate set for judging, or the possibility of communication between subject and agent prior to judging. Some studies can be considered adequate for results obtained from independent judging but not from subjects' judging. Of the 12 remaining 6, or 50%, were significant. Clearly the adequacy or otherwise of the methods used makes little difference to the apparent replicability of the technique.

Selective reporting may also be a relevant factor. If many experiments are carried out but those which seem to be giving chance results are abandoned before completion or are never written up for publication then there will be a bias towards significance in the literature. This bias may be introduced by subjects dropping out, by authors themselves, or by editors selecting submitted work for publication. Any estimates of the repeatability of a finding based only on published results may overestimate the true repeatability to an extent dependent on the size of that selection. Sargent has pointed out that a ganzfeld study is relatively time-consuming and arduous to perform as compared with many types of forced-choice psi experiments. They are therefore unlikely to be undertaken lightly, for example as student projects or informal experiments. In addition, since the first such study was reported only six years ago (Honorton and Harper, 1974) there has not been a long time in which large numbers of such studies

could have been carried out. One might therefore expect the number of unreported ganzfeld studies to be small, and certainly smaller than for many other types of psi experiment. Finally the method used varies comparatively little.

For all these reasons it seemed that it might be possible to find out about the majority of unpublished ganzfeld studies and to compare the success rate of those with that of the published ones. In this way an estimate of the extent and importance of selective reporting could be made. The published data was surveyed, as already noted, and a questionnaire was used to find out about unpublished work. The results for each were then compared.

METHOD

A questionnaire was sent to any person who it was thought might have been associated with ganzfeld studies. This asked about studies not completed, and those completed but not published. The complete questionnaire appears in the appendix.

The questionnaire was initially sent to 81 people, most of whose names and addresses were obtained from the Parapsychological Association member's list. For those who had published ganzfeld work a list of references was appended to check that none had been omitted. In response to information given by some respondents a further five questionnaires were sent out, making 86 in total.

RESULTS

40 completed questionnaires were returned, giving a return rate of 47%. These elicited a total of 32 further ganzfeld studies, reported by 15 respondents. Of these 12 studies were reported as not completed. These included seven "in progress" with the intention to complete and five others. The other reasons for not completing were "methodological defects" (2), "results not significant" (2), "demonstration to the media, aborted" (1), "high attrition rate and overly complex design-analysis requirement" (1), and "departed for graduate school" (1) (the total is more than five since respondents could give more than one reason for not completing the study). In no case was "results

not significant" given as the sole reason for failing to complete and therefore no selection at this stage was apparent (assuming thruthful answers).

20 completed but unpublished studies were listed. Reasons for not publishing were as follows:

Intending to publish	13
Methodological defects	5
Results not significant	1
Other and:	
Never intended for publication	2
Awaiting modified confirmation	1

The 20 unpublished studies included both significant and non-significant results (significance on primary analysis for ESP in the ganzfeld only was asked for). The results of one could not be evaluated for various reasons given by the author. Of the remaining 19, 7 were reported to have given results significant at the .05 level on the primary analysis. This is 37% which compares with 58% (by my estimate) in the published data. As might be expected, the difference is in the direction of more significance in the published work, but a chi-square test comparing significant and non-significant for published and unpublished shows no significant difference (see table 2).

The adequacy of the method used is also relevant. Authors may be more willing to report significant findings if the method is sound, and the method may affect the findings. For any analysis here it is necessary to assume that authors correctly reported whether their methods were adequate or not and this assumption may be unwarranted and is hard to test. However, assuming so, we find that 14 of the 19 studies were adequate and of these 5, or 36%, obtained significant results. This provides no evidence that methodological weaknesses were responsible for significant findings in the unpublished any more than in the published work.

DISCUSSION

The proportion of 'significant' overall psi scores for unpublished ganzfeld studies is comparable to that found in the published ones. We

TABLE 2
Proportions of 'significant' studies, published and unpublished

	Total	Number of significant studies	Percentage of significant studies
Published	31	18	58%
Unpublished	19	7	37%
Total	50	25	50%

cannot be sure that all unpublished have been detected here, but certainly all those from the major parapsychological laboratories have been (assuming honest replies), and there would seem to be no a priori reason to suppose that the proportion would be markedly different in those which were undetected by this survey. In fact, in order to make a difference to the conclusions reached here we would need to find a further four non-significant unpublished but completed studies, and no significant ones, to add to those included here. It seems unlikely, though certainly possible, that with further investigation the position would change in this way.

Also relevant is to consider the overall significance. Two years ago Honorton (1978, p.87) pointed out that "even if we make the absurd assumption that there are ten . . (unreported failures)..for each significant study, the observed results would still be significant at p=0.02". Of all the studies included here, both published and unpublished, 25 out of 50 provided results significant at the .05 level. If these 25 were to be the selected best of a larger number of studies there would need to be a further 450 undetected non-significant studies.

We may therefore conclude with some certainty that the bias introduced by selective reporting of ESP ganzfeld studies is not a major contributor to the overall proportion of significant results,

and the apparent success of the technique. If the positive results are not due to psi but are spurious this must be for some reason such as methodological errors or overanalysis undetected here, unconscious biases, untruthful answers, or deliberate fraud, or other such factors. It is not due solely to selective reporting.

ACKNOWLEDGMENT

I would like to thank all those respondents who completed and returned the questionnaire so promptly.

ABSTRACT

Estimates of the replicability of ESP in the ganzfeld based on the published data alone range from 50% to 58%. In a survey of unpublished ganzfeld studies a further 32 studies were elicited. Of 20 completed but not published 37% obtained significant results in the primary analysis, a proportion not significantly different from that in the published data. It was concluded that selective reporting is not a major contributor to the apparent success and repeatability of ESP in the ganzfeld.

REFERENCES

Honorton, C. 'Psi and internal attention states'. In Wolman, B. (ed.) 'Handbook of parapsychology', Van Nostrand Reinhold, 1977.

Honorton, C. 'Psi and internal attention states: Information retrieval in the ganzfeld'. In 'Psi and states of awareness', Parapsychology Foundation, 1978.

Honorton, C. & Harper, S. 'Psi mediated imagery and ideation in an experimental procedure for regulating perceptual input'. J. A.S.P.R., 68, 156-168, 1974.

Sargent, C.L. 'Repeatable significance and the significance of repeatability'. Paper presented at the 3rd International Conference of the Society for Psychical Research, Edinburgh, 1979.

Sargent, C.L. Personal communication, March 1980 (a).

Sargent, C.L. Personal communication, October 1980 (b).

Present address:
Pear Tree Cottage
Greyfield Road
High Littleton
Avon, BS18 5YB
England