# Dr.Stat, a computerized statistical aid for psychology students

Marijke Engels-Freeke\* en Dirk Paul Flach\*\*

#### Abstract

In 1988 the development of a computerized tutorial program for education statistics was started at the University of Amsterdam. In this article, this program will be described. Also several other aspects of this computer program will be discussed. An important part of the article will focus on a description and evaluation of the program itself, as well as on experiences of teachers and students. We will also pay attention to the background of the program: reasons for developing the program, the development process itself, the way the program was set up, the area of statistics is addressed by the program, and how the program works.

### Introduction

Statistics is a difficult subject for many students. The traditional method of teaching statistics, which consists of lectures based on an introductory textbook, often fails to engender an adequate understanding of statistics. In view of the difficulties experienced by students, in the 1980's the statistics course at the Department of Psychology of the University of Amsterdam was supplemented with seminars. Following the introduction of the voluntary seminars, approximately half the psychology students participated. In the seminars, various topics, identified by the students as difficult, were explained in less technical terms and additional exercises were given to help the students to come to grips with the subject matter. In 1989, these seminars were, initially partially and later on completely, replaced by an interactive computer program called Dr.Stat. Several events led to the development of Dr.Stat and the ultimate abolishment of the seminars.

First, the University Administration Board had made additional grants available for the development and implementation of projects aimed at improvement and innovation in education. With respect to educational innovation the Board emphasized the importance of

<sup>\*</sup> Department of Psychological Methology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam.

<sup>\*\*</sup> Department of Experimental Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam.

keeping up with recent developments in computer technology and its application in education. In this period, the Department of Experimental Psychology was using a Computer-Assisted Education program (C.A.E. program) to assist in teaching psychological instrumentation. The positive results obtained with this C.A.E. program revealed the potential of this computer-based approach, and warranted further and extended use of the program. Consequently, the University Administration Board honored the request for a grant to develop a C.A.E. program to aid first year students in learning statistics.

The development of Dr.Stat was also stimulated by financial cuts of the faculty budget. The first year seminars were threatened to be discontinued due to the high costs involved, even though the usefulness of these groups for the students was evident. The large number of students participating annually in the labor-intensive seminars not only showed that additional teaching was indispensable, but it also showed that the high development costs of Dr. Stat would be outweighed by later savings in labor costs.

Finally, a very suitable authoring system was available for the development of an interactive statistics program and expertise in using this system was available. The group of the Computer Assisted Education-program at that time provided a Beta-test-site for the new authoring environment Authorware (then called Course of Action).

From the start Dr.Stat was intended to supplement the compulsory part of the statistics course and not to replace the entire statistics course. The material of the statistics course consisted of three books (Van den Brink & Koele, 1985, 1986, 1987) which covered the compulsory topics during the first year of the study. These topics were explained during lectures provided by two teachers of the Department of Methodology, who also had written the books. The books and lectures were supplemented by seminars which were run by two teaching assistants. Both the teachers and the teaching assistants expressed their satisfaction with this setup. In addition, the feedback provided by the students and the examresults demonstrated the adequacy of the setup. Replacing the entire course with an interactive computer program seemed both unnecessary and undesirable. Both teachers expressed doubts about the feasibility of addressing complex statistical concepts thoroughly and sufficiently by means of a computer program. They feared that the necessarily less thorough explanation of the statistical concepts provided by a computer program might result in insufficient understanding. An interactive computer program was judged to be useful to rehearse the subject matter and to provide additional practice with the concepts that were explained in the books and the lectures.

#### Development

The team responsible for the development of Dr.Stat consisted of programmers, statisticians, and a graphic designer. In addition, one of the teachers collaborated intensively with the team during the first year of the development. The development of a program that would meet the technical and educational requirements demanded close cooperation among the various disciplines. The programmers set out to study the books used in the statistics course. An inventory was made of the topics, as well as the relevant exercises and examples, that were to be included in the computer program. Next, a study was made of the problems that students encountered in studying the various statistical topics. The program developers made an inventory of the problems which were addressed during the first year seminars, and of the solutions and explanations discussed by the teaching assistants. The explanations and solutions that appeared not to pose a problem for the students served a starting point for the explanation of concepts within the computer program. Misunderstandings that occurred frequently helped to identify topics that had to be addressed explicitly and extensively in the program.

The results of the research were used to draw up a more detailed plan. Subsequently, the teacher provided feedback concerning this plan. After incorporating this feedback, the entire plan was divided into several modules. Each module dealt with issues relating to a single statistics topic, such as correlation, or measures of centrality. Each module was passed back and forth between the programmers and the teacher until agreement was reached that the module met the requirements and was acceptable for the students.

The program was implemented and made available to the students in a number of stages. The first year (1989-1990) of its use, Dr.Stat consisted of eight modules, which dealt with the descriptive part of statistics. The next year, 1990-1991, 11 modules were completed and added to the program, making up the current total of nineteen modules.

#### The subject matter and method of Dr.Stat

The aim of the project was to make available to the students a computer program to assist them in studying statistics. The objectives of the program were to present the required subject matter in a clear manner and to enable the students to work independently. It was required that the program should be structured in such a way that it would teach students how to solve problems, avoid common misunderstandings, and to engender sufficient understanding of statistics. These aims were achieved by adopting a method closely related to that presented by Berlin and White (1988). These authors proposed that subject matter should be presented at a concrete level, before moving on to a more formal level. Inbetween, the subject matter was also presented at a graphic level. This resulted in the presention of a topic at four sequential levels.

At the first level the topic is presented as a concrete problem avoiding the abstraction that students of statistics often find difficult. The second level of presentation is aimed at presenting and explaining statistical problems in graphical terms. The third level is more formal in that it includes a more abstract presentation of statistical concepts, notations and equations. The fourth and final level consists of several problems and exercises that provide the student with the opportunity to test his or her knowledge and understanding of the subject matter. Each module of Dr. Stat contains several topics that are presented at the four sequential levels mentioned. An example of the four sequential levels is given in this article.



Illustration 1. Concrete level



Illustration 2. Graphical level



Illustration 3. Formal level



Illustration 4. Exercises

Although some topics are less suited to this approach, this method was used as a point of departure in designing the program. In principle, each topic is presented at the four sequential levels. The majority of the modules deals with several statistical topics in succession.

The result of this approach is a hierarchic structure of the program. Subjects build upon earlier presented knowledge, and a fixed route is followed throughout the program. The developers felt that subjects were closely related ('a prerequisite hierarchy', Ohlsson, 1987). To accomplish this, mastery learning is used within Dr. Stat - a certain subject has to be completely understood before one can move on with the next subject. It is still possible, however, for students to jump to a certain part.

Each module is concluded with some additional exercises and a summary of the subject matter treated in the module. The modules and the topics within each module are presented in a hierarchical manner, which facilitates the accumulation of acquired knowledge. The hierarchical method of presentation is taken directly from the presentation in the books used in the statistics course. The content of the modules is shown in Table 1.

module	contents	chapter book	estimated time spent
Dr.Stat 1	Measures & Scales	2	1.0
Dr.Stat 2	Tables & Graphics	2	3.0
Dr.Stat 3	Measures of centrality	3	1.5
Dr.Stat 4	Measures of dispersion	3	3.0
Dr.Stat 5	Skewness	3	0.3
Dr.Stat 6	Transformations	3	1.5
Dr.Stat 7	Correlations	4	4.0
Dr.Stat 8	Regression	5	3.0
Dr.Stat 9	Probability theory	6	6.0
Dr.Stat 10	Random Variable	7	5.0
Dr.Stat 11	Probability distribution	8	9.0
Dr.Stat 12	Sample distributions	9	4.0
Dr.Stat 13	Estimation theory	10	5.0
Dr.Stat 14	Theory of testing	11	6.0
Dr.Stat 15	T-test	14	3.0
Dr.Stat 16	Tests for correlations	15	4.0
Dr.Stat 17	Medians en Proportions	16	4.0
Dr.Stat 18	Ranking numbers	17	3.0
Dr.Stat 19	Chi-square test	18	4.0
Dr.Stat Exam	1	-	6.0

Table 1: The modules of Dr.Stat, the contents of each module, the corresponding chapters in the books 'Van den Brink en Koele, Statistiek deel 1, 2 en 3', and the estimated houres spent studying the module.

### **Technical development**

Dr.Stat was developed with the authoring system Authorware Professional. This authoring system has facilitated the development of several important features of the program. Due to the modular structure Dr.Stat can easily be ported to the Windows platform.

Dr.Stat was initially developed for the Macintosh platform, more specifically the Macintosh Plus computer with a monochrome display at a resolution of 512 by 364 pixels. Although the presentation mode has not been updated, many graphics and animations have been included to illustrate various statistical concepts and problems.

# Use of Dr.Stat

The program starts with an introductory module which provides the student with an opportunity to familiarize himself or herself with the computer and the use of Dr.Stat. Subsequently the student follows the material in a fixed order. After entering his or her name, a student is automatically returned to the position in the program, where he or she stopped at the

previous session. However, this fixed order can be changed by the student. The student can return to an earlier chapter, or skip a module. Furthermore, the specific presentation of the subject matter by the program depends on the number and the kind of mistakes a student makes. If a student makes many mistakes, he or she will receive extra explanation and more specific feedback related to the mistakes made. Each module starts with the presentation of the expected duration of the module (see Table 1). This exected duration is computed as estimated time spent. Then the material is presented. Active involvement in the material is encouraged by presenting problems and exercises based on the subject matter of the module. Finally, at the end of the complete program a number of practice exams are administered.

### **Features of Dr.Stat**

The main objective of the Dr.Stat program is to help students understand statistics without additional assistance. No teacher is present when the students are using the program. To achieve this instructor independence, the program is designed to be as self-contained as possible. It is possible to use the program without the books on which the program is based. Several features were implemented to facilitate the use of the program.

### Feedback

When a student solves a problem or answers a question, he or she receives immediate feedback. The developers have tried to make the feedback as comprehensive as possible. The way feedback is given is very much dependent on the specific question, which makes it difficult to give general feedback principles. Nevertheless, the following aspects play an important role in the way feedback is given:

- In the event of an incorrect answer, the student is first encouraged to try again. Sometimes a hint is given to help the student to find the correct solution.
- With every question, usually several 'likely' errors are made, and for each of these likely errors the program provides explicite feedback.
- If the likely error is a solution to part of the problem (like the denominator of a ratio) the question will be rephrased to match the other part.
- If the student consistently fails to produce the correct answer, a short explanation of the problem is presented. Finally, the way to arrive at the correct response is explained. Typically, this happens after three to five tries.

# Calculate

The student may choose to respond to a problem by giving an equation rather than a number, so he/she does not need a calculator. The program evaluates the answer and judges the correctness of the solution. For example, as the answer to the calculation of a z-score one can enter (48-46.7)/2.367 instead of 0.549. Following this equation response, the resulting answer appears on the screen supplemented with judgement and feedback.

#### Tables

Within the program standard statistical tables are available which can be consulted anywhere in the program. An answer can be provided by copying and pasting an entry in a statistical table. The available tables relate to binomial probabilities, normal curve areas, chi-square distribution, student t-distribution, and Fisher-Z transformation of Pearson's product moment correlation coefficient.

# Logging

The program keeps track of each student's use of the program. The main goal is to offer students the possibility of continuing the program where they left off. This is a useful feature as students distribute the 70 hours that it takes to work through the entire program over several sessions. In order to keep track of students' progress, students are asked to enter their name and a personal number at the beginning of each session. The program stores data concerning individual students in a database and uses these data to return the student to the stage in the program where the student had last entered.

#### Students' results

The time spent on the different modules and the number of problems solved correctly is registered for each student. At any desired moment, students can request an overview of their results and compare these with the average students' results. The results of the individual student are presented graphically so that the student can get an immediate impression of his or her progress.

### Mailing system

The registration of the name and the number at the beginning of each session also serves to connect the student to the e-mail system. No special networking requirements are needed for this text-based system. As soon as a student registers within the program, he or she is able to send mail to other students and to the teachers, who are also registered within the program. The

aim of the e-mail system is to enable to students to request help from an instructor when they have trouble understanding a statistical problem, and to report problems or errors encountered in using Dr. Stat. As a result of this mailing system, each module of Dr.Stat can be continuously updated and corrected, because it facilitates localization and correction of the errors reported by recording the place in the program from which the students sent their e-mail messages.

#### Glossary

Dr.Stat contains a glossary of statistical terms that can be consulted. If a student requires additional explanation concerning a given statistical concept, he or she is referred to the module in which the term is explained extensively.

### Dr.Exam

From the feedback provided by the students when the program was used for the first time, a demand for additional exercises at the end of the program became apparent. To meet this demand, an additional module, called Dr. Exam, was added making up a total of 20 modules. Dr. Exam provides the students with the opportunity to do an exercise exam. The exam score is calculated immediately and feedback is provided on request. The program can present fixed exams, or create a new exam by randomly choosing exam items. Although Dr.Exam is a part of Dr.Stat, it can be used as an independent program.

#### In-house and out-of-house use

When Dr.Stat was first implemented, the program was meant to be used in the computer rooms at the Faculty. However it quickly became apparent that there was not enough room to accommodate the four to five hundred new students who enroll annually. Therefore, in 1992 the project TeleCOO was started, in which Dr.Stat was used by the students from their homes. Using a computer and modem in their homes, students were able to log on to the faculty network to use Dr.Stat. Overall, the results of this project were very positive; students liked the possibility of working at home instead of at the Faculty. However, this approach also had certain drawbacks such as the requirement of supporting several modem configurations and the time to establish connections. Therefore, it was decided to make the program available to students on floppy disks, so that they could install the program on their computers at home. This way students could still use the program at home, but without having to log-on to the faculty system. The possibility of working at home is increasingly popular with students; more

than half of the students who use Dr.Stat use the program at home. Dr.Stat is also in use in a some other Psychology Faculties in The Netherlands.

# The evaluation of Dr.Stat

Meijer (1989) investigated the effectiveness of the first eight modules of Dr.Stat, after they were used in the curriculum in 1989-1990. The first eight modules concerned descriptive statistics. The students who made use of Dr. Stat generally enrolled into the seminars that covered the rest of the material (i.e., other subject matter than descriptive statistics). This rendered impossible a simple comparison of exam results of subjects who had used Dr. Stat of students who had followed the seminars, because the students who made use of Dr.Stat the second part of the course followed the seminars. An exploratory comparison among students who did and students who did not work with Dr.Stat, and a comparison of students who did and students who did not work with Dr.Stat were 1.36 points higher (13.01 versus 11.65) than the average of the students who had not ( $t_{164} = 2.45$ , p = 0.02). The average exam results of students who had followed seminars were 1.51 points higher (13.03 versus 11.52) than the average of the students who had not done so ( $t_{164} = 2.79$ , p = 0.01).

The effectiveness of Dr.Stat was again investigated in the academic year 1992-1993. At that time all the modules from the program had been completed and the seminars had been discontinued. Students who had followed at least 15 of the 19 modules were compared with students who had not used Dr.Stat. On average, the exam results of students who had used Dr.Stat were 1.5 points higher (14.65 versus 13.14) than average results of the students who had not done so ( $t_{157} = 3.35$ , p = 0.001).

Because a possible experiment of the effectiveness of Dr.Stat has to be conducted in the curriculum and the students have free choice to follow Dr.Stat, it is not possible to make use of an experimental design. Although it is likely that the weaker and insecure students made use of the extra support provided by Dr.Stat, the results are open to alternative conclusions. Moreover, Dr.Stat is confounded with the lectures: it would be difficult to sort out how much better results are due to Dr.Stat and how much of the gain was due to lectures and seminars. To rule out the alternative explanations one could pretest all the students or make use of a covariate such as intelligence of school results. But there are still a few problems: Should the pretest measure the initial statistical level of the students (which is likely to be low for all students at

the beginning of the curriculum) or should it measure intelligence and/or may be motivation? In which way the frequency of attending lectures and/or working with Dr.Stat should be accounted for in the investigation? Future investigation should be aware of these problems to exclude all alternative conclusions.

### Experiences of students

Reactions of students to Dr.Stat have been consistently positive. This is probably best illustrated by the intensity with which the program is used. About half of the students population uses the program at one point during the statistics curriculum. This can be considered to be a large number as the use of Dr. Stat is voluntary. For example in 1997, about 500 students started the first year of Psychology. 123 Students received a copy of Dr. Stat for home use; even more students may actually have used the program at home, since students were allowed to make copies of the program themselves. At the same time, 162 students used parts of Dr. Stat in the computer rooms at the University.

The students' experience in using the program was also evaluated in a more systematic manner. Meijer's research showed that students regarded both Dr.Stat and the seminars as useful and pleasant, although the appreciation of the seminars was slightly higher.

Dr. Stat was evaluated annually by having the students complete a questionnaire concerning their experience with the program. Once the development of the program had neared completion, evaluation using the questionnaire was discontinued. Students can still communicate their opinions about the program by means of the e-mail system, or by contacting the C.A.E coordinator. The majority of the eactions remain positive. In particular the possibility of using the program at home is much appreciated. The students' appreciation of the statistical program has increased since the full program became available.

#### Professors' experiences

Initially the participating professors of statistics were not very enthusiastic about the development of Dr.Stat. However they are now satisfied that the program does provide a good means to rehearse and to deepen understanding of the statistical subject matter. They do believe that the program is too limited to clarify more difficult concepts, because the subtle differences, that are often the crux of an explanation, are not easy or impossible to realize on a screen. The communication in a lecture is more natural and more flexible and discussion is possible. For these reasons the lectures and the books are still judged to be indispensable.

#### **Future developments**

The positive experiences with the Dr.Stat program has encouraged us to continue to develop the system. At present, three avenues of further development are being considered: translation of the program into other languages, incorporation of internet facilities, and the enhancement of the interface.

In 1996 a project aimed at translating the program was started with funding from the European Union (the SOCRATES project). This project involves a collaboration between the Universities of Amsterdam, York and Santiago de Compostela. So far, the first eight modules have been translated into Spanish and English. Additional funding to support further translation of the program is currently being sought.

A second avenue of further development became necessary with the translation of the program. Now that translated versions of Dr. Stat have appeared, students at an increasing number of universities will be able to use the program. At the same time, the growing dissemination of the program will make it difficult to retain control over the process of updating, maintenance, and debugging. In addition, the existence of various versions of Dr. Stat may result in a divergence of these versions. To prevent this from happening, we intend to use the internet for the distribution of updates and possibly to maintain core functionality of the program.

We are now considering updating the interface of Dr. Stat. The interface has remained virtually unchanged throughout the development of the program and is beginning to show its age. Main drawbacks are that it is still in black and white, and at a low resolution (512 by 364 pixels). So far, it has not been a problem because the program was designed with these limitations in mind and it is widely illustrated. Nevertheless, we want to keep Dr. Stat in pace with time.

### References

- Balk, V.A. (1994). Statistiek voor studenten Psychologie. In: M. J. A. Mirande (Ed.). De kwaliteiten van computerondersteund onderwijs: voorbeelden uit het hoger onderwijs (p 197-214). Bussum. Coutinho.
- Berlin, D. F., & White, A. L. (1988). A model for research, design, and selection of instructional uses of technology in mathematics and science education. In: *Proceedings of* the 27th International Conference of ADCIS. (p. 234-238)
- Meijer, K. (1989). Evaluatie van het eerste deel van het C.O.O.-project Dr.Stat. Stageverslag, UvA, Faculteit Psychologie, Amsterdam.

Ohlsson, S. (1987). Some principles of intelligent tutoring. In: Lawler, R.W. & Yazdani, M. (Eds). Artificial Intelligence and Education (Vol. 1): Learning Environments & Tutoring Systems. Norwood, Ablex Publishing.

Van den Brink, W. P., & Koele, P. (1985) Statistiek, deel 1, Datareductie. Meppel. Boom.
Van den Brink, W. P., & Koele, P. (1986) Statistiek, deel 2, Theorie. Meppel. Boom.
Van den Brink, W. P., & Koele, P. (1987) Statistiek, deel 3, Toepassingen. Meppel. Boom.

Ontvangen: maart 1998 Geaccepteerd: januari 1999

54