

# TESTAN: An integrated modular system for personality assessment and test development on MS-DOS personal computers

ALEXANDER G. SHMELYOV  
*Moscow State University, Moscow, Russia*

TESTAN (TEST and ANalysis) is an authoring shell and integrated psychometric package for development of computerized multiple-choice questionnaires, tests, and personality scales on personal computers that use MS-DOS. TESTAN allows (1) writing and editing items, interpretation messages, keys, norms, a bank of profiles, and so on, (2) collecting data in on-line or paper-and-pencil modes, (3) selecting the most discriminating items by means of correlation and factor analysis for practical use, (4) validating test scales and items according to external criteria and expert ratings, and (5) making multifactor assessment decisions after testing. TESTAN can be used to test student conceptual knowledge in any area in the form of multiple-choice questions. This report describes essential functional properties and facilities of TESTAN for psychometrists and applied psychologists.

Recently, some examples of specialized instrumental software have appeared that allow nonprogrammer users to make computerized tests (see review in Bartram, in press). Most of the systems are not yet menu-driven and do not provide direct access to all test components and results of testing, as far as I have been able to determine from the reports about such systems as, for example, the Computer Assisted Research Procedure (CARP; Backer-Johnsen, 1993) or the Microcomputer Assessment System (MAS; Hays et al., 1992). The system described in the present article, TESTAN (TEST and ANalysis), does not yet have a Windows interface, but it is already an easy-to-use, menu-driven integrated system that has contextual help and a set-up mode that allows users to connect the system to their favorite text editors.

The general rationale behind the development of TESTAN was to develop an open, modular, integrated, operational environment. All test components, data, and intermediate results (e.g., raw scores) are stored in text files with evident formats that can be analyzed not only by system-supplied tools but also by users' own statistical programs. They also can be easily exported to any well-known statistical package (Systat, SAS, SPSS, etc.). On this point, I agree with the principle declared by Schneider (1991), the author of Micro Experimental Laboratory (MEL), and have tried to refrain from using worksheets and record files (the strategy more appropriate for passive

users of ready-made programs) and to accept an orientation on text files (the strategy more appropriate for active users, developers of new methods and techniques).

## General Structure and Purposes

There are three different principal parts of the system: (1) MAINTEST, for editing test components, collecting, scoring, and interpreting test data, (2) TESTPLUS, for psychometric analysis of validity and reliability of the test as a whole and of separate items, including exploratory and confirmatory factor analysis and cluster HCS analysis (see Figure 1), and (3) EXPAN, for decision making and parallel expert rating.

The main application of TESTAN is as computer support for pilot experiments during psychological test development or test design. But TESTAN can be used for practical work if the user develops a special test for his or her special problem situation, as well as a tool for laboratory practical works for teaching students in psychometrics and test design. It is the current technological basis for a course on psychological testing at Moscow State University that I have taught for many years. After collecting test data, TESTPLUS teaches psychometric analysis by guiding students through the main traditional psychometric procedures.

## User Interface

The main idea behind the user interface in TESTAN is to use the rather developed skills in operating a standard text editor that most nonprogrammers possess. The user can link to the system any text editor (working with ASCII files) that he or she prefers. This feature allows the text files prepared for the paper-and-pencil version of a test to be used for the computerized version with little modification. Such use of text files is impossible in most systems based on worksheet technology, such as *Question Mark* (1992). The user of TESTAN needs only to surround the numbers of different test

---

The author would like to thank Vladimir Pokhilko, Vladimir Ryjov, Lewis Goldberg, John Digman, John Johnson, Nick Hammond, Jean McKendree, Boele DeRaad, and David Bartram for their contributions to the project. Correspondence should be addressed to A. G. Shmelyov, ul. Hertzena 4, kom.24, "Human Technologies," Psychological Faculty, Moscow State University, Moscow, 103009, Russian Federation (e-mail: shmel@ahumantec.msk.ru).

Note: The author has a direct financial interest in some of the software described in this paper.—*Editor*

**TESTPLUS (C) A. Shmelyov, Ht, Moscow, 1993. MAIN MENU.**

**Current test: Test name** **Current base: Base name**

NORMS	VALIDITY	RELIABILITY	FINEITY	ITEMS
<b>Distribut.</b> <b>Normality</b> <b>Robustness</b> <b>Table norms</b> <b>Criter. norm</b> <b>Stand. scale</b>	<b>Crit. input</b> <b>External v.</b> <b>Internal v.</b> <b>FA (scales)</b> <b>Confirm. FA</b>	<b>Half-split</b> <b>Alpha</b> <b>Retest input</b> <b>Retest rel.</b>	<b>SD input</b> <b>SD (scales)</b> <b>Cns. key inp.</b> <b>Consistency</b>	<b>Reliability</b> <b>Ext. validity</b> <b>Retest rel.</b> <b>Fidelity (SD)</b> <b>Exp. grouping</b>

**Exploratory FA**  
**Cluster analysis**  
**Congruency**

**Commentary line for the current item of pull-down menu.**

**Enter - Run. F1 - help. F4 - view result. F8 - check format.**

Figure 1. Main menu of the second subsystem TESTPLUS.

questions by square brackets and to put the special symbol “#” in front of the number of each item, and no more. This is possible to do very easily by means of the REPLACE ALL operation in MINIEDIT— a text editor supplement to TESTAN that supports the WORD STAR standard set of commands and works with a mouse and clipboard. MINIEDIT allows the files with the certain extensions specified by the main menu-driven monitor of TESTAN to be picked up.

A similar approach based on a text editor was used in MAS (Hays et al., 1992). But TESTAN requires fewer special symbols and less work in preparing text files for on-line testing. This relative ease of editing fits the requirement of iterative editing work on test items during test development and the educational purposes in laboratory work with the students.

In item writing, the test developer is limited only by the size of the monitor screen. The programs can recognize (and highlight) the screen coordinates for the number of the current alternative selected by the test taker for answering.

There is direct access to all test components (the item inventory, instructions, scale names, keys, norms, interpretation messages), which are saved in separate text files. Manipulating text files (instead of worksheets) becomes possible because of the presence of a special “format checking” routine that can be run by the function key F8 in any item of the TESTEDITOR submenu. This routine informs the user about any errors in file formats. It is possible to encode items and keys into binary files in order to protect the confidential information from being viewed by potential test takers.

Of course, there are some worksheets in the modes where they are necessary. For example, for entering the data from answer sheets (from paper-and-pencil forms of tests), TESTAN has a special “Blank Editor”—an electronic worksheet that can be automatically reformatted in accord with the number of items and the format of the answer sheet. There is also a special “Key Editor,” which represents the content of any item and any scale at the same time and highlights a “key response” right on the screen.

TESTAN creates report files (for individuals and for groups) that can be browsed and edited into final reports with the text editor before printing out. TESTAN works under MS-DOS. It was written in Borland Turbo-Pascal 6.0 with the Turbo-Vision package. The development of the Windows version is under way.

#### Modular Architecture

The user can keep in separate subdirectories different versions (different files) for each component of the same test (e.g., particular norms for male and female subjects). Each component can be loaded by means of the standard “peak-up window” with names of files. The number of tests that the user can keep and load from the hard disk into the system is not limited.

It is possible to prepare items in graphics in the popular standard of PCX library, which is well known to those who work with Paint Brush, Deluxe Paint, or similar graphic editors on personal computers. Of course, the programs

for presenting textual and graphic items are different in the system. But they are both based on the same standard of interface for on-line testing.

To save time, the test-on-line module can realize a simple algorithm of adaptive testing that ceases random presentation of items once the current percentage of correct answers reaches the upper or lower points (i.e., they cut the ranges for "good-fit" and "bad-unfit" scores or marks). The initial critical points can be specified by the user in the special control file. The longer the test procedure lasts, the closer these two points become to each other. This mode serves as a useful tool for testing conceptual knowledge (i.e., for achievement tests). If there is not a special control file (with parameters for adaptive testing) among files of test components (in the disk environment), the test-on-line module will present items in ordinary nonrandom sequential order as they are ordered in the file of items.

Another example of flexible, modular architecture is presented in the interpretive module: (1) if there is a special file of indicative (or typical) profiles, the program will calculate the similarity of the current factor profile to all indicative ones from this file; (2) if there is a special file with texts of integrative portraits (interpretive messages) for similar indicative profiles, these portraits will be included in the file of the current report.

The scoring interpretive module is separated from on-line testing routines as a separate executable file. Psychometric routines and expert ratings are located in separate modules as well. The user can make his or her own configuration of the suitable subset of "exe-files," depending on the situation (only testing or testing and feedback). In comparison with complete configuration (about 1 MB of hard disk), the miniconfiguration occupies not more than 200K and can be easily distributed on all terminal computers in computer classrooms or run from floppy disks as well.

As a result of the modular architecture, TESTAN can be used in combination with other systems. For example, it takes several minutes to pick up text files with items generated by a system such as *Norton Test Maker* (1993) and to run test-on-line mode under TESTAN with the new material. There are four levels of priorities in access to different parts and modules of the system (developer, administrator, operator, and client) to prevent access to confidential information for unauthorized people.

### Interpretive Facilities

I cannot pay special attention here to the standard and expectable details of the system. For example, scoring uses the traditional linear product "keys \* answers." It is easy to check the results of TESTAN by means of manual scoring.

The new property of TESTAN is a variety of interpretive models. There are four different forms and sources of interpretive information:

1. The standard factor scoring gives the user the standardized factor profiles (in the form of a bar chart) based on linear or converting table norms on sten-scale.

2. There is a file with "per-factor" interpretive messages. It is not sufficiently exact information for professionals,

but it is exact enough for generating the simple psychological feedback to the clients (subjects).

3. There is a changeable file with "indicative profiles." The user can obtain automatically the empirical typical profiles by means of cluster analysis for subjects. But it is not necessary to use only the empirical profiles. It is possible to build "abstract" hypothetical profiles. Each profile can be compared with the current profile of the subject. The significantly similar and contrast (opposite) profiles are pointed out in the output files prepared for professionals. In this way, we have more integrative interpretations.

4. The same structure of data was used in the unchangeable file—the "personality traits bank." It contains 480 personality trait names in the English version (Goldberg, 1992) and 2,090 names in the Russian version (Shmelyov & Pokhilko, 1993). By means of the special linear-matrix operator (convertor), the factor profile of any test can be transformed automatically into the standard profile (based on the Big Five factors in the English version and on the 15 original Russian factors in the Russian version of the personality traits taxonomy). The personality traits that show significant similarity with the empirical Big Five profile (of the current subject) are pointed out in the output file under the headline "Associated personality traits." This information is not final: the user can delete irrelevant traits from the report file by means of text editor.

But how can the user obtain the linear-convertor for his own new test? Of course, this procedure cannot be completely automated. To obtain a new convertor for your new test, you should evaluate your factor names by means of 50 or 25 bipolar Big Five markers (Goldberg, 1992; Johnson & Ostendorf, 1993). It is possible to do this in on-line mode in EXPAN (an EXPert ANalysis supplement to TESTAN). The narrative reports generator (Butcher, 1987) cannot be present in the universal configuration of TESTAN because of its close relatedness to the special content of a particular test or questionnaire.

### Iterative Cycle "Test–Analysis–Test"

If you use a standard statistical package you have no chance to get the results (e.g., factor loadings) in the format that can be read automatically by your computer testing shell. The specialized psychometric statistical package TESTPLUS gives the following contributions:

1. You can save the factor loadings (or biserial correlations with the total scores) in the format of new keys for your test. Then you can reduce the number of items extracted to only the informative items from the initial pool. Thus, in the case of analysis of a large pool of items (up to 240), TESTAN gives you a chance to go through several iterative cycles of "test–item analysis–test" very quickly. All steps of the "sequential strategy for test development" are available in TESTAN in the automatic mode (Jackson, 1970).

2. The names of the items of the hierarchical menu in TESTPLUS refer immediately to the psychometric topics: NORMS, RETEST RELIABILITY, EXTERNAL VALIDITY, ITEMS, SOCIAL DESIRABILITY, ALPHA COEFFICIENT, and so on. It is much easier to find the

suitable routine in TESTPLUS, especially for beginners in psychometrics (see Figure 1).

3. A special ITEM REDUCTION routine in TESTPLUS makes it possible to remove irrelevant (bad) items automatically not only from files of items and keys but from the data bank as well. As a result, you can have at once new norms for the reduced version of your test. This is the most powerful facility for test construction work. As a rule, it needs to be used after the item analysis is performed.

The main profit from using TESTPLUS in comparison with other specialized psychometric packages (i.e., CASPER; Chambers & Grice, 1991) is the possibility of analyzing much larger pools of items (up to 240 items vs. up to 30 items in CASPER).

On the basis of the psychometric package TESTPLUS, I have already developed many powerful multifactor tests and questionnaires, including a Russian-modified version of 16PF (Cattell, Eber, & Tatsuoka 1970), the original Russian 18-scales Test of Social Mentality (Shmelyov, 1991), the original Russian Test of Managerial Potential (Shmelyov, Grebenuk, Soloveychik, & Lepekha, 1993), and a Russian test SPOT of selection agents for insurance companies (Shmelyov & Krymov, 1994).

### Decision Making

EXPAN works under TESTAN as a satellite program. It allows the user (1) to attribute the different weights to different factors and to calculate the total ratings for different persons in multifactor (multicriteria) situations, (2) to check the expert validity of a new test (It is possible to obtain a cube of data— $60 \times 40 \times 40$ , subjects  $\times$  factors  $\times$  judges. The Alpha coefficient for the agreement between experts is used as a measure of expert reliability.), (3) to connect your test factor names to the Big Five personality factor markers and to get a rich interpretation material to your test, and (4) to perform on-line personality scaling in the paradigm of “semantic differential” or “repertory grids” (with supplied personality constructs) and to reconstruct individual and group semantic spaces.

For all items in all three parts of TESTAN, you can have contextual help from the corresponding chapter of the built-in manual. Because it is also a text file, you can print it out in order to get a hard copy of the manual.

### System Requirements

TESTAN can be installed on any personal computer with a hard disk, 640K RAM, and MS-DOS 3.0 or higher. The complete configuration takes about 1 MB of hard disk space. A color monitor is not necessary. The graphic testing module TESTDIAL requires a VGA graphic card.

### Availability and License Agreement

There are shareware and work versions of TESTAN that differ in their limits on numbers of items, scales, subjects, criteria, and profiles (see Table 1). To receive the English shareware version, you should send a blank formatted HD 3.5-in. diskette along with a return postage-paid mailer to A. Shmelyov, c/o CTI-Centre for Psychology, University of

**Table 1**  
**Numbers of Items, Scales, Subjects, Criteria, and Profiles for the Shareware and Work Versions of TESTAN**

Version	Items	Scales	Subjects	Criteria	Profiles
Shareware	60	5	60	2	50
Work	240	17	500	15	500

York, York, YO1 5DD, UK (e-mail: ctipsych@york.ac.uk). If you like the shareware version, you can send a signed order (your personal signature will be your licensing agreement) to the address in York and, at the same time, send an e-mail message to the author in Moscow. Information about the prices for different configurations of the work version and details about the licensing agreement are available upon request.

### REFERENCES

- BACKER-JOHNSEN, T. (1993). A general program for computer-based testing: Psychology software news. *Newsletter of the CTI-Centre for Psychology in York*, 3, 75.
- BARTRAM, D. (in press). Computer-based assessment. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (Vol. 9). London: Wiley.
- BUTCHER, J. M. (1987). *Computerized psychological assessment: A practitioner's guide*. New York: Basic Books.
- CATTELL, R. B., EBER, H. W., & TATSUOKA, M. M. (1970). *Handbook for the sixteen personality factor questionnaire (16PF)*. Champaign, IL: Institute for Personality & Ability Testing.
- CHAMBERS, W. V., & GRICE, J. W. (1991). CASPER: A personal computer package for exploring psychometrics. *Behavior Research Methods, Instruments, & Computers*, 23, 79-81.
- GOLDBERG, L. R. (1992). The development of markers for the big five factor structure. *Psychological Assessment*, 4, 26-42.
- HAYS, R. D., GILOGGLY, J. J., HILL, L., LEWIS, M. W., BELL, R. M., & NICHOLAS, R. (1992). A Microcomputer Assessment System (MAS) for administering computer-based surveys: Preliminary results from administration to clients in an impaired-driver treatment program. *Behavior Research Methods, Instruments, & Computers*, 24, 358-365.
- JACKSON, D. N. (1970). A sequential system for personality scale development. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology* (Vol. 2, pp. 61-96). Orlando, FL: Academic Press.
- JOHNSON, J. A., & OSTENDORF, F. (1993). Clarification of the five-factor model with a bridged big five dimensional circumplex. *Journal of Personality & Social Psychology*, 65, 563-576.
- Norton Test Maker* [Software manual] (1993). New York: Norton.
- Question Mark* [Software manual] (1992). London: Question Mark Computing.
- SCHNEIDER, W. (1991). Equipment is cheap, but the field must develop and support common software for psychological research. *Behavior Research Methods, Instruments, & Computers*, 23, 114-116.
- SHMELYOV, A. G. (1991). [The psychology of political counter-standing: A test of social mentality]. *Psychologicheskii Journal*, 5, 26-36. (In Russian. Also published as Shmelyov, A. G. [1995]. Test of social world view and attitudes of the populations toward the reforms in Russia. *Journal of Russian & East European Psychology*, 33, 6-25.)
- SHMELYOV, A. G., GREBENUK, A. A., SOLOVEYCHIK, A. S., & LEPEKHA, T. R. (1993). [The validity of the Test of Managerial Potential. *Vestnik Moscovskogo Universiteta*]. *Psychologia*, 3, 12-20. (In Russian)
- SHMELYOV, A. G., & KRYMOV, A. A. (1994). [SPOT—the computerized 7-scale test for insurance agents' selection]. Unpublished computer software manual. (In Russian)
- SHMELYOV, A. G., & POKHILKO, V. (1993). A taxonomy-oriented study of Russian personality-trait names. *European Journal of Personality*, 7, 1-17.