Animated diagrams in teaching statistics

KARL F. WENDER and J.-SEBASTIAN MUEHLBOECK
University of Trier, Trier, Germany

In this study, we investigated whether computer-animated graphics are more effective than static graphics in teaching statistics. Four statistical concepts were presented and explained to students in class. The presentations included graphics either in static or in animated form. The concepts explained were the multiplication of two matrices, the covariance of two random variables, the method of least squares in linear regression, α error, β error, and strength of effect. A comprehension test was immediately administered following the presentation. Test results showed a significant advantage for the animated graphics on retention and understanding of the concepts presented.

The use of computers in academic teaching has become increasingly popular. Technical developments in computer technology facilitate the use of diagrams, graphics, animations, and interactive animations in talks and lectures, going far beyond the limitations of the traditional blackboard.

It is widely accepted that diagrams assist the comprehension of presented materials. Traditional teaching has relied on static examples and illustrations for aiding the understanding and retention of abstract concepts.

Films and animated graphics are essential elements of education in the natural sciences (Strittmatter, 1994). The improvements of computer technology applications ease the integration of movement, films, and animations in instructional presentations, which help draw and sustain the attention of the spectator. The attention-drawing effects of movements are considered to be automatic and innate, because they serve an evolutionary purpose.

One might argue about the benefits of animations, however, on the Internet. As one browses the World-Wide Web, one finds many examples that demonstrate how distracting animations can be if they are not relevant. By drawing the attention of the viewer, they can even impede the elaboration of more important information and become a "source of distraction" (Mayer, Heiser, & Lonn, 2001). Levin, Anglin, and Carney (1987) concluded that the use of additional graphics "for cosmetic reasons" should be abandoned. It is true that these ideally sustain the viewer's interest. But excessive cognitive load, resulting from too much information being presented, may lead to a decrease of cognitive elaboration. Contemplation of animated displays demands additional cognitive effort, which may surpass the spectator's information-processing capacity and impair learning (Lewalter, 1997). Since cognitive load is a function of previous knowledge, only a thorough analysis of domain-specific knowledge

can ensure a profitable match between the graphical content and the learner's information-processing capacity.

Pictures contain spatial information, which is crucial in many mental processes. Originally proposed by Johnson-Laird (1983) in his theory of mental models, this has been recently confirmed by Knauff, Mulack, Kassubek, Salik, and Greenlee (2002), who have demonstrated the influence of visuospatial imagery on the performance of logical reasoning tasks using fMRI studies.

Mayer and Sims (1994) emphasize the property of pictures as memory aids in the learning of descriptive concepts. Spatial constellations and sequences are supposed to account for increased retention in the learning of procedural concepts. The effects of animations on learning have been extensively studied by Schnotz, Boeckheler, Grzondziel, Gaertner, and Waechter (1998).

Research so far indicates that animations are more advantageous for overall retention of subject matter if the concepts or rules taught contain a temporal course or progression, movement, or spatial relations (Rieber, Boyce, & Assad, 1990). This applies to concepts whose dynamics are hard to imagine. In procedural instructions, animations not only help to identify relevant objects, but also demonstrate the course of the action (Palmiter & Elkerton, 1993). In their meta-analysis, Betrancourt and Tversky (2000, p. 326) point out that an "animation is likely to be useful when the learning material entails motion, trajectory, or change over time, so that the animation helps to build a mental model of the dynamics." Generally speaking, if an animation is likely to depict more distinctly and clearly features that are to be integrated into a viewer's mental model, its benefits are to be superior to those conveyed by static graphics. These would require additional explanations or ancillary symbols (e.g., arrows) to specify interdependencies or sequences, thereby filling the display with additional information. By extension, successful animations may also facilitate logical inferences and transfer.

Weidenmann (1994) proposed a classification of pictures into realistic (photos or representational drawings), analogous (illustrations of objects or scenes), and logical

Animations can be obtained by writing an e-mail to the first author. Correspondence should be addressed to K. F. Wender, FB1-Psychology, University of Trier, 54286 Trier, Germany (e-mail: wender@uni-trier.de).

pictures. Whereas analogue diagrams schematize objects or parts of objects that exist in the real world, logical pictures are culturally determined illustrations depicting abstract structures, relations, quantities, and processes, as in structure or flow charts, for example.

In analogue diagrams, it is quite common to have animations that show, for example, how a pulley works (Hegarty, 1993). Logical diagrams depict abstract concepts and their relationships. So far research investigating the effects of animated graphics has been mainly focused on their use as realistic or analogous pictures. A possible beneficial effect of the use of animated graphics for abstract subject matter has not yet been empirically confirmed.

In this paper, we are primarily interested in the use of animations in teaching concepts in multivariate statistics which are inherently logical. Several authors provide animated Java applets in their on-line tutorials for teaching statistical methods (Lane, 1999; Lane & Tang, 2000; Schulmeister, 1995). In this domain, however, there are very few empirical studies that test whether or not animated diagrams are superior to their traditional static counterparts. Given the findings mentioned above, mathematical concepts seem to be especially suitable for investigation since the dynamics of multivariate statistical concepts are hard to imagine.

The concepts presented in our experiment were thought to fulfill the requirements for learning material as stated by Betrancourt and Tversky (2000). In addition, we took the possible differences in spatial abilities into account. We expected high spatial ability to facilitate the construction of a mental model, rendering animations more beneficial for low spatial ability students.

Given the additional effort required in the design of animated material, there is a need for theoretical conceptions and empirical research, since instructional success does not necessarily seem to accompany the use of such material. But the possibility of animations having a detrimental effect should be ruled out.

METHOD

Participants

The experiment was conducted with participants from a class of graduate students from our Department of Psychology. The course chosen was a one-semester 2-h lecture on multivariate statistics. There were 112 participants, all of whom were registered as full-time psychology students.

Procedure

At the beginning of the first lecture, the students were asked to participate in an experiment. They were told that its purpose was to obtain representative association norms for a sample of German nouns. The data were to be collected under two different conditions. The whole group was randomly divided into two subgroups; one group stayed in the lecture hall, while the other group was led into an identical room next door.

The latter group was given the word-association questionnaire, whereas the first group was told that they would review a few methodological concepts from previous classes. After about 45 min, the tasks were changed. Now the second group answered the ques-

tionnaire, and the first group reviewed the methodological concepts. This procedure was used to conceal the aim of the experiment, which was to investigate the effect of the animations.

After the word-association questionnaire, each group completed two subtests of a standardized intelligence test measuring spatial ability (Intelligence Structure Test 2000; Amthauer, Burkhard, Liepmann, & Beauducel, 1999). Spatial ability was measured as a control.

Materials

We selected four concepts for experimental investigation: multiplication of two matrices, the covariance of two random variables, the method of least squares in linear regression, and the concepts α error, β error, and strength of effect.

For each of these concepts, an animated presentation was developed using Macromedia Flash 5. For example, a scatter plot of data points was used to portray the method of least squares (see Figure 1). Then a regression line appeared. For each data point, the vertical distance from the regression line (i.e., the residual) was indicated by a short line. Then the residuals expanded to squares, which moved to the right and merged into a large square representing the sum of squares. Then the regression line rotated slowly, changing the residual for each data point and, at the same time, the corresponding square. Consequently, the sum of squares varied in size, depending on the slope of the regression line. The dot on the curve, depicted at the bottom right, moved synchronously with the sum of squares, thereby indicating the minimum.

Similar animations were developed for the remaining three concepts. For α error, β error, and strength of effect, two normal distributions were drawn (see Figure 2). Colored areas showed the error probabilities and the resulting power. To illustrate the interdependence of these criteria, α , δ , and N subsequently varied with dynamically adjusting β and power areas and distribution shape. For each of the animated presentations, static versions for the static condition were obtained by taking screen shots.

In the first group, the static condition was presented first, and the word association questionnaire with the spatial ability testing followed. The second group started with the association questionnaire and the spatial ability tests before viewing the animated presentation. Both presentations were given by the first author and conveyed the same information. The rationale behind the sequencing was the attempt to rule out the possibility of getting better results in the first group simply because of effects of fatigue in the second group. That is, we wanted to be conservative with respect to our hypothesis. Immediately after the presentation of the methodological concepts, the participants received a 10-item comprehension test. This was the dependent variable in the experiment. The questions were formulated to test for overall retention and understanding of the concepts presented. Three questions assessed matrix multiplication; another three addressed the concept of covariance. One open-ended question related to the method of least squares. The remaining questions assessed the understanding of the α error, β error, and power concept in significance testing. Questions included a simple computation of two given matrices and three open-ended questions that could be answered in one or two sentences—for instance, “Why does the power increase if a larger sample size is used?” The other questions could be answered in a few words. Three questions were in multiple-choice format.

RESULTS

First, the results from the two subtests of spatial ability were combined to a single score. Second, a median split separated participants into two groups: a high spatial ability (HSA) group and a low spatial ability (LSA) group. The data were analyzed by a 2×2 analysis of variance, with *animation* (animated vs. static) and *spa-*

Figure 1. Screen shot of the animation used to illustrate the method of least squares. (See text for further explanation.)

tial ability (HSA vs. LSA) as factors. The score in the comprehension test served as the dependent variable. The possible scores ranged from 0 to 10.

The factor *animation* showed a significant effect [$F(1,111) = 9.09, p < .003$]. This factor accounted for 44% of the variance ($\eta^2 = .44$). The results are shown in Figure 3. The factor *spatial ability* also reached significance [$F(1,111) = 9.41, p < .003$]. HSA students scored higher in the comprehension test.

The interaction did not reach significance. However, LSA participants tended to profit more from the animation [$F(1,111) = 1.81, p = .18$]. This is illustrated in Figure 3, the mean difference for the LSA group being larger than that of the HSA group. And, finally, we did not find any gender differences.

DISCUSSION

In teaching abstract concepts, such as those in statistics, it is quite common to use diagrams for illustration. Our results show that animation of abstract subject matter can have a beneficial effect in the teaching of statistics. For both groups in our experiment, the same depictions were utilized and explained in the same manner. Nevertheless, the animated graphics led to higher performance. When designing the animations, we tried not to clutter the presentations, but to emphasize the relationships between the essential subconcepts. Therefore, we animated causal or functional relationships—that is, the growth of the sum of squares as a consequence of the changing slope of the regression line.

Figure 2. Screen shot of the animation explaining α error, β error, and strength of effect.

Figure 3. Average scores for the animated and the static groups, separated by spatial ability. HSA, high spatial ability. LSA, low spatial ability.

The participants were not aware of the real nature of the experiment. Thus we ruled out the possibility of the experimental treatment changing expectations and motivations on the part of the students and influencing their performance. A known experimental goal tends to instigate higher motivation in the treatment group. On the other hand, the static group may try to compensate for an anticipated deficit. It is difficult to predict which of these effects will occur. Therefore, we tried to avoid these problems by concealing what we actually tested for, and we revealed the aim of the experiment only after the administration of the comprehension test.

There is a growing interest in issues concerning the requirements for appropriate instructional design of animated visualizations. Lowe (2001) investigates alternatives to realistic presentations of the behavior of complex dynamic systems. Taking into account the learners' perceptual and cognitive capacities, he introduces systematic steps for designing animations.

Animation of abstract concepts is different from animating realistic or analogue pictures. When one is working with realistic or analogue pictures, it is more obvious which parts to animate. Therefore the authoring of animations for logical diagrams is not quite as straightforward as it is for analogue diagrams. We think it would be beneficial for the understanding of abstract concepts if animations were focused on functional relationships of subconcepts, making them explicit by portraying interdependencies. This, however, needs further investigation. It would be especially desirable to develop methods that permit an analysis not only of abstract concepts, but also of the relationships between their components.

REFERENCES

- AMTHAUER, R., BURKHARD, B., LIEPMANN, D., & BEAUDUCEL, A. (1999). *Intelligenz Struktur Test IST 2000*. Göttingen: Hogrefe-Verlag.
- BETRANCOURT, M., & TVERSKY, B. (2000). Effect of computer animation on users' performance: A review. *Travail Humain*, **63**, 311-329.
- HEGARTY, M. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory & Language*, **32**, 717-742.
- JOHNSON-LAIRD, P. N. (1983). *Mental models. Toward a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.
- KNAUFF, M., MULACK, T., KASSUBEK, J., SALIK, H. R., & GREENLEE, M. W. (2002). Spatial imagery in deductive reasoning. *Cognitive Brain Research*, **13**, 203-212.
- LANE, D. M. (1999). The Rice Virtual Lab in Statistics. *Behavior Research Methods, Instruments, & Computers*, **31**, 24-33.
- LANE, D. M., & TANG, Z. (2000). Effectiveness of simulation training on transfer of statistical concepts. *Journal of Educational Computing Research*, **22**, 383-396.
- LEVIN, J. R., ANGLIN, G. J., & CARNEY, R. N. (1987). On empirically validating functions of pictures in prose. In H. A. Houghton (Ed.), *The psychology of illustration* (Vol. 1, pp. 116-135). New York: Springer-Verlag.
- LEWALTER, D. (1997). Kognitive Informationsverarbeitung beim Lernen mit computerpräsentierten statischen und dynamischen Illustrationen [Cognitive information processing in learning with computer-presented static and dynamic illustrations]. *Unterrichtswissenschaft*, **25**, 207-222.
- LOWE, R. (2001). *Perceptual and cognitive challenges to learning with dynamic visualizations*. Retrieved June 18, 2002, from <http://www.iwm-kmrc.de/workshops/visualization/lowe.pdf>.
- MAYER, R. E., HEISER, J., & LONN, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, **93**, 187-198.
- MAYER, R. E., & SIMS, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, **86**, 389-401.
- PALMITER, S., & ELKERTON, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human-Computer Interaction*, **8**, 193-216.
- RIEBER, L. P., BOYCE, M. J., & ASSAD, C. (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer Based Instruction*, **17**, 46-52.
- SCHNOTZ, W., BOECKHELER, J., GRZONZDZIEL, H., GAERTNER, I., & WAECHTER, M. (1998). Individuelles und kooperatives Lernen mit interaktiven animierten Bildern [Individual and cooperative learning with interactive animated pictures]. *Zeitschrift für Pädagogische Psychologie*, **12**, 135-145.
- SCHULMEISTER, R. (1995). *Grundlagen hypermedialer Lernsysteme: Theorie-Didaktik-Design* [Principles of hypermedia learning systems: Theory-didactic-design]. Frankfurt: Addison Wesley.
- STRITTMATTER, P. (1994). Wissenserwerb mit Bildern bei Film und Fernsehen [Knowledge acquisition using pictures from film and television]. In B. Weidenmann (Ed.), *Wissenserwerb mit Bildern* (pp. 177-194). Bern: Hans Huber.
- WEIDENMANN, B. (1994). Informierende Bilder [Informative pictures]. In B. Weidenmann (Ed.), *Wissenserwerb mit Bildern* (pp. 9-58). Bern: Hans Huber.