# USE OF BOTH SUM OF RANKS AND DIRECT HITS IN FREE-RESPONSE PSI EXPERIMENTS 

By George P. Hansen and Jessica Utts


#### Abstract

Free-response ESP experiments have used a number of different sta tistical indices for evaluation. Although reseachers have been criticized for using multiple indices without appropriate adjustments, there are good reasons for using several different indices. Some of these reasons are described herein.

This report discusses the use of both the sum of ranks and the direct hit measures. A procedure correcting for the dual analysis is described. This method, however, solves only one of the potential problems that may arise with multiple analysis. A BASIC computer program implementing this procedure is presented, and a table given for experiments that have four items in a target pool and sample sizes of $10,20,30$, and 40 trials.


A number of methods have been developed to statistically evaluate free-response ESP experiments. Burdick and Kelly (1977) have described two general approaches, atomistic and wholistic. Atomistic methods evaluate discrete portions of a response generated by a subject attempting to use ESP. Some recent procedures have been presented by Honorton (1975), Jahn, Dunne, and Jahn (1980), and May, Humphrey, and Mathews (1985). A wholistic approach involves comparing a subject's total response with members of the judging pool (the actual target and some number of decoys). The items in the pool are given rankings or ratings to indicate the level of correspondence with the subject's response. This might be done by the subject or by an outside judge who is blind to the correct match.

Wholistic approaches include the sum-of-ranks statistic (Solfvin, Kelly, \& Burdick, 1978), direct hits, binary hits, and standardized ratings (Stanford \& Mayer, 1974). In the psi ganzfeld work, the di-rect-hits measure has been the statistic most frequently used, but the sum-of-ranks method has also seen wide acceptance (Honorton, 1985).

## Which is the Best Statistic?

The best statistic is a matter of some question, and a number of investigators have commented on the effectiveness of various mea-
sures. Stanford and Sargent (1983) have described the standardized ratings as being "particularly psi-sensitive" (p. 319). Sargent (1980) has called the binary hit measure "very insensitive" (p. 7).

It should be kept in mind that the actual power of a test will depend on how psi occurs and on how the judging is done. There are a number of ways ESP could manifest. For instance, perhaps it is an "all-or-none" phenomenon. In that case, a subject would either get a direct hit or not; and, if not, his attempt would not be any more likely to get a second place rank than it would a fourth place. Perhaps psi operates so that a first place ranking is more likely than a second place, and a second place is more likely than a third place, and so on. Perhaps the effect operates so that the target is simply more likely to be placed in the top half of the rankings than in the bottom half.

If psi operates so that there are mostly low-valued ranks in an experiment, then the sum-of-ranks statistic would be the best to use. But if psi operates in an all-or-none fashion (an extrachance number of first place ranks but all other rankings evenly distributed), then the direct-hits measure is more powerful. Before any valid statement can be made about power comparisons of various statistical analyses, the assumed mode in which psi occurs must be specified. McConnell (1958) pointed this out in a similar situation. (Although it is beyond the scope of this paper, such power comparisons can be studied with the use of simulation, which is now feasible because of the widespread availability of microcomputers [e.g., Hansen, 1986].)

It is unclear just how ESP should be expected to manifest in free-response situations. In fact, it may depend on a variety of factors. For instance, Honorton and Schechter (1986) presented data from ganzfeld research indicating that personality type is one factor that affects how psi scoring is distributed. Another factor that may influence the outcome is the target pool itself. If several items within a pool are somewhat similar, a judge may experience some confusion; as a result, a target may receive a second place rank rather than a first place rank. For instance, if several pictures of people are included in the pool and the subject's response mentions people, the judge may have some difficulty. In such situations, the sum-of-ranks index may be more powerful than the direct-hits index. If a judge is inexperienced, he or she may pay attention only to the best match and neglect the others. In such cases, a second place target rank may be no more likely than a last place ranking; in this instance, the direct-hits measure would be more sensitive than the sum of ranks.

## Reasons for Several Measures

There are a number of reasons an investigator might wish to use several indices to check for statistical significance. As just mentioned, psi may manifest in several ways; a researcher may want to allow for several possibilities. Another reason one may desire to use multiple indices is that discrete distributions can impose a more severe requirement than intended. As an example, in Table A. 1 published by Solfvin, Kelly, and Burdick (1978), with $N=10$ and $R=$ 4 , a sum of ranks of 18 or less corresponds to a probability of .033 ; the next highest sum (19) corresponds to a probability of .06 . In such circumstances, a second appropriate criterion can be allowed for significance testing while keeping the significance level at or below. 05.

When similar cases arose previously, the Bonferroni method (e.g., Rosenthal \& Rubin, 1984) could be used to correct for multiple analysis (e.g., Honorton, 1985). This procedure is conservative because its $p$ values are derived from an extreme negative association between the measures. For many cases, this is unrealistic. In a typical ganzfeld experiment, for instance, a large number of direct hits would be associated with a small sum of ranks; Hyman (1985) found these to be highly correlated.

## Use of Both Direct Hits and Sum of Ranks

In the following text, we present a method enabling the use of both direct hits and sum of ranks together to test for statistical significance but avoiding the excess conservatism (and loss of statistical power) associated with the Bonferroni method. A computer program was written to calculate exact bivariate probabilities assuming that both the direct-hits and sum-of-ranks indices were checked. The appendixes contain a description of the calculation procedure and rationale, an annotated computer program in Applesoft ${ }^{\text {©00 }}$ BASIC, and a table that gives examples.

Table 1 in Appendix A gives results of calculations for experiments with $10,20,30$, and 40 trials, each with four items in a target pool (and therefore a probability of a direct hit of .25). To save space, we show only limited regions of the distributions. It should be noted that the computer program can handle many more cases than the four examples.

There are two ways to use the table. The " $p$-value" approach is to observe the data first, then read the combined level of significance from the table. For example, in a 10 -trial experiment with four items in a judging pool, observing six direct hits and a sum of ranks of 16 would give a $p$ value of .02137 . This means that the probability of observing six or more direct hits, or a sum of ranks of 16 or less, or both, is .02137 under the null hypothesis.

The second approach to using the table is to decide in advance which pairs will constitute a significant result. This approach has the advantage of allowing the experimenter to place more emphasis on one of the measures over the other.

Suppose that you wish to conduct a 20 -trial experiment with direct hits as the primary analysis. The cumulative binomial distribution (from a table or calculation) indicates that nine direct hits are needed (for a probability of .04093 ). Fight direct hits gives an associated probability of .10181. Consulting part B of Table 1 in Appendix A indicates that a second criterion can be allowed (i.e., a sum of ranks of 39 or less). This gives an overall probability value of . 04611 . Thus, if the experiment resulted in only six direct hits but a sum of ranks of 39 , it would still be significant because one of the prespecified criteria was met.

As another example, suppose that you wish to conduct a 30 -trial experiment and to weigh each criterion approximately equally. In that case, you might specify the criteria of either 13 or more direct hits or a sum of ranks of 63 or less for an overall probability of .03946. Individually, the probability of 13 (or more) direct hits is .02159 , and that of a sum of ranks of 63 (or less) is .02992 . If the Bonferroni method was used, the conservative $p$ value would be $.02159+.02992$ instead of the exact $p=.03946$.

One word of caution must be inserted here. One must decide in advance whether only direct hits or sums of ranks will be used, or both. It is not legitimate to use the best of the $p$ values resulting from each of the two separate measures and the combined measure. These methods also require trials that can properly be considered independent (see Kennedy, 1979, for discussion).

## Appendix A

Calculation Procedure and Rationale

The calculation for exact bivariate probabilities is made by considering a multinomial distribution where all possible outcomes of a trial are equiprob-
able (i.e., a target is equally likely to be assigned any of $R$ ranks). Nearly all elementary books on probability discuss this (e.g., Harris, 1966). For $N$ trials, there are $R^{v}$ possible assignments of the ranks. The maximum sum obtainable would be $N \cdot R$ (all complete misses); the minimum sum would be $N$ (all direct hits).
Let $n_{k}$ designate the number of targets given a $k$ th rank, $n_{\star}$ being a nonnegative integer.

$$
\begin{equation*}
N=n_{1}+n_{2}+\ldots n_{R} \tag{1}
\end{equation*}
$$

The two statistics of interest are the sum of ranks ( $M$ ), given by

$$
M=1 \cdot n_{1}+2 \cdot n_{2}+\ldots+R \cdot n_{R}
$$

and

$$
D=\text { number of direct hits }=n_{1} .
$$

Notice that these are both functions of ( $n_{1}, n_{2}, \ldots, n_{R}$ ). Therefore, the combined probability associated with particular values of the pair $(D, M)$ is the sum of the probabilities of all sets of ( $n_{1}, n_{2}, \ldots, n_{R}$ ) for which those particular values are obtained.
It follows that to compute those probabilities, we need to enumerate the probabilities for all sets ( $n_{1}, n_{2}, \ldots, n_{R}$ ) that satisfy Equation 1. Then, as we enumerate each set, we compute ( $D, M$ ) and accumulate the probability for that pair.

As a small example, suppose that $N=2$ and $R=2$. Then $n_{1}+n_{2}=2$. The possible sets of ( $n_{1}, n_{2}$ ), their probabilities (under the null hypothesis), and the associated ( $D, M$ ) are:

| $\left(n_{1}, n_{2}\right)$ | Probability | $(D, M)$ |
| :---: | :---: | :---: |
| $(0,2)$ | $1 / 4$ | $(0,4)$ |
| $(1,1)$ | $1 / 2$ | $(1,3)$ |
| $(2,0)$ | $1 / 4$ | $(2,2)$ |

In this case there is only one set of $\left(n_{1}, n_{2}\right)$ corresponding to each $(D, M)$ pair. In general, for each ( $D, M$ ) pair we would add the probabilities of all sets of $n_{K}$ 's resulting in that pair.

The remaining question is how to enumerate and compute probabilities for the sets ( $n_{1}, n_{2}, \ldots, n_{R}$ ). Under the null hypothesis, each trial is equally likely to receive any of the $R$ ranks, so the set ( $n_{1}, \ldots, n_{R}$ ) has a multinomial distribution with equal probabilities, each $1 / R$. Thus,

$$
\begin{equation*}
P\left(n_{1}, n_{2}, \ldots, n_{R}\right)=\frac{N!}{n_{1}!n_{2}!\ldots n_{R}!}\left(\frac{1}{R}\right)^{v} \tag{2}
\end{equation*}
$$

The challenge in writing the program was to find a way to systematically
Exact Probability of Obtaining Either the Indicated Number of Direct Hits (or More) Or The Indicated Sum of Ranks (Or Less) Or Both

| A. For 10-trial experiment with $p$ (hit) $=.25$ |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Direct hits | Sums of ranks |  |  |  |  |  |  |  |
|  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 0 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 |
| 1 | . 94369 | . 94369 | . 94369 | . 94369 | . 94369 | . 94369 | . 94369 | . 94369 |
| 2 | . 75597 | . 75597 | . 75597 | . 75597 | . 75597 | . 75597 | . 75597 | . 75597 |
| 3 | . 47441 | . 47441 | . 47441 | . 47441 | . 47441 | . 47441 | . 47441 | . 47441 |
| 4 | . 22412 | . 22412 | . 22412 | . 22412 | . 22412 | . 22412 | . 22412 | . 22424 |
| 5 | . 07813 | . 07813 | . 07813 | . 07813 | . 07813 | . 07813 | . 07833 | . 07964 |
| 6 | . 01973 | . 01973 | . 01973 | . 01973 | . 01973 | . 01997 | . 02137 | . 02629 |
| 7 | . 00351 | . 00351 | . 00351 | . 00351 | . 00371 | . 00475 | . 00815 | . 01628 |
| 8 | . 0004 | . 0004 | . 0004 | . 0005 | . 00107 | . 00280 | . 00701 | . 01582 |
| 9 | . 00003 | . 00003 | . 00007 | . 0002 | . 0009 | . 00276 | . 00701 | . 01582 |
| 10 | . 00000 | . 00001 | . 00006 | . 0002 | . 00009 | . 00276 | . 00701 | . 01582 |
|  | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| 0 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 |
| 1 | . 94369 | . 94369 | . 94369 | . 9437 | . 94375 | . 94395 | . 94454 | . 94592 |
| 2 | . 75597 | . 75598 | . 75607 | . 75651 | . 75805 | . 7622 | . 7712 | . 7874 |
| 3 | . 47445 | . 4748 | . 47643 | . 48168 | . 49464 | . 52041 | . 56306 | . 62286 |
| 4 | . 22508 | . 22864 | . 23908 | . 26275 | . 30615 | . 37278 | . 4604 | . 56107 |
| 5 | . 08469 | . 09826 | . 12673 | . 17564 | . 24727 | . 33914 | . 44478 | . 55546 |
| 6 | . 03855 | . 06294 | . 10366 | . 16338 | . 24222 | . 3377 | . 44454 | . 55546 |
| 7 | . 03234 | . 05993 | . 10266 | . 16318 | . 24222 | . 3377 | . 44454 | . 55546 |
| 8 | . 03223 | . 05993 | . 10266 | . 16318 | . 24222 | . 3377 | . 44454 | . 55546 |
| 9 | . 03223 | . 05993 | . 10266 | . 16318 | . 24222 | . 3377 | . 44454 | . 55546 |
| 10 | . 03223 | . 05993 | . 10266 | . 16318 | . 24222 | . 3377 | . 44454 | . 55546 |


| $m$ |  |  |
| :---: | :---: | :---: |
| n |  |  |
| $\bar{m}$ |  |  |
| m |  |  |
| N |  |  |
| N | -TNT~N~N~N~N O- <br>  |  |
| N |  <br> 为 |  |
| $\stackrel{1}{\sim}$ |  |  |
| 魚 |  |  |

B. For 20-trial experiment with $p($ hit $)=.25$

|  |
| :---: |
|  O <br>  |
|  |
|  |
|  |
|  <br>  <br>  |
|  <br>  <br>  |

69

|  <br>  |  |
| :---: | :---: |
|  |  <br>  <br>  |
|  |  <br>  <br>  |
|  |  <br>  <br>  |
|  |  <br>  <br>  |
|  |  ox <br>  |
|  |  <br>  <br>  |
|  |  <br>  <br>  |
|  |  |



enumerate all possible sets of $n_{k}$ 's. Notice that an arrangement can be specified by writing down $N+R-1$ "slots" and placing "pegs" in $R-1$ of them. The number of slots before the first peg is $n_{1}$, between the first and second peg is $n_{2}$, and so forth. For example, if $N=2$ and $R=2$, we would have three slots and one peg, resulting in:

$$
\frac{1}{\ldots} \text {, or } \underbrace{}_{1} \text {, or } \ldots \frac{1}{}
$$

corresponding to ( $n_{1}, n_{2}$ ) of $(0,2),(1,1)$, and ( 2,0 ), respectively.
Using this concept, the program systematically moves the pegs until all sets of $n_{k}$ have been enumerated. The number of sets is:

$$
C_{R-1}^{N+R-1}=\frac{(N+R-1)!}{N!(R-1)!}
$$

For each set, $D$ and $M$ are calculated, and the probability given by Equation 2 is added to the appropriate cell in Table 1.

## Appendix B

Annotated Computer and Program in Applesoft BASIC

10

INPUT "NUMBER OF TRIALS $>$ "; T
INPUT "NUMBER OF RANKS>"; R
REM Maximum $\mathrm{R}=8$
REM Maximum T depends on R and computer memory
DIM D (T,R * T)
$\mathrm{X}=\mathrm{T}:$ GOSUB 6000: TFAC $=\mathrm{FAC}$
FOR Pl $=1$ TO T +1
IF R $=2$ THEN GOSUB 2000: NEXT : GOTO 600
FOR P2 = P1 + 1 TO T +2
IF $\mathrm{R}=3$ THEN GOSUB 2000: NEXT : GOTO 560
FOR P3 = P2 + 1 TO T + 3
IF R $=4$ THEN GOSUB 2000: NEXT : GOTO 550
FOR P4 = P3 + 1 TO T +4
IF $\mathrm{R}=5$ THEN GOSUB 2000: NEXT: GOTO 540
FOR P5 = P4 + 1 TO T + 5
IF R = 6 THEN GOSUB 2000: NEXT : GOTO 530
FOR P6 = P5 + 1 TO T +6
IF $\mathrm{R}=7$ THEN GOSUB 2000: NEXT : GOTO 520

FOR P7 $=\mathrm{P} 6+1 \mathrm{TO} \mathrm{T}+7$
IF $\mathrm{R}=8$ THEN GOSUB 2000: NEXT
NEXT P6
NEXT P5
NEXT P4
NEXT P3
NEXT P2
NEXT P1
$\mathrm{TT}=\mathbf{R} \wedge \mathrm{T}$
REM This section cumulates counts in each cell to
REM determine the cumulative distribution
FOR I $=$ TTOT $*$ R
FOR J = 0 TO T
IF $\mathrm{J}=0$ THEN $\mathrm{V}=0$ : GOTO 850
$\mathrm{IF} \mathrm{I}=\mathrm{R} * \mathrm{~T}$ THEN $\mathrm{V}=0$ : GOTO 850
$\mathrm{V}=0$
FOR II = I + 1 TOR *T
FOR JJ $=0$ TO J -1
$\mathrm{V}=\mathrm{V}+\mathrm{D}(\mathrm{JJ}, \mathrm{II})$
NEXT JJ
NEXT II
$\mathrm{D}(\mathrm{J}, \mathrm{I})=(\mathrm{TT}-\mathrm{V}) / \mathrm{TT}$
NEXT J
NEXT I
REM This section prints results
FOR J = T TO R * T
FOR I $=0$ TO T
PRINT D (I,J);"
NEXT I
PRINT
NEXT J
END

REM
$\mathrm{P}(1)=\mathrm{P} 1: \mathrm{P}(2)=\mathrm{P} 2: \mathrm{P}(3)=\mathrm{P} 3: \mathrm{P}(4)=\mathrm{P} 4: \mathrm{P}(5)=\mathrm{P} 5$ :
$\mathrm{P}(6)=\mathrm{P} 6: \mathrm{P}(7)=\mathrm{P} 7$
$\mathrm{P}(\mathrm{R})=\mathrm{T}+\mathrm{R}$
REM This loop determines 'spaces' between 'pegs'
FOR I $=1$ TO R
$\mathrm{X}(\mathrm{I})=\mathrm{P}(\mathrm{I})-\mathrm{P}(\mathrm{I}-\mathrm{I})-\mathrm{I}$
$\mathrm{X}=\mathrm{X}(\mathrm{I})$
GOSUB 6000
F (I) = FAC
$\mathrm{FAC}=\mathrm{FAC}+\mathrm{LOG}(\mathrm{K})$
NEXT K

REM For $T=10$ and $R=4$, estimated running time is

## NEXT I

$\mathrm{S}=0: \mathrm{TS}=\mathrm{TFAC}$
REM This loop calculates number of arrangements
REM (N!/(nI! n2! ... nR!)
FOR I = 1 TO R
$\mathrm{S}=\mathrm{S}+\mathrm{I} * \mathrm{X}(\mathrm{I})$
$\mathrm{TS}=\mathrm{TS}-\mathrm{F}(\mathrm{I})$
NEXT I
$\mathrm{D}(\mathrm{X}(1), \mathrm{S})=\mathrm{D}(\mathrm{X}(1), \mathrm{S})+2.718281828 \wedge \mathrm{TS}$
RETURN

REM Factorial, Logarithms are used
IF $\mathrm{X}=0$ THEN FAC $=0$ : RETURN
IF $\mathrm{X}=1$ THEN FAC $=0:$ RETURN
$\mathrm{FAC}=0$
FOR K = l TO X
$\mathrm{FAC}=\mathrm{FAC}+\mathrm{LOG}(\mathrm{K})$
NEXT K

REM approx. 7.5 minutes on an Apple
REM For $T=40$ and $R=4$, estimated running time is
REM approx. 21 hours on an Apple

## References

Burdick, D. S., \& Kelly, E. F. (1977). Statistical methods in parapsychology. In B. B. Wolman (Ed.), Handbook of parapsychology. New York: Van Nostrand Reinhold.
Hansen, G. P. (1986). Monte Carlo methods in parapsychology. Paper presented at the Twenty-Ninth Annual Parapsychological Association Convention, August 5-9, Rohnert Park, CA
Harris, B. (1966). Theory of probability. Reading, MA: Addison-Wesley.
Honorton, C. (1975). Objective determination of information rate in psi tasks with pictorial stimuli. Journal of the American Society for Psychical Research, 69, 353-359.
Honorton, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. Journal of Parapsychology, 49, 51-91.
Honorton, C., \& Schechter, E. I. (1986). Ganzfeld target retrieval with an automated testing system: A model for initial ganzfeld success. Paper presented at the Twenty-Ninth Annual Parapsychological Association Convention, August 5-9, Rohnert Park, CA.
Hyman, R. (1985). The psi ganzfeld experiment: A critical appraisal. Journal of Parapsychology, 49, 3-49.

Jahn, R. G., Dunne, B. J., \& Jahn, E. G. (1980). Analytical judging procedure for remote perception experiments. Journal of Parapsychology, 44, 207231.

Kennedy, J. E. (1979). Methodological problems in free-response ESP experiments. Journal of the American Society for Psychical Research, 73, 1-15.
May, E. C., Humphrey, B. S., \& Mathews, C. (1985). A figure of merit analysis for free-response material. Paper presented at the Twenty-Eighth Annual Parapsychological Association Convention, August 12-16. Medford, MA.
McConnell, R. A. (1958). Continuous variable trials. Journal of Parapsychology, 22, 279-290
Rosenthal, R., \& Rubin, D. B. (1984). Multiple contrasts and ordered Bonferroni procedures. Journal of Educational Psychology, 76, 1028-1034.
Sargent, C. L. (1980). Exploring psi in the ganzfeld. Parapsychological Monographs, No. 17. New York: Parapsychology Foundation.
Solfvin, G. F., Kelly, E. F., \& Burdick, D. S. (1978). Some new methods of analysis for preferential-ranking data. Journal of the American Society for Psychical Research, 72, 93-109.
Stanford, R. G., \& Mayer, B. (1974). Relaxation as a psi-conducive state: A replication and exploration of parameters. Journal of the American Society for Psychical Research, 68, 182-191.
Stanford, R. G., \& Sargent, C. L. (1983). Z scores in free-response ESP methodology: Comments on their utility and correction of an error. Journal of the American Society for Psychical Research, 77, 319-326.

## Psychophysical Research Laboratories <br> 301 College Rd., East <br> Princeton, NJ 08540

Division of Statistics
University of California
Davis, CA 95616

