Visuals in Instruction: Functions of Screen Captures in Software Manuals[1]

Mark Gellevij, Hans van der Meij, Ton de Jong, Jules Pieters

Faculty of Educational Science and Technology
P.O. box 217
7500 AE Enschede, The Netherlands
gellevij@edte.utwente.nl

Abstract: This paper aims to answer two questions. The first is how screen captures in software manuals should be used to support learning. Four functions of screen captures in software manuals are presented. Three studies report the influence of these functions on the effectiveness and efficiency of the learning process. The second question is how users of visual manuals process combinations of text and pictures in memory. Two theories, Dual Coding theory and Cognitive Load theory seem to be in conflict here. Based on cognitive load measurements in two studies, Dual Coding theory initially gives the best explanation for text-picture processing.

Introduction

The use of computers is well integrated into our day-to-day lives for both professional as well as for 'fun' purposes. Despite this integration, working with a computer or computer based application at first almost never goes flawlessly. Although computer programmers and interface designers strive after user interfaces that are as intuitively as possible, some sort of instruction nearly always appears to be required to learn how to work with it. Producers of computer hardware, software, and computer applications therefore most often offer some sort of user support to their customers. A usual type of support for these products are software manuals.

Instructional material in which the use of various media is integrated improves learning more than instruction with only one medium (Mayer, 1999; Mayer & Sims, 1994; Mayer, Moreno, Boir & Vagge, 1999; Mayer & Gallini, 1990; Robinson, Robinson, & Katayama, 1999; Kulhavy, Stock, Peterson, Pridemore, & Klein, 1992). Especially Mayer and his co-authors have repeatedly shown that users benefit from a multimodal approach, whose most common form is that of a mixture of words and pictures. The most common type of visualization in software manuals is a screen capture (Van der Meij & Gellevij, 1998). Using software manuals in which screen captures are combined with text are therefore expected to be an improvement over purely textual manuals.

In this paper, the use of screen captures in software manuals is studied to answer two research questions. The first question is how screen captures should be used and designed to improve learning and reduce training time. The goal is to provide guidelines on how to improve software manuals by the use of screen captures. The second question is how a combination of text and pictures is processed in memory. Two theories on this topic are seemingly in conflict. Dual Coding theory predicts that learning is enhanced by such combinations while cognitive load theory predicts that learning will be hindered.

Screen captures to improve software manuals

The goal of a computer manual is to learn people work with a computer program in a short period of time. One reason to examine how screen captures could improve manuals was the observation that manuals seemed to become increasingly visual. Especially in commercially successful manuals, pictures were as much used as text. This, in spite of the fact that it was as yet unknown if these visualizations did or did not enhance learning.

[1] This paper is a summary of the Ph.D. thesis : Gellevij, M.R.M. (2002). *Visuals in Instruction: Screen Captures in Software Manuals.* Enschede: Twente University Press. Prof. Dr. de Jong and Prof. Dr. Pieters were the promotors of this thesis, Dr. van der Meij the assistant promotor.

To test this, an explorative study was carried out, in which the use and effects of two visual manuals (which designs were based on commercially successful manuals) and a textual (control) manual are compared (Gellevij, Van der Meij, de Jong, & Pieters, 1999). The main difference between the two visual manuals was the use of partial screen captures (see Figure 1) in Stuur's 'Windows for children' manual (Stuur, 1996), and of full screen captures (see Figure 2) in Gardner and Beatty's 'Guided Tour' manuals (Gardner & Beatty, 1994). Users of the manual with partial screen captures learned less than users of the textual and the visual manual with full screen captures. No differences were found between these last two manuals. From this, it was concluded that it does matter which type of screen captures (full or partial) are used, but that screen captures not necessarily improve a manual. After all, the textual and visual manual with full screen captures worked equally well.

The findings of this explorative experiment made us reconsider the way to deploy screen captures in manuals. The visual manuals used in the explorative study were designed according to what can be called a genre-approach. That is, the screen capture design chosen (in this case either partial or full screen captures) is used consistently throughout the manual without the purpose of having these screen capture support particular goals. Taking such a goal-based perspective was considered to provide for good opportunities to improve a manual. A detailed analysis of the instruction in a manual according to this goal-based approach lead to four specific functions for the use of screen captures in software manuals. Screen captures support:

- switching attention between the manual and the computer screen;
- the development of a mental model of the computer program;
- the identification and localization of screen objects, and
- the verification of screen states.

In three experiments, it was tested whether the use of screen captures to support one or more of these functions indeed improved the manual. These studies all compared a textual manual with a visual manual in which screen captures were used to support a particular function. Users were asked to work with the manual for a while. During these practice-sessions, their behavior was observed or recorded by log-registration, and time was measured that users needed to work through (parts of) the manual. Tests were administered after the practice-sessions.

Results from the experiment is which it was observed how often users switch their attention between the manual and the computer screen showed that users looked up very frequently (Gellevij & Van der Meij, 2002) (see Figure 3). In almost all the action steps in which a switch could be expected, users indeed showed this behavior. This meant that a manual could hardly be improved by this function.

Using screen captures to support the development of a mental model showed clear results (Gellevij, Van der Meij, de Jong, & Pieters, 2002). Users of a visual manual (see Figure 4 and Figure 5) developed a stronger mental model that users of a textual manual. To test this mental model users were asked to solve problems with the computer program.

The same held for the function identification and localization of screen objects results (Gellevij, Van der Meij, de Jong, & Pieters, 2002). Users of a visual manual (see Figure 4 and Figure 5) scored better on test that measured their knowledge about objects on the screen (identification) and were faster in finding objects on the screen (localization).

Users of the visual manual (see Figure 6) also benefited from screen captured that supported the function verification of screen states results (Gellevij, Van der Meij, de Jong, & Pieters, in preparation). Based on the information in the manual, users had to check (verify) if this information was the same as the information presented on the computer screen, and if necessary, to change it. Users of a visual manual not only made less errors than users of a textual manual, they also used less time for it.

Processing text and pictures in memory

Information processing theories describe and hypothesize about processes in the mind when storing and retrieving knowledge and skills. Such processes are described in so-called information processing models of which the common features are joined in the Modal model (Bruning, Shaw, & Ronning, 1995). This model divides human cognition in three memory-systems: sensory memory, short-term memory (also called working memory), and long-term memory. Sensory memory very briefly stores information that comes to us through our senses. That information is next processed in short-term memory, to ultimately be encoded and stored in long-term memory.

The bottleneck in this process seems to be short-term memory because of its limitations. It can hold information for rather short intervals (about a minute), and its storage capacity is limited to 7 plus or minus 2 items or elements. This stands in contrast to long-term memory, which can hold information perhaps for life, and has an enormous capacity (Miller, 1956; Gleitman, 1986).

Two specific theories on information processing focus on short-term memory when text and pictures are to be processed simultaneously. These theories are of particular interest since they seem to give contradictory predictions about the influence of instruction on learning when text and pictures are combined. The theories are: Dual Coding theory and Cognitive Load theory.

According to Dual Coding theory, the basic reason why better learning occurs when words and pictures are combined, lies in how the learner processes information in working memory. Dual Coding, or dual processing, presupposes that when both verbal and visual information is presented, these two types of information are connected in working memory, and this referential connectivity, in turn, contributes to the construction of a strong mental model (Paivio, 1990; Mayer, 1999)[2]. Support for these assumptions can be found in the so-called 'contiguity effect', which means that text and visuals that are presented in a coordinated way lead to better learning results than text and visuals that are presented separately (see e.g., Mayer & Sims, 1994). In later work Mayer and Moreno (1998) go a step further and propose that working memory consists of two distinct systems: a verbal and a non-verbal system. Verbal information only concerns narrative, and both visual and textual information are seen as belonging to visual information. According to Mayer and Moreno (1998) textual and visual information are initially both represented in the non-verbal system. The textual information is then translated in auditory form and further processed in the verbal system. The research presented in this book started from this assumption that text and pictures are being processed in two distinct but connected working memory systems.

Cognitive Load theory builds on the idea that there is only one working memory with limited capacity. When people face a task that is already quite difficult, 'additional' instructions may therefore be more of a burden than an aid. An example of a 'burden' is that learning to use a computer program requires the use of more information sources and the handling of more devices than just the instructional material. Users must not only process the manual, they must also attend to the keyboard, mouse, and computer screen. Cognitive Load theory indicates that this situation poses two potential risks: redundancy and split attention effects (Chandler & Sweller, 1991; Sweller, 1994; Sweller & Chandler, 1994).

Redundancy occurs when the manual presents screen captures that the user can also see on the computer monitor. The *redundancy hypothesis* predicts that offering the same information twice, as in the case of a depicted computer screen, requires the user to also process this information twice. This double processing takes up unnecessary memory space, space that could have been used for processing other information.

Split attention effects occur when people must attend to different sources of information simultaneously. The *split attention hypothesis* predicts that in such a situation learning is hampered when these sources must be integrated. According to Sweller and Chandler, split attention effects obstruct learning when a text fragment is needed to understand a picture and vice versa because the user then has to process two distinct information sources at the same time.

In conclusion, according to Cognitive Load theory, the use of text-picture combinations in a manual while working with the computer at the same time, is *not* a preferable instructional format. The screen captures are redundant because the same information is already presented on the computer screen. Therefore, they increase the load on working memory, and are thus likely to increase training time and decrease learning. In contrast, Dual Coding theory predicts that instruction with text-picture combinations exceeds text-only information. Thanks to the distinctly different nature of text and pictures and their separate processing of content similar information, the information is better integrated, resulting in a stronger mental model and, more generally, strengthened learning.

In the last two studies in which screen captures functions were tested, it was also examined whether cognitive load differed while using either a textual or visual manual (Gellevij, Van der Meij, de Jong, & Pieters, 2002, in preparation). Cognitive load was measured by asking users to score the difficulty of specific tasks on a continuing scale from 'very difficult' to 'very easy'. In one of these studies, also response time was measured by letting users

[2] We see no essential differences between the views of Paivio and Mayer that is functional for the reasearch carried out. Our interpretation of these sources is that Mayer's Dual Processing theory builds on Paivio's Dual Coding theory.

shut off a sound that was played now and then while they worked with the manual. A long response time was considered high cognitive load, short response time was seen as low cognitive load. From Cognitive Load theory, an increase in cognitive load and/or worse learning in the visual condition was expected because of redundancy. From Dual Coding theory, better learning was expected in the visual condition, with no explicit prediction for cognitive load.

All findings showed equal cognitive load for users of the textual and visual manual. This, in combination with the learning effects found support the explanation of text-picture processing by Dual Coding theory. Users of the visual manual learn more than users of the textual manual while their cognitive load remains the same.

Summary and Discussion

The first question was how screen captures should be used to improve a manual. Findings of the research presented show that a goal-based, functional approach in using screen captures leads to a more effective and efficient manual. By supporting specific functions with screen captures, users learn more in less time.

The second question was how a combination of text and pictures is being processed in memory. Based on the findings, the explanation by Dual Coding theory seems most valid. This conclusion should be treated with caution. According to Dual Coding theory, textual information is being processed in the verbal memory system, despite the fact that written text is actually non-verbal information. It assumes that the written text is being translated into 'spoken' format and than further processed in verbal memory. In case both text and pictures are actually processed in the same (non verbal) memory system, the explanation by Dual Coding theory does not hold.

References

Bruning, R.H., Shaw, G. J., & Ronning, R. R. (1995). *Cognitive Psychology and Instruction (2nd Edition)*. Engelwood Cliffs, New Jersey: Prentice Hall.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*, 293-332.

Gardner, D. C., and Beatty, G. J. (1994). *Visuele leermethode Windows 3.1 [Windows 3.1: The visual learning guide]*. Utrecht: A.W. Bruna Uitgevers B.V.

Garmonsway, G. N. (1970). *The Penguin All English Dictionary*. Paris: Bordas.

Gellevij, M.R.M., & Van der Meij, H. (2002). Screen captures to support switching attention. *IEEE Transactions on Professional Communication, 45*.

Gellevij, M., Van der Meij, H., De Jong, T, & Pieters, J. M.(in preparation). Verifying screen states.

Gellevij, M., Van der Meij, H., De Jong, T, & Pieters, J.M. (2002). Multimodal versus unimodal instruction in a complex learning context. *Journal of Experimental Education, 70,* 215-239.

Gellevij, M., Van der Meij, H., De Jong, T, & Pieters, J. (1999). The effects of screen captures in manuals: Textual and visual manuals compared. *IEEE Transactions on Professional Communication, 42,* 77-91.

Gleitman, H. (1986). *Psychology (2nd edition)*. New York: Norton.

Kulhavy, R. W., Stock, W. A., Peterson, S. E., Pridemore, D. R., & Klein, J. D. (1992). Using maps to retrieve text: a test of conjoint retention. *Contemporary Educational Psychology, 17*, 56-70.

Mayer, R. E. (1999). Designing instruction for constructivist learning. In: Reigeluth, C.M., *Instructional-Design theories and models Volume II: A new paradigm of* instructional theory. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*, 715-726.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for Dual processing systems in working memory. *Journal of Educational Psychology, 90*, 213-330.

Mayer, R. E., Moreno, R., Boir, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology, 91*, 638-643.

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.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Nowaczyk, R. H., & James, E. C. (1993). Applying minimal manual principles for documentation of graphical user interfaces. *Journal of Technical Writing and Communication, 23*, 379-388.

Paivio, A. (1990). *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

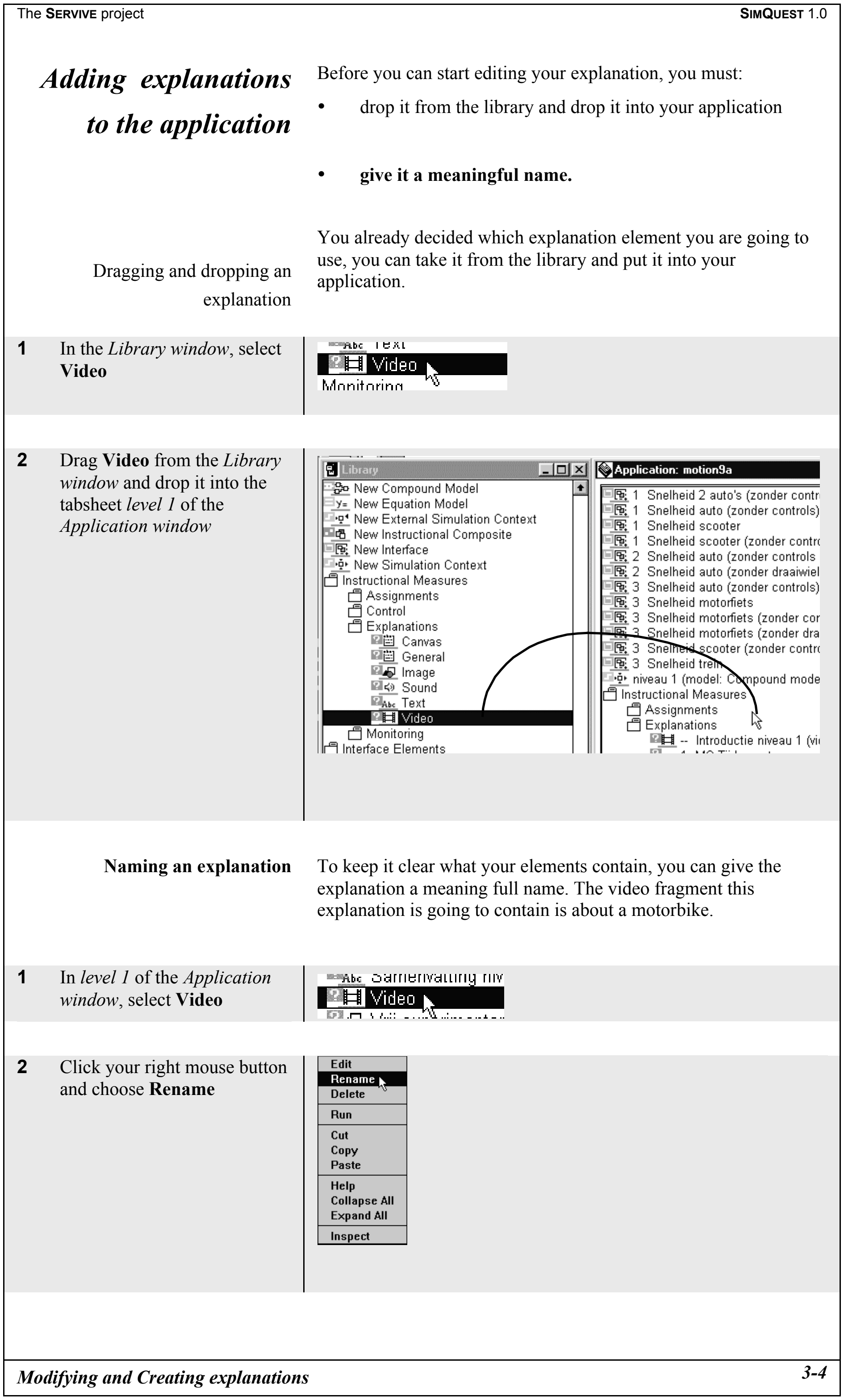
Robinson, D. H., Robinson, S. L., & Katayama, A. D. (1999). When words are represented in memory like pictures: evidence for spatial encoding of study materials. *Contemporary Educational Psychology, 24*, 38-54.

Stuur, A. (1996). Windows voor kinderen: deel2 [Windows for children: part 2]. Utrecht: A.W. Bruna Informatica.

Sweller., J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295-312.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185-233.

Van der Meij, H & Gellevij, M. (1998). Screen captures in software documentation. *Technical Communication, 45,* 529-543.



The **SERVIVE** project **SIMQUEST** 1.0

Adding explanations to the application

Before you can start editing your explanation, you must:

- drop it from the library and drop it into your application
- **give it a meaningful name.**

Dragging and dropping an explanation

You already decided which explanation element you are going to use, you can take it from the library and put it into your application.

1 In the *Library window*, select **Video**

2 Drag **Video** from the *Library window* and drop it into the tabsheet *level 1* of the *Application window*

Library
New Compound Model
New Equation Model
New External Simulation Context
New Instructional Composite
New Interface
New Simulation Context
Instructional Measures
Assignments
Control
Explanations
Canvas
General
Image
Sound
Text
Video
Monitoring
Interface Elements
Application: motion9a
1 Snelheid 2 auto's (zonder contr
1 Snelheid auto (zonder controls)
1 Snelheid scooter
1 Snelheid scooter (zonder contr
2 Snelheid auto (zonder controls
2 Snelheid auto (zonder draaiwiel
3 Snelheid auto (zonder controls)
3 Snelheid motorfiets
3 Snelheid motorfiets (zonder cor
3 Snelheid motorfiets (zonder dra
3 Snelheid scooter (zonder contr
3 Snelheid trein
niveau 1 (model: Compound mode
Instructional Measures
Assignments
Explanations
-- Introductie niveau 1 (vi

Naming an explanation

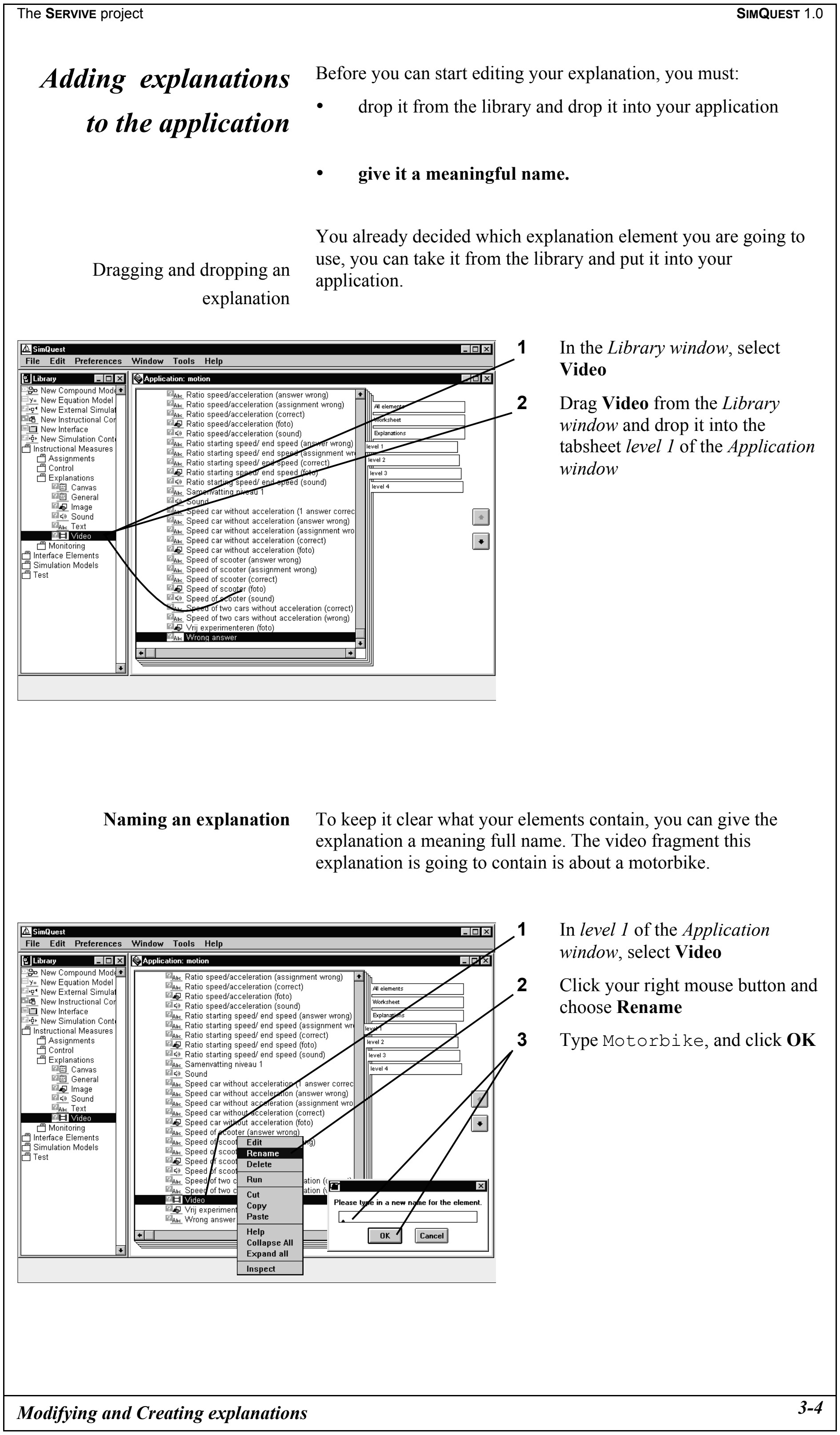
To keep it clear what your elements contain, you can give the explanation a meaning full name. The video fragment this explanation is going to contain is about a motorbike.

1 In *level 1* of the *Application window*, select **Video**

2 Click your right mouse button and choose **Rename**

Modifying and Creating explanations ***3-4***

Figure 1 Example page of a visual manual with partial screen captures



The SERVIVE project SIMQUEST 1.0

Adding explanations to the application

Before you can start editing your explanation, you must:

- drop it from the library and drop it into your application
- **give it a meaningful name.**

Dragging and dropping an explanation

You already decided which explanation element you are going to use, you can take it from the library and put it into your application.

SimQuest
File Edit Preferences Window Tools Help
Library
New Compound Model
New Equation Model
New External Simulat
New Instructional Cor
New Interface
New Simulation Cont
Instructional Measures
Assignments
Control
Explanations
Canvas
General
Image
Sound
Text
Video
Monitoring
Interface Elements
Simulation Models
Test
Application: motion
Ratio speed/acceleration (answer wrong)
Ratio speed/acceleration (assignment wrong)
Ratio speed/acceleration (correct)
Ratio speed/acceleration (foto)
Ratio speed/acceleration (sound)
Ratio starting speed/ end speed (answer wrong)
Ratio starting speed/ end speed (assignment wrong)
Ratio starting speed/ end speed (correct)
Ratio starting speed/ end speed (foto)
Ratio starting speed/ end speed (sound)
Samenvatting niveau 1
Sound
Speed car without acceleration (1 answer correc
Speed car without acceleration (answer wrong)
Speed car without acceleration (assignment wro
Speed car without acceleration (correct)
Speed car without acceleration (foto)
Speed of scooter (answer wrong)
Speed of scooter (assignment wrong)
Speed of scooter (correct)
Speed of scooter (foto)
Speed of scooter (sound)
Speed of two cars without acceleration (correct)
Speed of two cars without acceleration (wrong)
Vrij experimenteren (foto)
Wrong answer
All elements
Worksheet
Explanations
level 1
level 2
level 3
level 4

1 In the *Library window*, select **Video**

2 Drag **Video** from the *Library window* and drop it into the tabsheet *level 1* of the *Application window*

Naming an explanation

To keep it clear what your elements contain, you can give the explanation a meaning full name. The video fragment this explanation is going to contain is about a motorbike.

Please type in a new name for the element.

1 In *level 1* of the *Application window*, select **Video**

2 Click your right mouse button and choose **Rename**

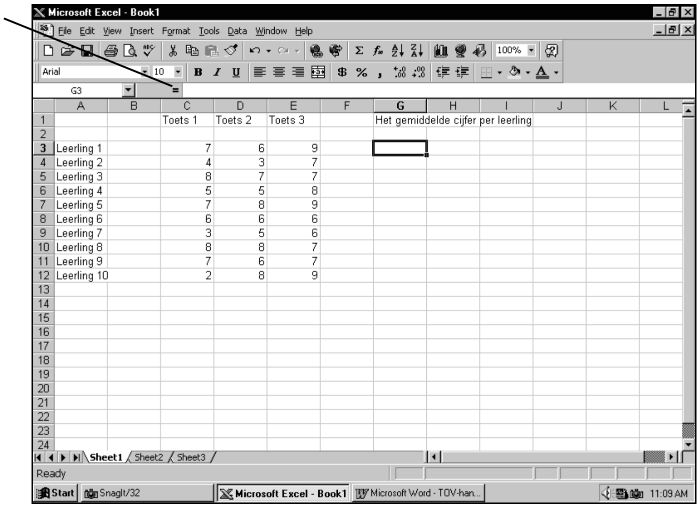
3 Type `Motorbike`, and click **OK**

Modifying and Creating explanations ***3-4***

Figure 2 Example page of a visual manual with full screen captures

Visual manual

1. Click on =



	A	B	C	D	E	F	G
1			Toets 1	Toets 2	Toets 3		Het gemiddelde cijfer per leerling
2							
3	Leerling 1		7	6	9		
4	Leerling 2		4	3	7		
5	Leerling 3		8	7	7		
6	Leerling 4		5	5	8		
7	Leerling 5		7	8	9		
8	Leerling 6		6	6	6		
9	Leerling 7		3	5	6		
10	Leerling 8		8	8	7		
11	Leerling 9		7	6	7		
12	Leerling 10		2	8	9		

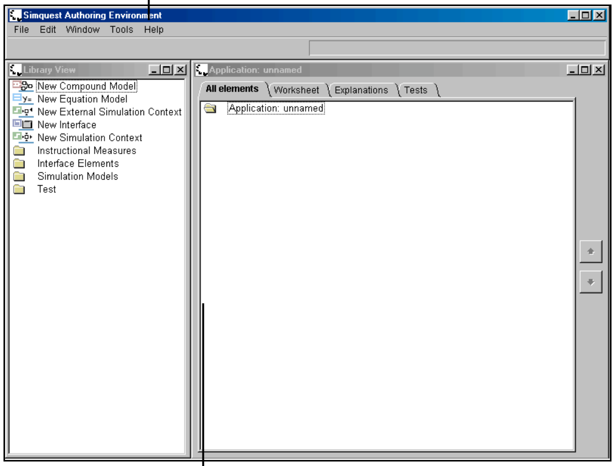
Textual manual

1. Click on the **= sign** in the formula-bar

Figure 3 Screen capture to support *switching attention between manual and screen*, and the textual counterpart

Visual manual

***SimQuest* menu and status**
With this window you can operate the general functions of SimQuest, like saving an application or exiting SimQuest. This window also shows the status of operations SimQuest is performing, for example the loading of an application.



Library
The library window contains standard elements you can use to create a new application.

Application
The application window shows the elements that you have put into your appl application window also shows the structure of your application.

Textual manual

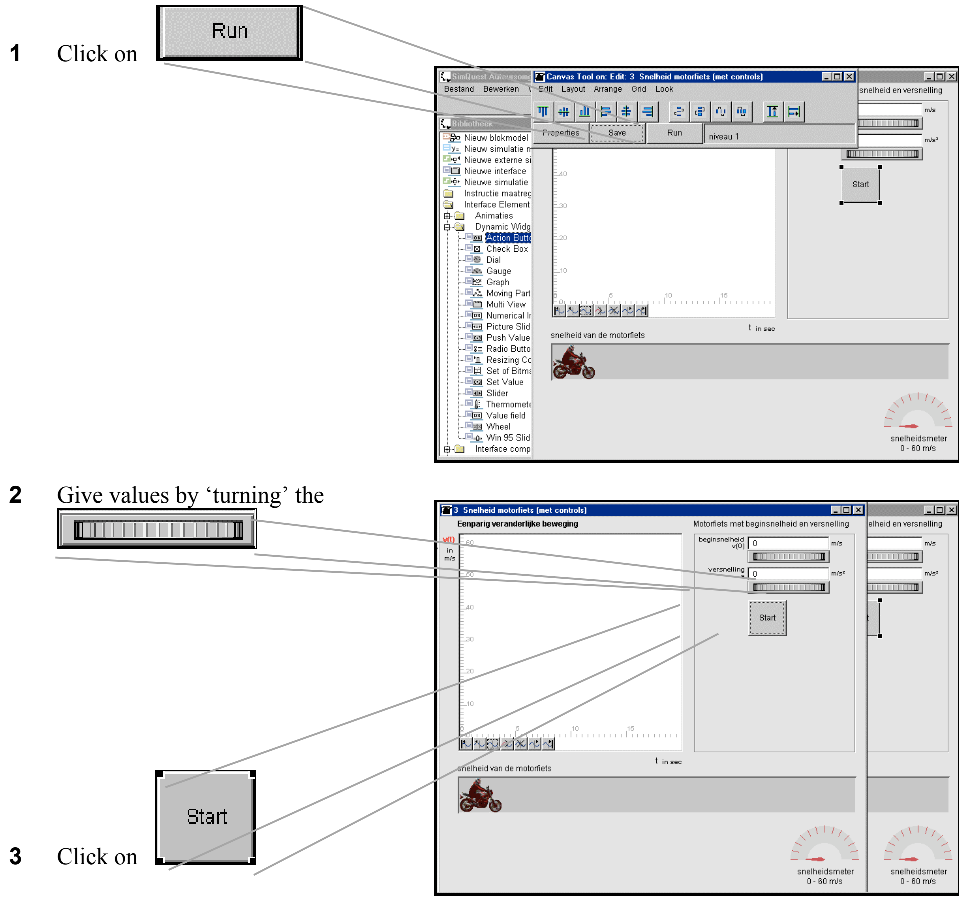
***SimQuest* menu and status (window on top)**
With this window you can operate the general functions of SimQuest, like saving an application or exiting SimQuest. This window also shov of operations SimQuest is performing, for example the loading of an application.

Library (window bottom left)
The library window contains standard elements you can use to create a new application.

Application (window bottom right)
The application window shows the elements that you have put into your application. The application window also shows the structure of your

Figure 4 Screen capture in **explanatory sections** for the functions *developing a mental model of the program* and *identifying and locating window elements and objects*, and the textual counterpart

Visual manual



1 Click on Run

2 Give values by ‘turning’ the

3 Click on Start

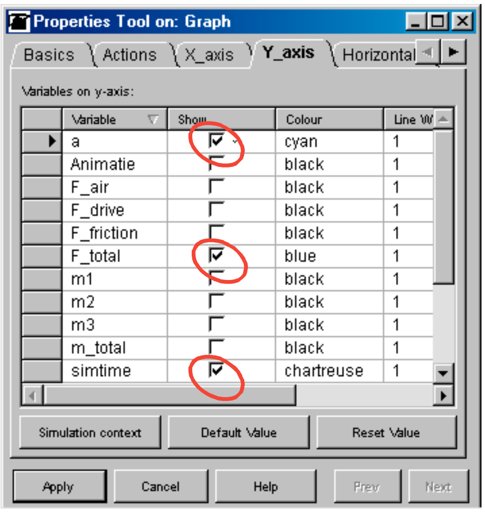
Textual manual

1 Click on **Run** in the window *Canvas Tool*
2 Give values by ‘turning’ the **Wheel**
3 Click on **Start**

Figure 5 Screen capture in **procedural sections** for the function *developing a mental model of the program* and *identifying and locating window elements and objects*, and the textual counterpart

Visual manual

1 Check, in the *Properties Tool on: Graph*, if the boxes in the *show* column of the variables **a** , **F_total**, and **simtime** are marked



Variables on y-axis:

Variable	Show	Colour	Line W
a	☑	cyan	1
Animatie	□	black	1
F_air	□	black	1
F_drive	□	black	1
F_friction	□	black	1
F_total	☑	blue	1
m1	□	black	1
m2	□	black	1
m3	□	black	1
m_total	□	black	1
simtime	☑	chartreuse	1

If not, mark the appropriate boxes by clicking in them

Textual manual

1 Check, in the *Properties Tool on: Graph*, if the boxes in the *show* column of the variables **a** , **F_total**, and **simtime** are marked

If not, mark the appropriate boxes by clicking in them

Figure 6 Screen capture for the function *verifying screen states*, and the textual counterpart