Program Abstracts/Algorithms

An SPSS program to compute subjective organization

JOSEPH K. KAZEN and HAJIME OTANI

Central Michigan University, Mt. Pleasant, Michigan

This paper describes an SPSS program that computes four common subjective organization measures: Pellegrino's (1971) ARC' for bidirectional pairs and unordered triplets, Tulving's (1962) measure for bidirectional pairs, and Bousfield and Bousfield's (1966) deviation measure. This program can compute these measures between any two tests for an unlimited number of to-be-remembered items.

Free recall methodology involves the presentation of a list of items, followed by instructions to recall as many items as possible in any order. In contrast to other methods, this method allows investigators the advantage of analyzing the way in which subjects organize the items in order to facilitate recall. Especially when a test is repeatedly administered, it becomes apparent that subjects are not recalling items randomly (see, e.g., Tulving, 1962). Even though a list may not include any obvious structure, subjects recall items in a relatively fixed order. This phenomenon is commonly known as subjective organization.

Unfortunately, the computation of subjective organization is excessively complex and laborious. In addition, there are many measures from which to choose (see Sternburg & Tulving, 1977, for a review). Ironically, the only computer programs thus far available compute measures not commonly used in the experimental literature (Bateson & De Vader, 1990) or are based on obsolete technology (Pellegrino, 1974).

In this paper, we provide a program that will compute some of the more commonly used subjective organization measures (as suggested by Murphy & Puff, 1982). These include: Pellegrino's ARC' for bidirectional pairs (ARC) and unordered triplets (TARC), Tulving's (1962) measure of subjective organization (SO) for bidirectional pairs, and Bousfield and Bousfield's (1966) deviation measure for bidirectional pairs (DS). (See Table 1 for the computational basis for these measures.) To facilitate data analysis, we decided to use the statistical package

SPSS-X (1988), which is a widely available package on mainframe systems. This gives the experimenter the option to implement the computed measures in all the analytical procedures that SPSS provides (e.g., ANOVA, MANOVA, multiple regression, etc.). Also, by using the WRITE OUTFILE command, the researcher can create a new data file that can be utilized by other packages with which the investigator is more familiar (see SPSS-X, 1988, for more information on this procedure).

Program Description

Changing a few parameters in the first nine lines of the program file enables the program to compute the four measures between any two tests for an unlimited number of tests and to-be-remembered items (see Table 2). For the program to work, recall performance must be coded properly in the data file. In addition, the commands in the program file that define the location and meaning of the variables in the data file must be modified.

Data file. To create a data file, performance on all the to-be-remembered items has to be recorded for each sub-

Table 1
Computational Basis of Subjective Organization Measures

Measures:

$$\text{ARC}' = \text{O(ITR)} - \text{E(ITR)}/\text{MAX} - \text{E(ITR)}$$

$$\text{DS} = \text{O(ITR)} - \text{E(ITR)}$$

$$\text{SO} = \text{O(ITR)}/C + 1$$

where

O(ITR) = number of observed intertrial repetitions

E(ITR) = expected number of intertrial repetitions

MAX = maximum number of intertrial repetitions

C = number of common items recalled on Tests t and $t + 1$

For the ARC':

$$\text{E(ITR)} = (N - X + 1)!(A)(M - X + 1 - R)/N!$$

$$\text{MAX} = (M - X + 1 - R)$$

where

M = number of items recalled on Test t

N = number of items recalled on Test $t + 1$

R = number of pairs or triplets from Test t that lost one or more items on Test $t + 1$

X = number of items in the subjective organization unit (e.g., 2 for bidirectional pairs and 3 for triplets)

A = parameter based on the direction of the units (e.g., 2 for bidirectional pairs and A! or 6 for unordered triplets)

For DS:

$$\text{E(ITR)} = 2C(C - 1)/MN$$

For SO:

O(ITR) = items recalled at the beginning or end on two successive trials are added to the usual O(ITR)

Note—ARC', adjusted ratio of clustering (Pellegrino, 1971); DS, difference score (Bousfield & Bousfield, 1966); SO, subjective organization (based on Tulving, 1962).

The development of this program was supported in part by National Science Foundation Grant USE9051323. This program was developed as a part of the first author's master's thesis project at Central Michigan University. We thank David Payne, Robert Proctor, and an anonymous reviewer for their helpful comments on the manuscript. Correspondence should be sent to H. Otani, Department of Psychology, Central Michigan University, Mt. Pleasant, MI 48859 (e-mail: hajime.otani@cmich.edu).

Table 2
First Nine Lines of Program File

```

1 FILE HANDLE TEST/ NAME='TEST DATA A'
2 DATA LIST FILE=TEST RECORDS=1
3 /1 ID 1-3 A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 4-13 B1 B2 B3
  B4 B5
4 B6 B7 B8 B9 B10 14-23 C1 C2 C3 C4 C5 C6 C7 C8 C9 C10
  24-33
5 COMPUTE ITEMS=10
6 RECODE A1 TO A10 (0=1000)
7 RECODE B1 TO B10 (0=1000)
8 VECTOR SX=A1 TO A10
9 VECTOR SY=B1 TO B10

```

ject. Table 3 displays an example. Assume that a subject was given a list of 10 items (A–J) and then completed three free recall tests. Also assume that this subject recalled 6 items in the order of A,B,C,D,E,F on the first test, B,A,C,D,F,H,I on the second test, and, finally, C,A,B,E,F,I,H on the third test. As can be seen in the example, recalled items should be coded with their output serial positions, and nonrecalled items should be coded as 0.

Once the data are coded, the data file can be created as in the example. The first three columns represent the subject's identification number. The next three sections (10 columns each) are the subject's performance on the three tests. Needless to say, if the list includes more than 10 items, each section will require more than 10 columns. If necessary, more than one line per subject can be used to code the data. In addition, other factors (e.g., experimental conditions, covariates, etc.) can be included in the data file, as long as they are defined properly in the DATA LIST command in the program file.

Program file. The program file must be modified to accommodate the specific requirements of the data file. This is done by adjusting Lines 1–4 in the program in Table 2. Lines 1 and 2 identify the data file to be read. The FILE HANDLE command assigns a nickname to the data file, and the NAME command identifies the data file to be read. In the example, the data file is called TEST DATA A, and its handle is TEST. On Line 2, DATA LIST FILE=TEST indicates where the variables are and assigns variable names to the data in the file. The RECORDS=1 command indicates that each case in the data file is represented by one line of data. If the data for each subject is more than one line, then the number must be changed. For instance, if three lines are needed to code each subject's data, the command must be changed to RECORDS=3.

The /1 in Line 3 indicates the beginning of data definition for the first line of data for each subject. The ID 1-3 indicates that the subject's identification number is found in columns 1–3 in the data file and is given the label ID. Recall scores for all three tests are found in Columns 4–33. The 10 scores representing performance on the three tests are given the alphanumeric names A1–A10, B1–B10, and C1–C10 for Tests 1–3, respectively. It is important to name scores within a test with the same characters and assign numbers ranging from 1 to the

number of to-be-remembered items. The characters representing different tests cannot be the same.

The next step is to modify the parameters in Lines 5–9. The COMPUTE ITEMS=10 on Line 5 indicates that the number of to-be-remembered items was 10. This number must be modified, depending on the number of items included in a list. The RECODE commands on Lines 6 and 7 simply convert the 0s found in the scores on the tests to be analyzed to 1,000. The value of 1,000 is arbitrary but is essential for the program to compute subjective organization. The variable names found on these lines have to be changed, depending on the tests to be analyzed. For example, the program in the table is set up to convert the zeros in Test 1 and Test 2 to 1,000. To analyze organization between Tests 2 and 3, Lines 6 and 7 have to be changed to RECODE B1 to B10 (0=1000) and RECODE C1 TO C10 (0=1000), respectively.

The two VECTOR commands on Lines 8 and 9 convert the variable names on the two tests to be analyzed into a form that the program can easily use. For instance, the program is currently set up to analyze subjective organization between Tests 1 and 2. The variables for Test 1 (A1–A10) and Test 2 (B1–B10) are converted by the VECTOR commands to the variables SX1–SX10 and SY1–SY10, respectively. It is important to enter the variables from the earlier test (t) in the VECTOR SX command on Line 8 and scores from the later test ($t + 1$) on the VECTOR SY command on Line 9. Thus, if one desires to compute subjective organization between Tests 2 and 3, the VECTOR commands have to be changed to VECTOR SX = B1 TO B10 on Line 8 and VECTOR SY = C1 TO C10 on Line 9.

Output. The sample output is shown in Table 4. In addition to the four measures (two ARCs, SO, and DS), the program will compute M and N, which are the number of items recalled on tests t and $t + 1$.

Limitations

Under some circumstances, SPSS will issue warnings that indicate that a division by zero has been attempted in computing ARC or TARC for a particular case. This usually occurs when recall is extremely low on a test or when a number of items recalled on test t are not recalled

Table 3
An Example of How to Create a Data File

1. Assume that Subject 1 recalled items in the following order:

```

Test 1: A B C D E F
Test 2: B A C D F H I
Test 3: C A B E F I H

```

2. Code the above protocol as follows:

	Items									
	A	B	C	D	E	F	G	H	I	J
Test 1	1	2	3	4	5	6	0	0	0	0
Test 2	2	1	3	4	0	5	0	6	7	0
Test 3	2	3	1	0	4	5	0	7	6	0

3. Type in the data file:

```
001123456000021340506702310450760
```

Table 4
Output Example for Subject 1

ID	OITR	EITR	MAXITR	M	N	R	ARC
1	2.00	.86	3.00	6.00	7.00	2.00	.53
ID	OTITR	ETITR	MAXTITR	M	N	TRIR	TARC
1	1.00	.29	2.00	6.00	7.00	2.00	.42
ID	OITRTL	C	SO				
1	2.00	5.00	.33				
ID	DSEITR	C	DS				
1	.95	5.00	1.05				

Note—ID, subject's identification number; ARC, TARC, SO, and DS are the 4 subjective organization measures. Other parameters are used to compute the measures.

on test $t + 1$. These warnings can be eliminated, in part, by changing the formulas that compute the maximum values on Lines 95 and 96 to COMPUTE MAXITR = $M - X + 1$ and COMPUTE MAXTITR = $M - 3 + 1$ (Pellegrino, 1971).

Availability

The program can be obtained by sending a blank 3.5-in. floppy to H. Otani, Central Michigan University, De-

partment of Psychology, Mount Pleasant, MI 48859. The program can also be obtained via e-mail (hajime.otani@cmich.edu).

REFERENCES

- BATESON, A. G., & DE VADER, C. L. (1990). A FORTRAN program to calculate proximity values and an averaged proximity matrix. *Educational & Psychological Measurement*, **50**, 131-134.
- BOUSFIELD, A. K., & BOUSFIELD, W. A. (1966). Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports*, **19**, 935-942.
- MURPHY, M. D., & PUFF, C. R. (1982). Free recall: Basic methodology and analysis. In C. R. Puff (Ed.), *Handbook of research methods in human memory and cognition* (pp. 99-128). New York: Academic Press.
- PELLEGRINO, J. W. (1971). A general measure of organization in free recall for variable unit size and internal sequential consistency. *Behavior Research Methods & Instrumentation*, **3**, 241-246.
- PELLEGRINO, J. W. (1974). A FORTRAN IV program for analyzing higher-order subjective organization units in free recall learning. *Behavior Research Methods & Instrumentation*, **4**, 215-217.
- SPSS-X USER'S GUIDE (3rd ed.) (1988). Chicago, IL: SPSS Inc.
- STERNBURG, R. J., & TULVING, E. (1977). The measurement of subjective organization in free recall. *Psychological Bulletin*, **84**, 539-556.
- TULVING, E. (1962). Subjective organization in free recall of "unrelated words." *Psychological Review*, **69**, 344-354.

(Manuscript received December 1, 1994;
revision accepted for publication June 6, 1995.)