A Multivariate Randomization Test of Association Applied to Cognitive Test Results
Albert Ahumada and Bettina Beard, NASA Ames Research Center

Abstract
Randomization tests provide a conceptually simple, distribution-free way to implement significance testing. We have applied this method to the problem of evaluating the significance of the association among a number (k) of variables. The randomization method was the random re-ordering of k-1 of the variables. The criterion variable was the value of the largest eigenvalue of the correlation matrix.

Introduction
The experimental data for which the randomization test was devised were collected to measure possible changes in cognitive abilities and in ratings of cognitive test difficulty following a simulated space ascent in a vibration-augmented centrifuge (Adelstein et al., 2009). The simulated space ascent study was done to evaluate the effects of G-load and vibration on display legibility (Beard et al., 2009). The cognitive study was "piggy-backed" onto the legibility study. Each of 11 participants was given a 5 test battery before and after the "ride" and were also asked to rate the difficulty/pleasantness of each test on 5 dimensions. The problem was then to assess the statistical significance of the correlations among the ratings and the tests.

Multiple Correlation Correction
One approach would be to look at the largest absolute correlation and correct for the number of correlations being considered. A strict Bonferroni correction for n multiple significance tests at joint level α is α/n for each single test (Benjamini & Hochberg, 1995). For a k = 5 test battery, there are k(k-1)/2 = 10 correlations to test. The simulation of Appendix 2 showed that the correction provided an accurate estimate of the significance of the maximum correlation for correlation matrices similar in size to ours constructed from independent Gaussian variables. However, the maximum correlation is relatively powerless against the alternative that the correlation is caused by a single common factor (Malevergne & Sornettea, 2004).

Largest Eigenvalue
Anderson (1958) shows that when all the correlations have the same value, the optimal test for whether there is a correlation is based on the largest eigenvalue or principal component of the correlation matrix. However, the distribution of the largest principal component is usually derived for asymptotic conditions (Johnstone, 2001) and for the case in which it comes from a covariance matrix rather than a correlation matrix (Harris, 2001).

The Randomization Test
Given that our ratings are not normally distributed and the simplicity of using a randomization test, we developed the test shown in Appendix 1. It also serves as a randomization test for bi-variate correlation. Anderson (1958) also shows that a likelihood ratio statistic for the hypothesis of independence against the general hypothesis is the product of the eigenvalues (i.e., the determinant of the covariance matrix). The function of Appendix 1 is easily modified to use the product of 1 to k eigenvalues as the criterion statistic for the randomization test.

References

Acknowledgments
Supported by the NASA Space Human Factors Engineering Project. Cynthia Null assisted with the literature search.
Appendix 1

A Matlab implementation of the randomization test.

```matlab
function [pr, s, u] = randeig1cor(x,nr)
% largest eigen value statistic randomization test
% based on the covariance matrix
% input x is N subjects by k scores matrix
% input nr is the number of randomizations
% output pr is the proportion of randomizations whose first
% eigen value exceeds or equals that of the sample
% output s has the sample eigen values
% output u has the sample eigen vectors

[N k] = size(x);
y = x;
[u s] = svd(corrcoef(y));
s1 = s(1);
pr = 0;
for ir = 1:nr
  for ik = 2:k
    y(:,ik) = y(randperm(N),ik);
  end
  [ui si] = svd(corrcoef(y));
  pr = pr + (si(1) >= s1);
end
pr = pr/nr;
```

Appendix 2

A Matlab simulation of the Bonferroni method for adjusting the significance level to test the significance of the largest correlation in a correlation matrix.

```matlab
% Test Bonferroni method for maximum correlation
N = 32; % so t df (N-2) will be in usual t table
N2 = N-2;
k = 5; % # of correlations is 5*4/2 = 10
p = 0.01; % significance level
pb = p/(k*(k-1)/2); % Bonferroni level
tb = 3.646; % cumulative t(1-pb/2,30)
% http://www.socr.ucla.edu/Applets.dir/T-table.html
% t = c/sqrt((1-c^2)/N2) => c = t/sqrt(t^2+N2)
cb = tb/sqrt(tb^2+N2); % 0.5541 Bonferroni criterion
nrep = 100000;
samp = zeros(nrep,1);
for irep = 1:nrep
  c = corrcoef(randn(N,k))-diag(ones(k,1));
  samp(irep) = max(abs(c(:)));
end
% Estimate actual significance
psamp = sum(samp>cb) % 975 1030 1035 (three runs)
psamp/nrep % 0.00975 0.01030 0.01035
mean([0.00975 0.01030 0.01035]) % 0.0101
% Estimate test criterion for p = 0.01
samp = sort(samp);
tsamp = 0.5*(samp(nrep*(1-p))+samp(nrep*(1-p)+1))
% 0.5531 0.5553 0.5554
mean([0.5531 0.5553 0.5554]) % 0.5546
```