ROLF SCHULMEISTER*

# TAXONOMY OF MULTIMEDIA COMPONENT INTERACTIVITY
## A CONTRIBUTION TO THE CURRENT METADATA DEBATE

The term "interactivity" frequently occurs in the literature about learning management systems and e-learning and plays a special role in drafts of metadata for learning systems. Interactivity in LOM for instance is simply scaled from 1 to 5 which submits the concept to subjective judgement. The paper tries to define interactivity as distinct from interaction, communication or navigation. Interactivity denotes the actice manipulation of learning objects by the learner. The paper proposes a scale for interactive multimedia objects according to two dimensions: manipulation the representation form of learning objects and manipulating the content of learning objects in order to receive feedback.

*Key Words:* learning objects, metadata, LOM, interactivity, taxonomy of interactive learning objects

* University of Hamburg, D, schulmeister@uni-hamburg.de

Anyone dealing with the development of modular multimedia learning systems who has had frequent recourse to the metadata drafts of Ariadne (Alliance of Remote Instructional Authoring and Distribution Networks of Europe; http://ariadne.unil.ch), IEEE LOM (Learning Objects Metadata project group of IEEE; http://grouper.ieee.org/ groups/ ltsc/ wg12), or other institutions, can't help being aggravated by the fact that these internationally discussed metadata concepts fail to define important categories like interactivity in multimedia applications and multimedia leaning systems in a form of any practical value for the construction and didactics of such learning systems. The Ariadne project, for example, defines interactivity in the form of a scale ranging from *high* to *low* (http://ariadne.unil.ch/Metadata/ariadne_metadata_v3final1.htm). Such a scale can only yield subjective entries from the developers of learning systems, since learning program designers would have to decide for themselves whether their programs require or provide a high, medium, or low form of interactivity. The Learning Objects Metadata (LOM) standard draft from the IEEE organisation, which is based on proposals by Dublin Core and Ariadne, contains the same formal definition of interactivity:

| | | Description | Value | Restricted/Open Vocabulary | Datatype | Notes |
|---|---|---|---|---|---|---|
| 5.1 | Interactivity Type | The flow of interaction between this resource and the intended user. In an expositive resource, the information flows mainly from this resource to the learner. Expositive documents are typically used for learning-by reading. In an active resource, information also flows from the learner to this resource. Active documents are typically used for learning-by-doing. Activating links to navigate in hypertext documents is not considered as an information flow. Thus, hypertext documents are expositive. | Single value | restricted vocabulary: 3=Active 4=Expositive 5=Mixed 6=Undefined | Vocabulary | Expositive documents include essays, video clips, all kinds of graphical material and hypertext documents. Active documents include simulations, questionnaires and exercises. |
| 5.2 | Learning Resource Type | Specific kind of resource, most dominant kind first. | ordered list (minmax: 10 items) | open vocabulary with best practice: 3=Exercise 4=Simulation 5=Questionnaire 6=Diagram 7=Figure 8=Graph 9=Index 10=Slide 11=Table 12=Narrative Text 13=Exam 14=Experiment 15=Problem Statement 16=Self Assessment | Vocabulary | * |
| 5.3 | Interactivity Level | The degree of interactivity between the end user and this resource. | Single value | Restricted vocabulary: 0=very low 1=low 2=medium 3=high 4=very high | String (minmax: 1 char) | * |

Tab. 1: IEEE unapproved standards draft, 6 February 2000 IEEE P1484.12/D4.0

IEEE classifies interactivity first by type of learning resource and second by the level of interactivity. Ariadne on the other hand knows only one category for interactivity, which is defined in the same way as the interactivity level in IEEE/LOM: as an ordinal (nominal?) 5-step scale from *very low* to *very high*.

Several critics have complained that such a definition is meaningless and of no practical help to the developers of learning systems and script writers. Depending on the subjective viewpoints and individual impressions of authors, interpreters, observers, users or evaluators, a high or low level of interactivity can refer to very different forms of a user's freedom of action or behaviour. One author may emphasize the frequency of interactivity, another the quality of interactivity, and a third the media character. We need a theoretical reference system, a qualitative framework of categories, to arrive at a taxonomy of interactivity that will meet with international agreement.

In this paper, I would like to propose a scale for the metadata type *interactivity*. My chosen starting point is a screen from a multimedia learning system, or a web page in a learning platform, containing multimedia components other than pure text in individual "containers" (frames). Multimedia components in this context mean images, diagrams, animation sequences, video clips, audio samples, or tables, formulas, *JavaApplets* and *Flash* programs. Using such a page as a point of reference, we have to ask how much freedom of action the author has granted to users of the page, or which types or levels of interactivity have been chosen for the multimedia component.

Authors or evaluators of learning programs frequently call such programs "interactive", although the web pages of the learning platform or the screens of their multimedia programs do not in fact contain interactive elements. A page like the one illustrated beneath contains only menus and buttons for navigating to other chapters or pages of the learning program.
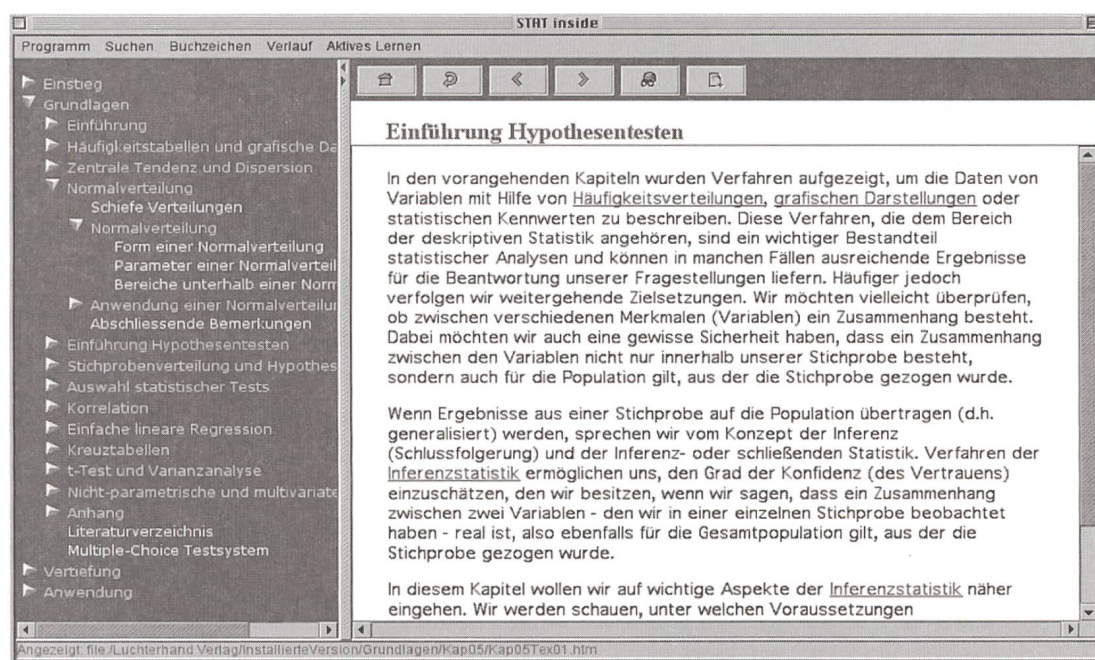
**Einführung Hypothesentesten**

In den vorangehenden Kapiteln wurden Verfahren aufgezeigt, um die Daten von Variablen mit Hilfe von Häufigkeitsverteilungen, grafischen Darstellungen oder statistischen Kennwerten zu beschreiben. Diese Verfahren, die dem Bereich der deskriptiven Statistik angehören, sind ein wichtiger Bestandteil statistischer Analysen und können in manchen Fällen ausreichende Ergebnisse für die Beantwortung unserer Fragestellungen liefern. Häufiger jedoch verfolgen wir weitergehende Zielsetzungen. Wir möchten vielleicht überprüfen, ob zwischen verschiedenen Merkmalen (Variablen) ein Zusammenhang besteht. Dabei möchten wir auch eine gewisse Sicherheit haben, dass ein Zusammenhang zwischen den Variablen nicht nur innerhalb unserer Stichprobe besteht, sondern auch für die Population gilt, aus der die Stichprobe gezogen wurde.

Wenn Ergebnisse aus einer Stichprobe auf die Population übertragen (d.h. generalisiert) werden, sprechen wir vom Konzept der Inferenz (Schlussfolgerung) und der Inferenz- oder schließenden Statistik. Verfahren der Inferenzstatistik ermöglichen uns, den Grad der Konfidenz (des Vertrauens) einzuschätzen, den wir besitzen, wenn wir sagen, dass ein Zusammenhang zwischen zwei Variablen - den wir in einer einzelnen Stichprobe beobachtet haben - real ist, also ebenfalls für die Gesamtpopulation gilt, aus der die Stichprobe gezogen wurde.

In diesem Kapitel wollen wir auf wichtige Aspekte der Inferenzstatistik näher eingehen. Wir werden schauen, unter welchen Voraussetzungen

*Fig. 1: Typical page from the program STATinside (Stadler 2001)*

Calling such an application interactive confuses navigation and interaction. Navigation only serves to control the order of pages, change the display or select a page to be viewed. Interactivity must be strictly distinguished from navigation. To me, interaction means controlling the object, subject or contents of a page.

## Proposal of a Taxonomy of Multimedia Components

In the following, I distinguish six levels of multimedia component modelling in a learning system, differentiated by the level of interactivity offered to the user, and illustrate these six levels with illustrations from several learning systems or programs:

### Level I: Viewing objects and receiving

The text uses pre-fabricated multimedia components which the user may look at (images, diagrams) or play (sound, video, Flash etc.). The viewer or listener has no way of influencing the component representation. At this level of interactivity, the user may only watch, read, or listen to the multimedia components. The multimedia components only serve for illustration or information. Their contents remain constant. Strictly speaking, this should be called Level 0 of interactivity, because — beyond calling up a picture or starting a playback — there is in fact no interactivity.
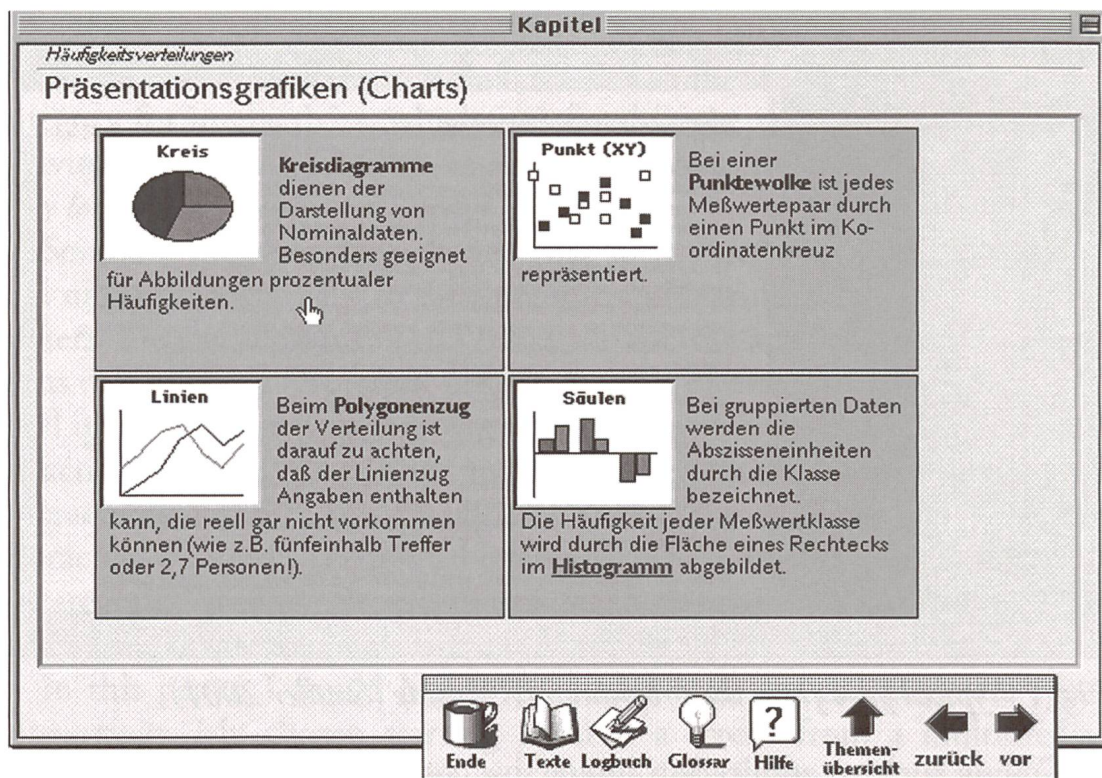
*Fig. 2: Screenshot from LernSTATS, a program for learning statistics (Schul-meister/Jacobs 1992-1996)*

This page merely informs students about which type of diagram is suited to which kind of data

## Level II: Watching and receiving multiple representations

At the next level, the multimedia components are pre-fabricated as well, but there are several variants for some of them. By clicking an image, the user may display another picture in the same frame (e.g. animated GIF), or, more generally speaking, the user can change the contents of the container by clicking on the multimedia component frame, selecting options or menu items, or via Hypertext links. Several versions of a diagram may be displayed in a sequence, or several music clips, video clips or animations played. At this level of interactivity, the result of interaction, i.e. the component representation, again does not allow anything beyond watching, and the multimedia components only serve for illustration or information.
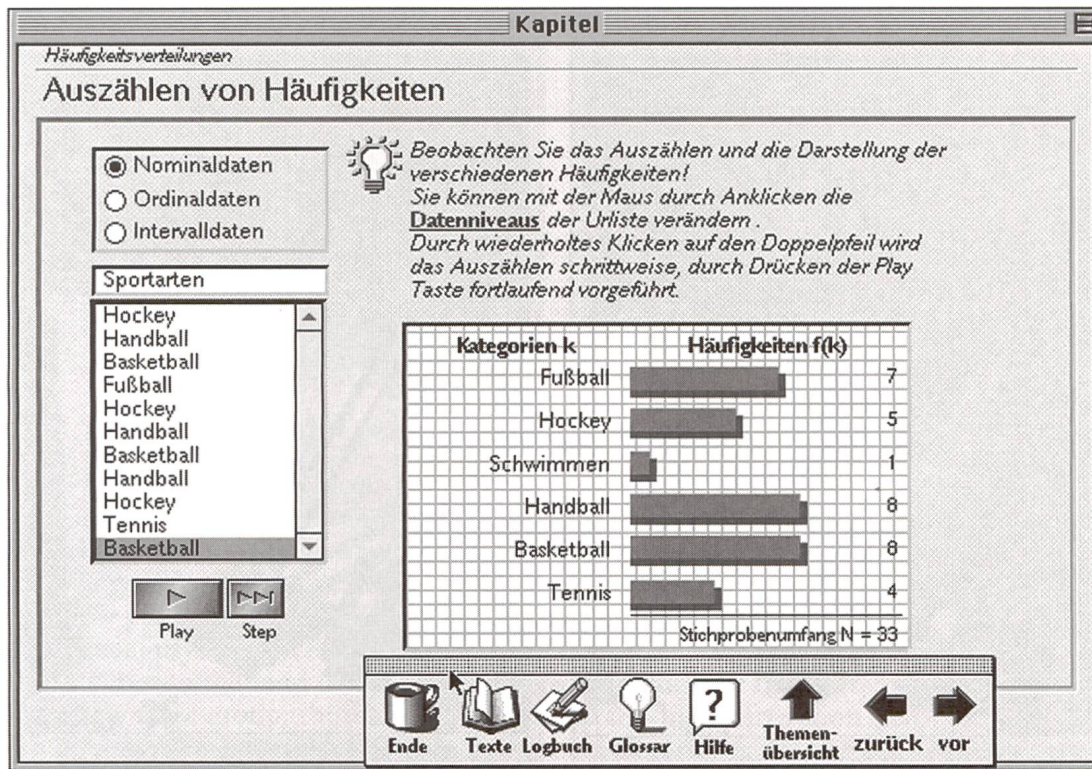
*Fig. 3: Multiple representations can be displayed by selecting a data set and pressing play (Schulmeister/Jacobs 1992-1996)*

The student can select several data sets and have a bar diagram displayed for each. This action can be repeated, but data and type of diagram cannot be changed.

## Level III: Varying the form of representation

In direct manipulation of the component, the user may for example scale a two-dimensional diagram, rotate the display of a three-dimensional animation, or jump to other segments of a video by clicking interactive objects in video sequences. This level of interactivity for the first time allows the user to feel in control of the multimedia component representation, to view the component from different perspectives or in different sizes or actively navigate within it. It must be stated however, that the three-dimensional object or interactive VR movie themselves remain unchanged, of course: the user action only manipulates the form of representation, not the contents. This level of interactivity is important for student motivation, even though the multimedia object remains constant.
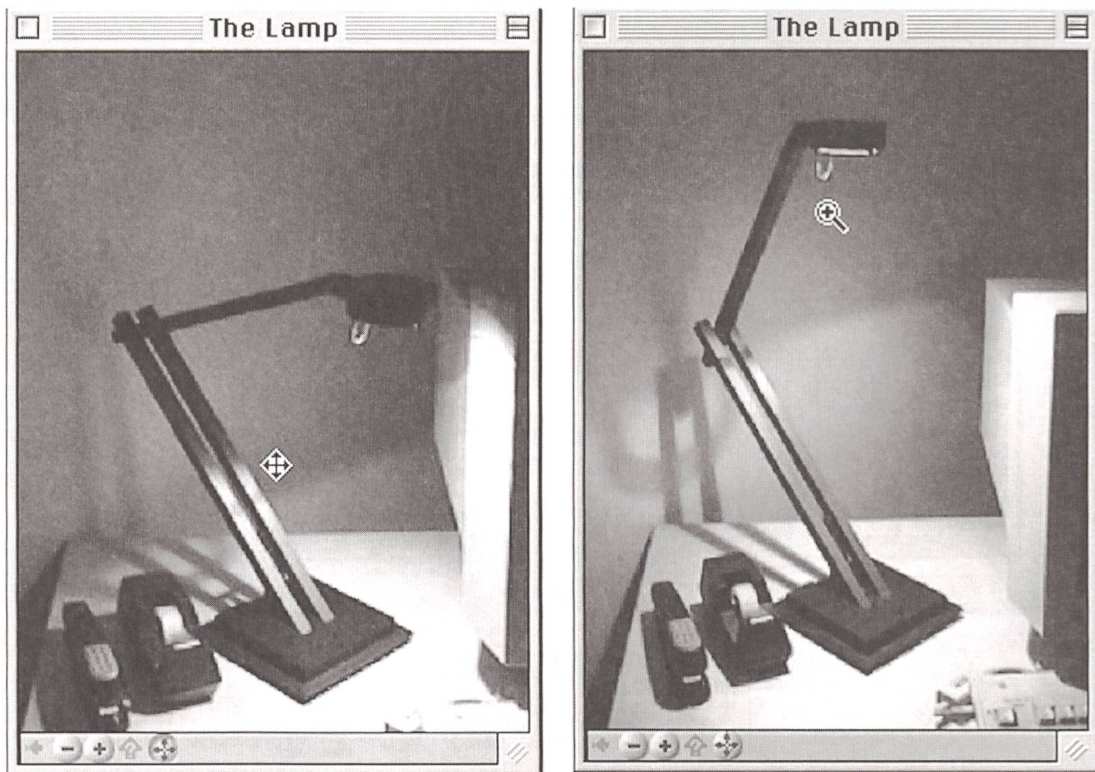
*Fig. 4: Direct manipulation of interactive 3D objects in VR movies (DeskLamp.mov, Apple Computer 1998)*

The student can move the video object (the desk lamp) with the mouse and thus manipulate the object representation and light beam. But the object itself remains the same.

Level IV: Manipulating the component content
The multimedia component contents are not pre-fabricated, but generated by the user on request. This does not apply to pictures and videos, but to diagrams, sound and animation, and representations generated by programs like Java or Flash. Within a certain set framework, the user can create new representations through newly entered data or variation of given parameters. An example is the statistical exercise on correlation displayed in figure 5. The user might also enter a text and have it processed by a speech synthesizer etc. In this way, a multimedia component can go through many representation forms and is not restricted to pre-fabricated objects. This level of interactivity allows the user to generate new representations, so that the multimedia components become visualizations of new relationships, assume heuristic functions for thought processes, and interact with the user's cognitive concepts.
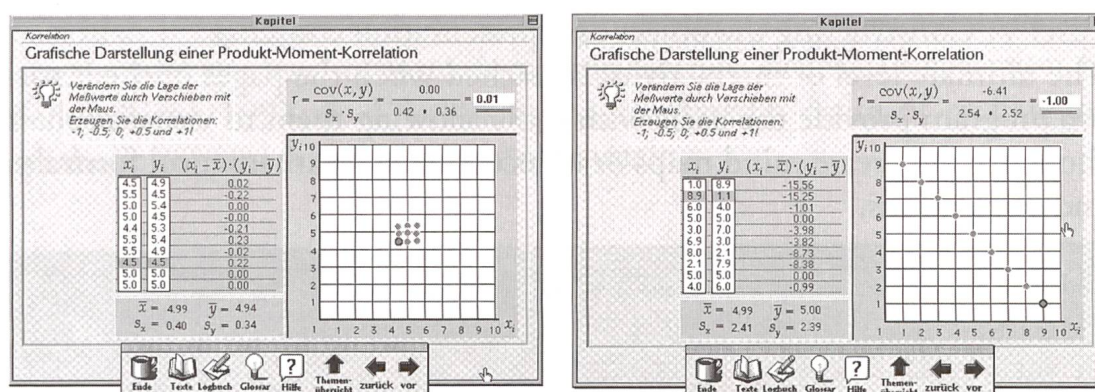
*Fig. 5: Two scatterplots created by dragging the dots in the plot (Schulmeis-ter/Jacobs 1992-1996)*

In this figure, you see two scatterplots created by students. The task of the student is the following: "Please drag the dots in the scatterplot with the mouse and try to construct a high/low positive/negative correlation." Why this unusual assignment?

Students of psychology, for example, when confronted for the first time with this type of exercise, very often tend to drag all dots into the middle of the scatterplot and assemble them in a close group, thinking that this illustrates high correlation (left picture in figure 5). They are surprised, however, to find out that this is not the case – indeed, their idea of high correlation results in zero correlation. The reason for this behaviour is that some students start with the naive cognitive assumption that correlation means something like nearness, neighbourhood or relation. But we know that this is not true. Correlation, on the contrary, depends on the co-variance of the data pairs and rows. It needs variance and a certain type of distribution or difference to form a correlation. Gradually, the students reach an understanding of the concept which results for instance in a diagonal row of dots (right picture in figure 5). The explorative space opened by such a highly interactive exercise gives students a chance to discover their own naive cognitive concepts, identify their misconceptions and gradually achieve a new understanding of the scientific concept of correlation.

As you can see, this exercise could only be created with a knowledge of the cognitive misconceptions of students, of the errors they make and the misunderstandings which arise in reading statistical textbooks. The exercises in LearnSTATS are specifically designed to address these problems. They don't tell the student what a statistical concept is, they rather invite the students to discover these concepts by themselves through applying their pre-formed cognitive patterns. Students are forced to realize their own cognitive concepts, are disturbed, and then try in a process of accommodation and assimilation (Jean Piaget) to gain back their equilibrium, to learn the statistical concept by modifying their cognitive concept.

This type of learning is called discovery learning (Jerome S. Bruner). Arranging learning situations or environments in such a way is totally different from an instructional type of teaching. Inventing exercises of this kind is not easy task. As I have said, one must have a knowledge of the naive assumptions of students and their cognitive mistakes in order to form an idea of the kind of exercise which might be useful.

At the same time exploration or discovery learning has a powerful potential for fostering the development of meta-learning capabilities.

## Level V: Constructing the object or representation contents

The ultimate level of interactivity is reached when the pages of the learning program provide the users with tools allowing them to visualize their thoughts and create mind maps or objects such as mathematical formulas and calculations.
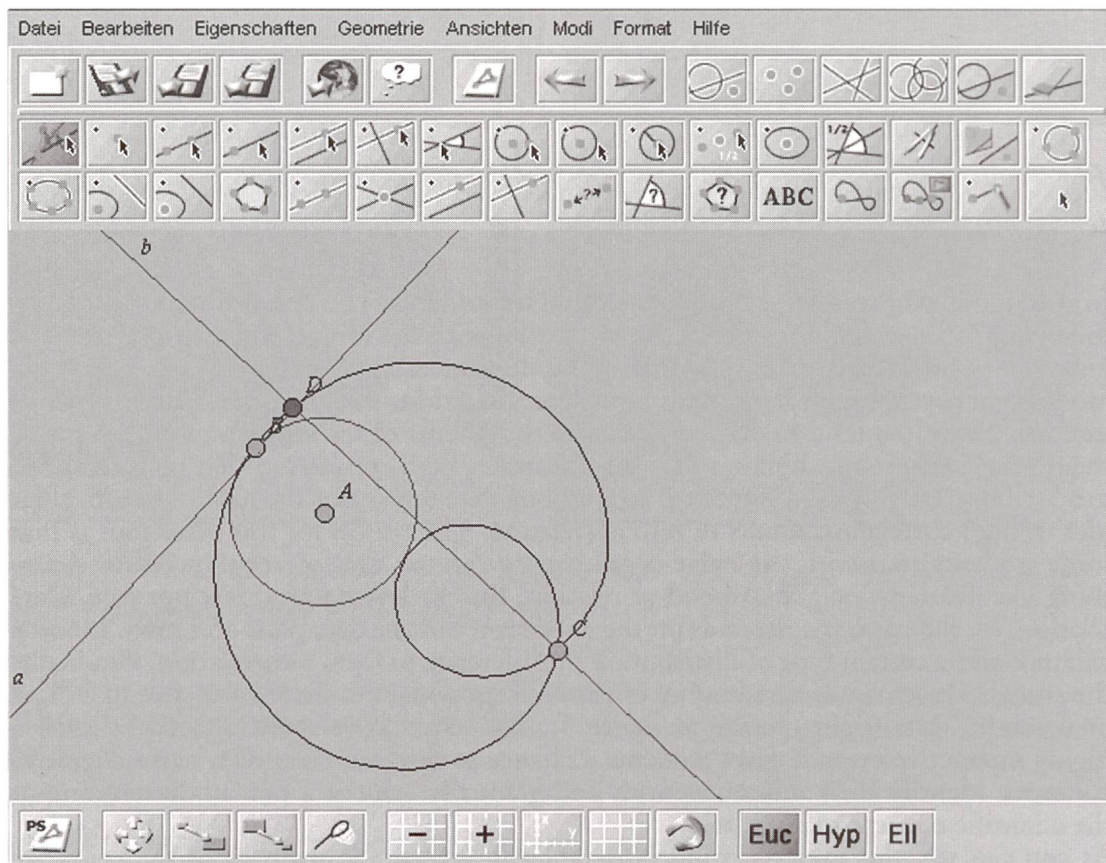


*Fig. 6: Geometry editor Cinderella with geometrical objects (Richter-Gebert/Kortenkamp 2001)*

## Level VI: Constructing the Object or Contents of the Representation and Receiving Intelligent Feedback from the System through Manipulative Action

In some disciplines, the development of systems with intelligent feedback is relatively far advanced, such as with mathematical editors and geometry programs – any context in which the symbolic contents of the semantic level can also be modelled as meaningful objects. That is case for most of natural science, but not for history, social science or arts. Interactivity on this level means that the "partner" computer or program is supplied with meaningful objects or actions which the program can interpret and

to which it can react with correlating meaningful actions. We have still not reached the level of human communication or social interaction (s. Schulmeister 1997, p. 40). But we have entered a realm of exchanging symbolic content within a restricted domain.



*Fig. 7: Interactive exercise with context sensitive feedback (Metzger, Schulmeister, Zienert 2002)*

This screenshot from "Die Firma 2. Deutsche Gebärdensprache interaktiv" [The Company 2. Interactive German Sign Language], a learning program for German Sign Language (in development), shows a grammar exercise for directional verbs in sign language. The tutor in the video gives instructions in sign language on how to place the furniture in the room on the right. Learners have two exercise windows at their disposal: The top picture contains a three-dimensional view of the room. The point of view can be changed by rotating the picture. The image below is a two-dimensional ground plan with the furniture items to be manipulated. Furniture can be moved and rotated. While the correct reception of the video cannot be controlled directly, this exercise is a practical way of checking whether the learner has understood the tutor's instructions correctly and to provide suitable and helpful feedback.

In order to promote the soundness of the taxonomy described, I am now going to test it against two other attempts at a scale, a scale of *visualization systems* and a scale of *program components*.

A Scale of Visualization Methods

The taxonomy of multimedia components interactivity shows some correspondence to the scale for degree of interactivity for visualization systems attempted by El Saddik (2001, p. 16):

1. Still images
2. Animated Pictures
3. Visualization with display adjustments for play, stop, speed etc.
4. Visualization selection and arrangement capabilites VCR for repeat, rewind etc.
5. Visualization with changing input, zooming and panning
6. Visualization with interactive decision points, e.g. changing data while running
7. Visualization generated by students (visualization construction kit)

It is my impression that levels 3 and 4 in El Saddik do not really designate a substantial difference. Both mean videos that can be manipulated by the user without changing the video itself. Levels 5 and 6 are also not clearly differentiated. Both mean videos that the user can manipulate interactively, by changing the display through zooming, rotating, etc., or by controlling representation and contents through manipulating other parameters. Moreover, the term *visualization* in El Saddik's concept is not quite clear. When talking of multimedia components in general, I don't think that there is a significant technical difference between a series of pictures (animated pictures in El Saddik) and video (visualization with display adjustments), since both constitute a series of individual images in time. Keeping this in mind, we arrive at the following scale for pictures and video:

Level I:     Viewing still pictures (El Saddik Level 1)
Level II:    Viewing video (including play, stop, speed, repeat, rewind etc.) (El Saddik Levels 2 + 3 + 4)
Level III:   Manipulating video display and viewing order (rotating, zooming, jumping to other parts of a video) (El Saddik Level 5)
Level IV:    Manipulating video or visualization contents through data input (El Saddik Level 6)
Level V:     Generating videos or visualizations through programs or data (El Saddik Level 7)
Level VI:    Receiving feedback on manipulations of visual objects (this level does not exist in El Saddik)

This shows clearly how closely El Saddik's concept corresponds to the taxonomy of interactivity level introduced here.

## A Scale of Program Components

Up to now, I have only used text/picture and text/video examples to illustrate interactivity levels. In the following, I will attempt to apply the same taxonomy to interactive programs in order to discuss the general soundness of this taxonomy for all kinds of multimedia components. For interactive programs, the taxonomy could look like this:

Level I: Automatic program execution
The user can start the program, which is then executed automatically. The user can only observe and watch the result, with no manipulation options.
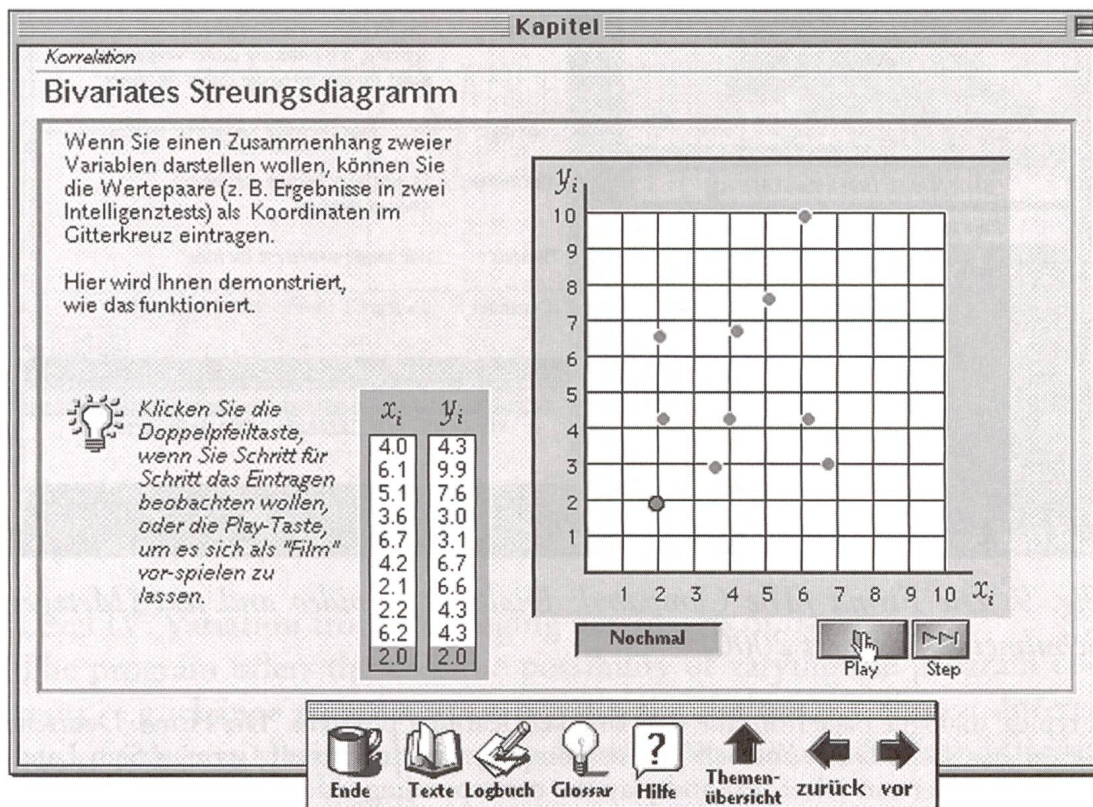


*Fig. 8: Automatic program execution: Correlation diagram (Schulmeister/Jacobs 1992-1996)*

A correlation diagram is created automatically from two lists of values for the variables x

and y. Users may repeat the process or watch it step by step, but they cannot influence the value lists or diagram type. Compare this example to the exercise in fig. 5, in which the user may move the dots in the correlation diagram to analyse the effect on the value pairs for the two variables.

## Level II: Multiple optional program executions
The program can be repeated by the user, with a choice of options, so that several variations are offered within the same exercise. But the user has no way of changing program execution or exercise contents.
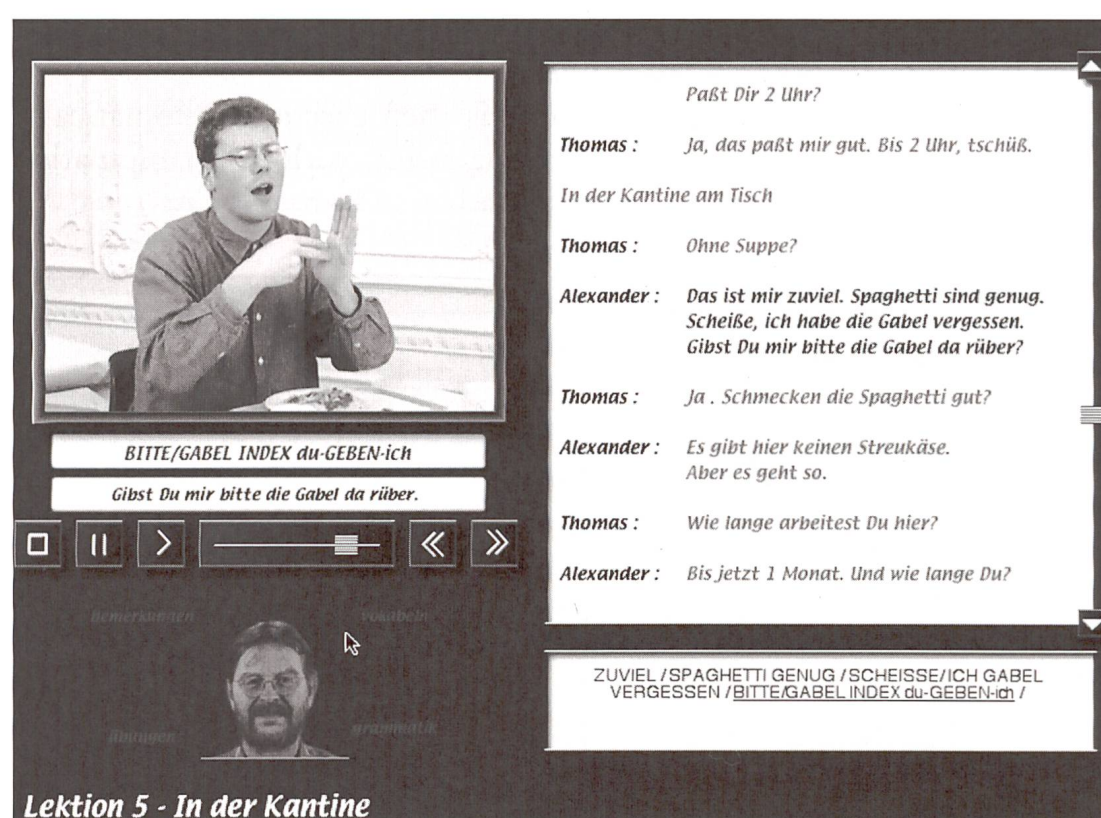


*Fig. 9: Die Firma [The Company]: Dialogue as video and text (Metzger, Schulmeister, Zienert 2000)*

A typical dialogue page from the sign language learning program "Die Firma. Deutsche Gebärdensprache Do It Yourself" [The Company. Do It Yourself German Sign Language]: The user can select text and play the corresponding video. The content remains the same, neither text nor video can be changed.

## Level III: Representation variation by program manipulation
The program offers the user the possibility of varying the program representation. There are several options to choose from, so that several varia-

tions are offered within the same exercise. But the user still has no way of changing program execution or exercise contents.
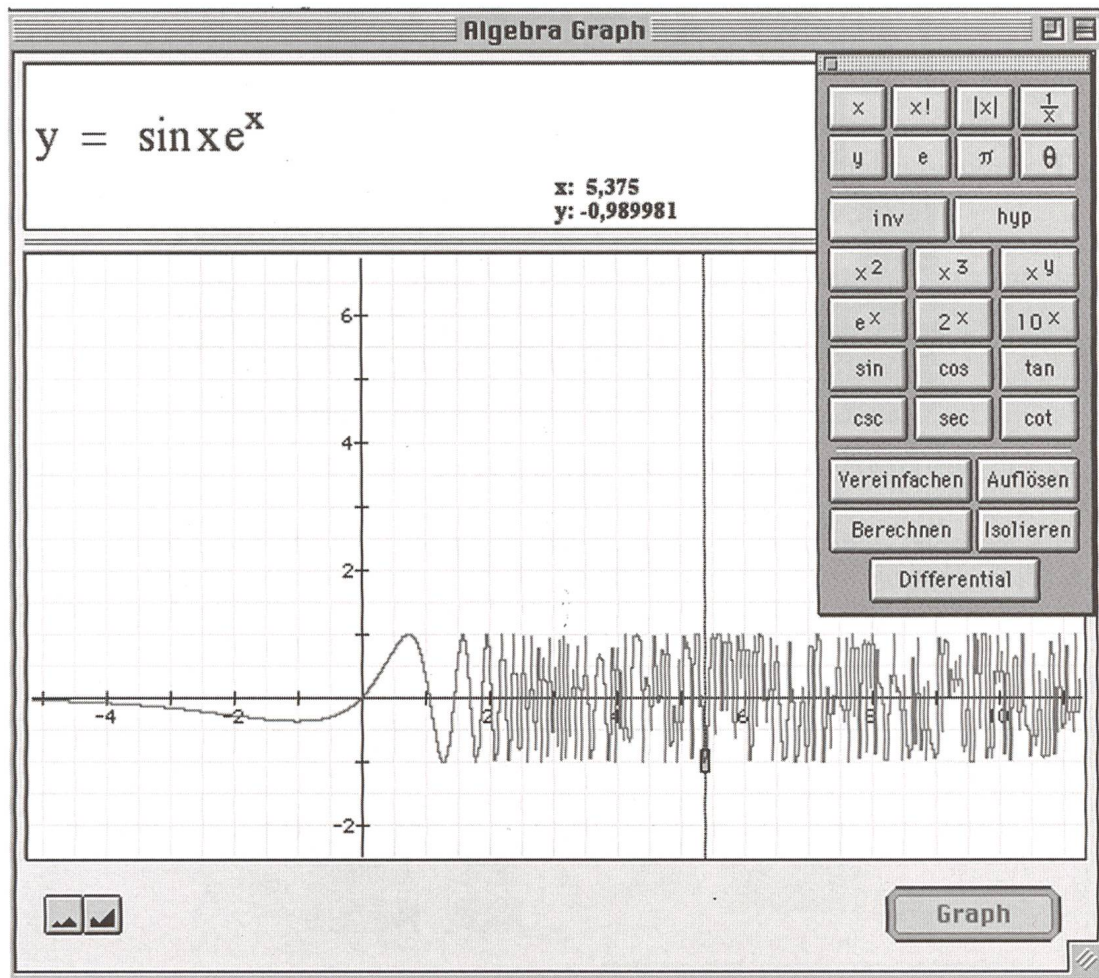


*Fig. 10: AlgebraGraph: Zooming, Rotating, Moving, Distorting (Avitzur 1994)*

Level IV: Variation trough changing parameters or data
The program offers the user the possibility of varying the program contents, e.g. change parameters in physics simulations, or choose different sets of values in statistics programs. In this way, the user can modify given objects and get new results.
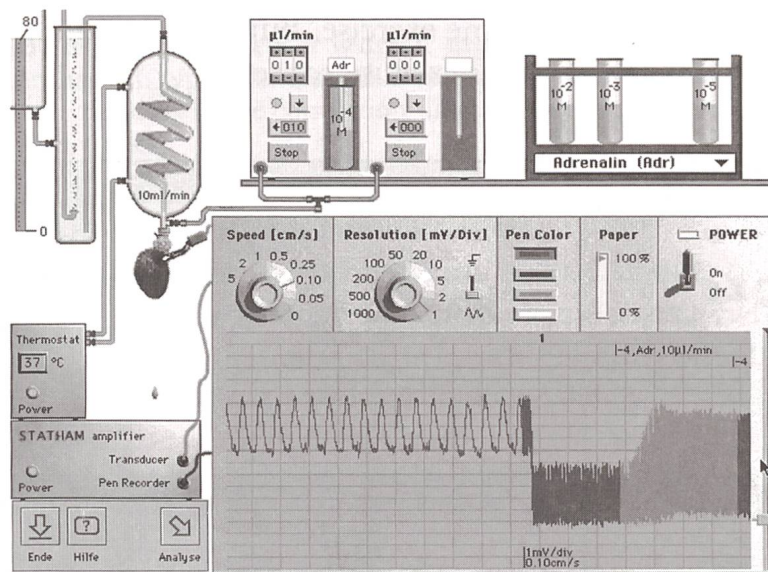
*Fig. 11: SimHeart — Laboratory (Hirsch et al. 1997)*

In this example, students of medicine are offered an artificial laboratory in which they can adjust controls and apply the instruments. As in the previous example, there is a choice of several options. But the student may change parameters for readings and calculations by manipulating the graphic objects (laboratory instruments).



*Fig. 12: Calculation of variance through data input in formulas (Schulmeister/Jacobs 1992-1996)*

In this example, the student can calculate variance by having the program calculate the raw data step by step and inserting the results into the formulas by way of Drag & Drop. The raw data can be changed.

## Level V: Constructing objects and generating processes

The program offers the user the possibility of generating the program contents, e.g. construct new environments in simulations, a biological or economic model, or initiate business processes in a virtual company. The program becomes a tool for the user to construct new worlds.
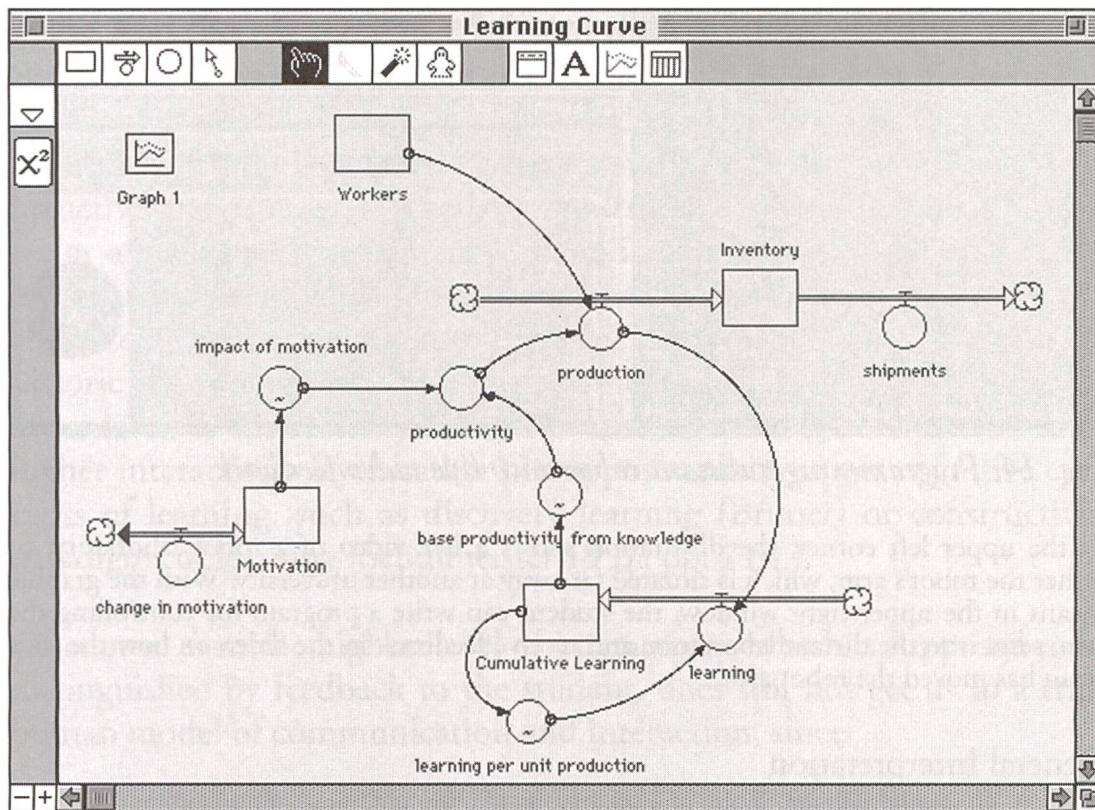


*Fig. 13: ithink editor: model editing (High Performance Systems)*

In this program, ithink, users can construct models of biological or physical processes and relationships with graphics. These models can then be set off like a program and provide time-related data in a number of iterative run-throughs. This example shows a model of the effects of the learning curve. The products are graphs and data.

## Level VI: Processes and programs with feedback

Programs which support the user in the construction of new worlds by providing intelligent feedback are still few and far between. First attempts are mostly found in the field of programming with programming languages.
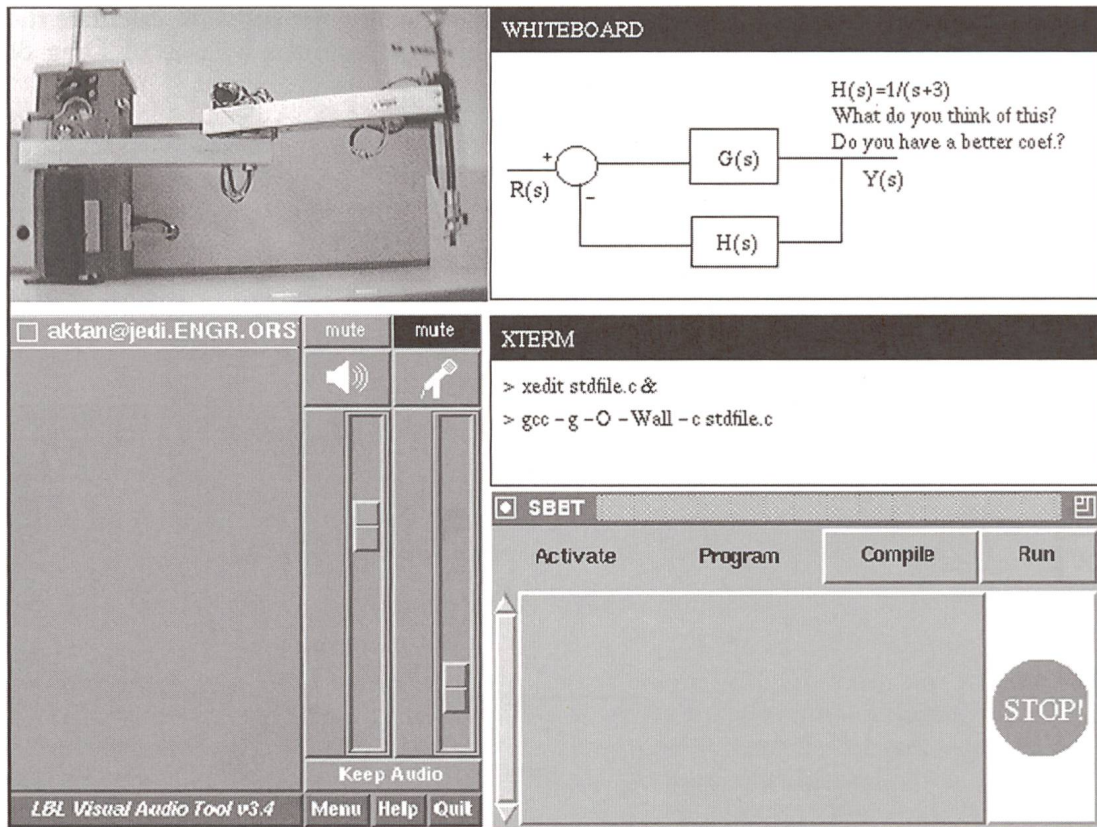
*Fig. 14: Programming a distant robot with live video feedback*

In the upper left corner, the illustration shows a live video of a robot laboratory, or rather the robot's arm, which is situated far away at another university. With the graphic means in the upper right window, the student can write a program for controlling the arm, send it to the distant laboratory, and receive feedback in the video on how the program has moved the robot arm.

## General Interpretation

In a more general sense, we must relate the user actions we call interactivity to the layers of multimedia space (s. Schulmeister 1996 and 1997). User actions, which on today's computers are still largely enactive interaction, connect event space with representation space and thus make the symbol space accessible.[1]

---

[1] The event space sends and receives all technical user actions, registers and controls all program processes, while the representation space supplies the interface to multimedia objects and contents, by way of windows, items, text display etc. The symbol space must be distinguished from both of these: this is a symbolic layer of the software or multimedia object containing the author's symbolic messages, the meaningful communications of the software.

A definition of multimedia thus consists of the dialogic, interactive component of the multimedia system, and of the interpretation and manipulation of multimedia objects by the learner. The learner triggers off events by manipulating multimedia objects. This presupposes a familiarity with the methods that multimedia objects trigger off in the representation level (ease of manipulation as an objective of multimedia designers). In doing this, the learner makes use of hypotheses about the methods triggered off by the objects in the deep structure. (Schulmeister 1997, p. 40)

Relating this to the interactivity levels, we arrive at the following conclusions: An increase of the interactivity level makes for a more varied event space, a more highly diversified representation space and a wider symbol space. And there is more to read from the levels of interactivity. Rhodes and Azbell (1985) distinguish three forms of interactivity design in learning environments: reactive, coactive and proactive design. Reactive design stems from the behaviourist stimulus-reaction paradigm, while proactive design assigns an actively constructive role to the user. One can see at a glance that the proactive share in interactivity rises with higher interactivity levels, while lower interactivity levels exhibit a more reactive character. The special charm of this scale is that it is compatible with the historic chronology of psychological theories of learning: The reactive lower levels of interactivity tend to assume a behaviourist character, while higher interactivity levels rather presuppose and transport cognitive concepts of learning, such as discovery learning (Bruner) or constructivist paradigms of learning (Schulmeister 1997, pp. 71ff.).

Of course, even the highest level of interactivity, which is supposed to be distinguished by feedback to the student, does still not get us to a truly human model of communication and interaction, since

"The reciprocity and symmetry of communication is what distinguishes real dialogue from the artificial dialogues of programs. I cannot agree with a program on the topic which has been predetermined by the author, I cannot cause the program to change its style of interaction and enter into a meta-communication. The reciprocity of communication is violated in human-program interaction." (Schulmeister 1997, p. 49)

The taxonomy of multimedia component interactivity in learning programs is a formal one. What is the interest in providing such a formal distinction of interactivity levels? We all know that many pedagogical hypotheses are linked to the concept of interactivity. As a developer of learning systems one might suppose, for example, that student motivation increases with higher interactivity levels. Such a hypothesis might be verified by way of an exercise offered in various levels of interaction. One

could also analyse in an experiment whether the alternative learning theory concepts made possible by higher levels of interactivity have an effect on learning efficiency and the quality of learning processes and learning results.

## References

APPLE COMPUTER. (1998). DeskLamp. Movie.

AVITZUR, R. (1994). Algebra Graph. Software. Frame Technology 1988-1991.

EL SADDIK, A. (2001) Interactive Multimedia Learning. Springer: Berlin u.a.

HIGH PERFORMANCE SYSTEMS. ithink. Software.

HIRSCH, M.CH.& BRAUN, H.A. (1997). Simulation des Herz-Versuchs. CD-ROM. Thieme.

METZGER, CH.; SCHULMEISTER, R. ET AL. (2000). Die Firma. Deutsche Gebärdensprache Do It Yourself. CD-ROM. Signum

METZGER, CH.; SCHULMEISTER, R.& ZIENERT, H. (2002). Die Firma 2. Deutsche Gebärdensprache interaktiv. CD-ROM. Signum

RHODES, D.M. & AZBELL, J.W. Designing Interactive Video Instruction Professionally. In: *Training and Development Journal* 39 (1985/12) 31-33

RICHTER-GEBERT, J.& KORTENKAMP, U.H. (2001). Die interaktive Geometrie-Software Cinderella. Buch und CD-ROM, Berlin: Springer.

SCHULMEISTER, R.; JACOBS, M., LernSTATS. (1992-1996). Software, Hamburg.

SCHULMEISTER, R. (1996). Grundlagen hypermedialer Lernsysteme. Theorie — Didaktik — Design. Addison-Wesley: Bonn, Paris u.a. ; 2. ed. Oldenbourg: München, Wien 1997, 3. ed. 2002.

STADLER, R. STATinside 1.0. (2001). Software. Luchterhand.