EPA2000: A multilingual, programmable computer assessment of off-line metacognition in children with mathematical-learning disabilities

ARMAND DE CLERCQ, ANNEMIE DESOETE, and HERBERT ROEYERS
University of Ghent, Ghent, Belgium

EPA2000 is a program for the assessment of off-line measured metacognitive skills and arithmetical performances in primary school children with mathematical-learning disabilities. The program is designed as a script engine. The concept makes it possible to modify and translate the test into different languages without reprogramming. A user-friendly script editor is built-in, with which all of the parameters of the test can be modified and translated in different languages.

Nearly 10% of primary school children have problems with mathematics, whereas 4% percent of them have mathematical-learning disabilities (LDs; see, e.g., Shalev, Manor, Auerbach, & Gross-Tsur, 1998). On the basis of the DSM–IV (American Psychiatric Association, 1994, pp. 46–50), *mathematical disabilities* can be defined as problems with mathematical calculation or reasoning, where the mathematical ability falls substantially below that expected for the individual's chronological age, measured intelligence, and age-appropriate education. The disturbance significantly interferes with academic achievement or with those activities of daily living that require mathematical skills. The difficulties in mathematical ability exceed those usually associated with any sensory deficit that may be present.

Children with mathematical LDs show some typical shortcomings in different phases and aspects of arithmetical problem solving (Desoete, Roeyers, & De Clercq, 2000). Some children make *input mistakes* of a perceptual or a phonetic type (McCloskey & Macaruso, 1995). These children confuse 6 with 9, 4 with 7, and 5 with 2, while having perceptual problems. They can also have problems with operator symbols (such as x, +, <, and >). Furthermore, in Dutch, *drie/vier* (three/four) and *zeven/negen* (seven/nine) often are mistaken, if phonetically confused. Other children with LDs especially make *elaboration* or *calculation mistakes* of a procedural type (McCloskey & Macaruso, 1995; Verschaffel, 1999). They answer 24 + 3 = 54 and 72 − 15 = 63. These children forget, in a multidigit addition problem, to start in the right column, to compute the sum of the digits in the rightmost column, to write the ones digit of the sum at the bottom of the column, and to carry the tens digit if any, and so forth. *Language-dependent* or *output mistakes* are also typical for some children with mathematical LDs (Campbell, 1998; Geary, 1993; McCloskey & Macaruso, 1995). These children have difficulties with all word problems and make number-writing confusions, since the spoken and written form is linguistically different. They write 710 or 71 for 17 (seventeen, or *zeventien* in Dutch), 370 for 73 (three-and-seventy, or *drieenzeventig* in Dutch). Children with mathematical LDs can also make *mental-representation-related mistakes* (Geary, 1993; Vermeer, 1997; Verschaffel, 1999). They answer 51 on the assignment *50 is 1 more than . . .* and 48 on the assignment *50 is 2 less than . . . .* Furthermore, some children have *arithmetic working memory* or *long-term memory* problems (Ackerman, Anhalt, & Dykman, 1986; Ashcraft, Fierman, & Bartolotta, 1984; McCloskey & Macaruso, 1995) and have trouble in retrieving arithmetical facts (multiplication tables, basic facts to 20). These children have problems remembering nonlogical (11 = eleven rather than oneteen, 12 = twelve rather than twoteen) or arbitrary facts. Furthermore, several studies have shown that many pupils with LDs lack valuable *metacognitive* skills used during the solution of arithmetical problems. All these items were included as subscales in the subsequently described EPA2000.

The *relationship between arithmetics and metacognition* was made clear by such authors as Carr and Jessup (1995), Lucangeli and Cornoldi (1997), and Verschaffel (1999). Metacognition, or the awareness of one's own cognitive functioning and the active monitoring of one's own cognitive processes (Flavell, 1976; Verschaffel, 1999), was recently studied in relation to mathematics (e.g., Borkowski, 1992; Montague, 1998). Prospective metacognitive *prediction*, or orientation, precedes control and guarantees that one will work slowly when exercises are new or complex and fast when tasks are easy or familiar, without overestimating one's performance (Brown, 1987; Masui & De Corte, 1999; Metcalfe, 1998; Vermunt, 1996). *Eval-*

uation refers to reflective verbalizations in retrospect, after the event has taken place, where children look at what strategies were used and whether they led to a desired result or not (Brown, 1987; Masui & De Corte, 1999; Vermunt, 1996).

Dynamic assessment refers, according to Lidz (1997), to the development of decision-specific information that most characteristically involves interaction between the examiner and the examinee, focusing on the learner's metacognitive processes and his or her responsiveness to intervention in a pretest–intervention–posttest administration format. Since some abilities seem trainable (Efklides, 1999), a pretest–posttest design, with the possibility for an short intervention (*kurzzeit lerntest*) study between both tests, seems most useful in the assessment of children with mathematical LDs (see also Rutland, 1995). No test is currently available for a dynamic assessment of metacognitive skills in the lower classes of elementary school.

This paper describes a program package, running under Microsoft Windows, for implementing a computerized instrument for metacognitive assessment. The target users are young children with LDs, although the test can also be used on children who perform moderately well, to measure their arithmetic and metacognitive skills. The contents of the test can be easily modified and translated by a psychologist without computer knowledge.

EPA2000

Development

The EPA was developed on the basis of a series of studies using a *semistructured metacognitive interview* with children in the second and third grades, inspired by the work of Lucangeli and Cornoldi (1997) and Tobias and Everson (1996). Analyses of the results of these studies (Desoete, Roeyers, & Buysse, 2000) showed the importance of metacognition in the arithmetic performances of young children. However, it became clear that open questions on metacognition were not so easily answered by young children. Furthermore, prediction and evaluation skills (*off-line measured metacognition*) seemed to be especially capable of differentiating good performers from moderate performers and children with LDs in this age group. Since both metacognitive parameters (prediction and evaluation) were measured before or after the solving of exercises, we labeled this metacognitive component *off-line metacognition*, in contrast with *on-line (measured) metacognitive skills*. Furthermore, since open questions were difficult for young children, different types of rating scales (a rating scale with numbers, faces to measure the degree of confidence [Vermeer, 1997], a color rating scale with traffic lights) were tested on children. The color rating scale appeared to be the most useful one for this age group with mathematical LDs.

The first version of the *EPA paper-and-pencil version* on off-line metacognition, with a color rating scale, consisted of 230 items. In the outcome-related assessment of metacognitive *prediction*, children were asked to look at mathematical exercises without solving them and to predict whether they would be successful in this task on a 4-point color rating scale. Children might predict well and do the exercise wrong, or vice versa. After completing the exercise, children *evaluated* how successful they had been, using the same 4-point color rating scale. In an analysis of data on 584 children (Desoete, Roeyers, & De Clercq, 2000), problems found to be unsuitable were replaced by other problems, and the number of items was reduced to 80. Children appeared to be able to handle the revised instrument very well. The psychometric characteristics of the EPA were analyzed. Different experts on arithmetics and on mathematical LDs were consulted, in order to increase the construct validity. As to the concurrent validity, Pearson product–moment correlation coefficients were computed between the arithmetic scores on the EPA and the scores on other arithmetic tests for these children ($n = 145$). A correlation of .56 ($p < .0005$) was found with the three other arithmetic tests frequently used in Belgium: the Kortijkse Rekentest (KRT; Cracco et al., 1995), the Vraagstukken (VT; Dudal, 1985), and the GRIPA (Catteeuw & Gheskiere, 1987). In addition, a correlation of .79 ($p < .0005$) was found between the EPA arithmetic scores and teacher ratings of arithmetic skills. Furthermore, Cronbach's alpha reliability analyses were conducted. Reliability coefficients of .88 were found. As to metacognition, different authors were consulted, to increase the construct validity. Furthermore, Cronbach's alphas of .79 and .73, respectively, were found for the prediction and evaluation scores of the EPA in the same sample.

A major disadvantage of the paper-and-pencil version, however, appeared to be the interference of arithmetic-solving processes in metacognitive skills and vice versa. It was observed that some children spent too much time on arithmetic problem solving, rather than on prediction or evaluation. Therefore, a *computer version (EPA2000)* was made with the same items of the EPA, which allowed us to measure the time needed for predictions and evaluations and made it possible to hide the given answer during the evaluation assessment. The EPA2000 was tried out on 30 children, and we have analyzed, up to now, data on 104 children. For this latter group, Cronbach's alpha was .91 for the arithmetic scores, .60 for the metacognitive prediction skills, and .77 for metacognitive evaluation skills. In a recent study with 80 children, we found the EPA2000 [$F(2,83) = 55.32, p < .0005$] to be able to differentiate third graders with and without mathematical LDs on prediction [$F(1,84) = 106.02, p < .0005$] and evaluation [$F(1,84) = 71.61, p < .0005$; Desoete, Roeyers, & De Clercq, 2000]. Furthermore, test–retest correlations of .81 ($p < .0005$) for the paper-and-pencil test were found in 30 children. In another study with 30 subjects, test–retest correlations of .80 ($p < .0005$) were found between the EPA and the EPA2000.

EPA2000 Computer Concept

In order to create a stable system with a simple user interface, the program is built as a script engine, made in Visual Basic 6, that executes a script file, written in plain ASCII text (see Figure 1). Every exercise is described in the program script as a number of elementary commands. The total command set consists of 25 commands. The whole script consists of about 2,400 lines. This is almost impossible for a nontrained programmer to modify. Therefore, a visual editor was created. This editor bundles all the elementary script commands for one exercise into one form on the screen. The concept of defining an application language as a well-defined set of scripting commands is frequently used. For example, it is the basis for HTML Web language, e-mail, and many others. As the applications become too complex, a user friendly front end is necessary. For the HTML language, there are Front-Page from Microsoft (Buyens, 1997), Dreamweaver (Towers, 1999) from Micromedia, and others. The method is also used in E-Prime (Schneider, 1998), where E-Studio is a comprehensive, graphic, experimental design environment that generates code in E-Basic, a full-featured object-oriented Basic compiler with many enhanced commands for behavioral research.

The EPA CLIENT User Interface

The introduction of the assessment to the children is designed as a cartoon. We used the cartoon figures, Willy and Wanda, of a well-known Belgian cartoonist, W. Vandersteen, whose cartoons have been translated into 25 languages. The cartoon text can be modified. For every cartoon picture, the text can be spoken through a sound card, by adding WAV files. This is especially useful for children with reading problems. Figure 2 gives a picture of the user interface that is used throughout the entire assessment. The upper side of the screen displays the arithmetical task and the possible answers. The response buttons are automatically adjusted in size, and the number of buttons is defined at design time in the script. The lower part contains the 4-point color rating scale, represented by four traffic lights. There are two parts to the test. In the first part, the question, the possible answers, and the 4-point rating scale presented by the traffic lights are displayed at the same time. In this part, the child can only answer if he/she thinks that he/she can or cannot answer the question in that exercise. After all the items are processed, every question is displayed without the rating scale. The child is then supposed to solve the exercise. After that, the exercise disappears, and the 4-point color rating scale (the traffic lights) appears again. After the test, it is possible to display a posttest. At the end of the test, the results are displayed and can be printed. A hidden control sequence is added to the test. It is activated by pressing a control–shift sequence. The sequence is unlikely to be hit accidentally by the child, but it allows the experimenter to interrupt in order to abort the test, choose a different testing language, and so forth.
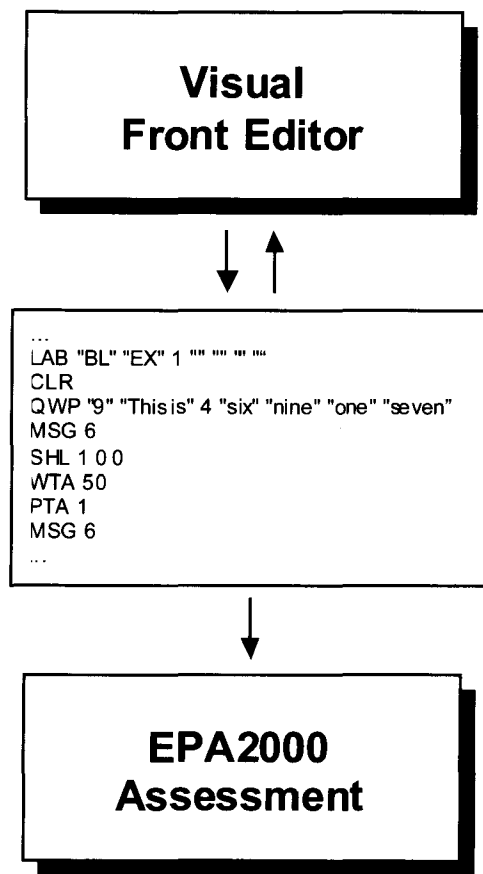
The Script Engine

Each script step in EPA2000 is kept as general as possible. One command describes the text to be displayed in the exercises on the screen, another command controls the buttons that are to be pressed from the question and the 4-point color scale, and there is a command for displaying cartoon pictures and the associated text balloon in the introduction. Below, a brief description of the commands is given. A detailed description is given in the source of the program.

Every line has the following structure:

COM arg1 arg2 ...

COM is always a three-letter word and is the command. Arg1 arg2 ... are additional arguments. Arguments can be text between quotes or numbers. All of the parts, such as the header of the script file, the different exercises, and the commands describing the processing of the data, start with the LAB command. The first argument of LAB indicates the type of block; the second is a comment. The next arguments are block specific and describe extra prop-
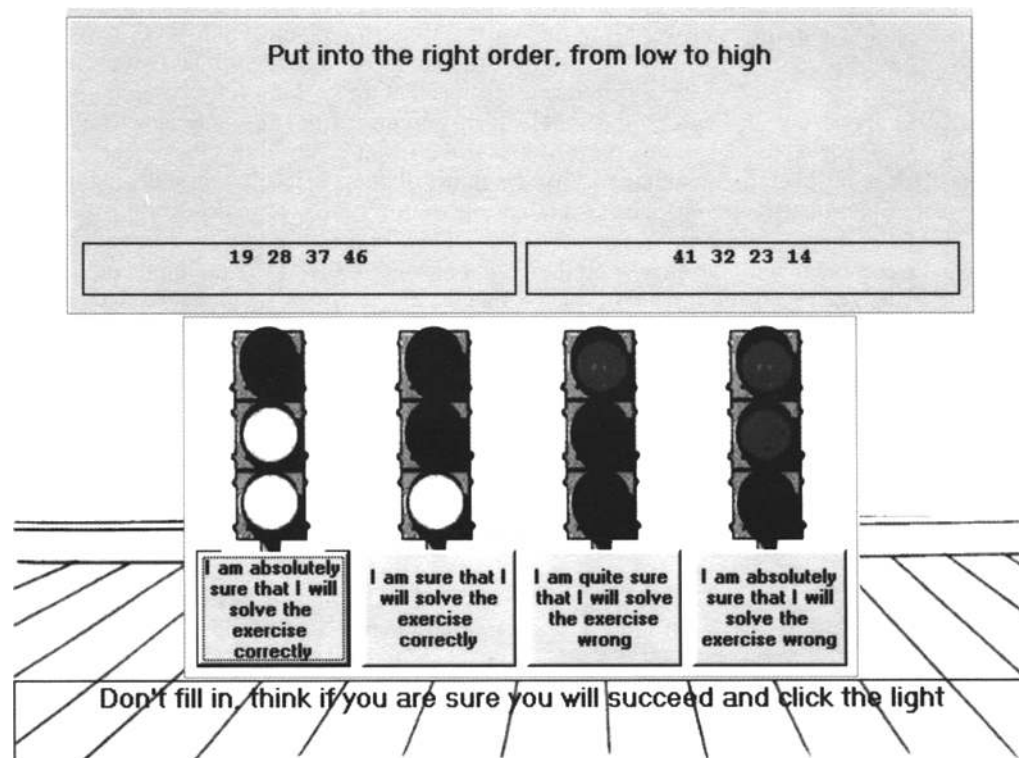

Figure 1. The EPA2000 construction.

Figure 2. The user interface used throughout the assessment.

erties of the block, such as the button with the good answer for an exercise and the subscale to which it belongs. We will now describe three relevant command sequences: the header of the script, a sequence for describing an exercise, and a sequence for representing the calculated results.

LAB "ST" "General program info" 0 "" "" "" ""
MSF 1 "name"
. . .
MSW 1 "Click an answer with the mouse"
MSW 2 "Fill in your First Name please"
MSL 1 1 "I am absolutely sure that I will solve the exercise correctly"
BGR "i\bg0.gif"
SLP 50
LAB "CT" "Intro Willy" 0 "" "" "" ""
IMG "i\intro_s.gif" 230 290 205 0
IMS 155 24 247 20 "Hallo, I am Willy"
WAV "En\Intro_wil.wav"
SLP 50
. . .
LAB "FM" "Do NOT Change" 0 "" "" "" ""
FRM ""

MSF, MSW define the text for the different messages that will be used in the test. MSL 1 1 are the labels for the first group of the first button in the 4-point color scale. BGR indicates the image file to be used as the background. SLP 50 lets the program "sleep" for 5 sec. LAB "CT" indi-

cates a command sequence for a cartoon. IMG describes the cartoon image and its place on the screen, IMS describes the text and the place of the cartoon balloon, and WAV is an optional sound file with the spoken text. FRM is a special command sequence that displays a form for obtaining information concerning the child (day of birth, grade, etc.).

A typical sequence for handling an exercise might look like this:

LAB "BL" "EX" 19 "A2" "G1" "OSC" ""
CLR
QWP "message text" "prompt" NumberOfButtons "Buttonlabel1" ...
WTA 50
PTA 2
SHL 2 0 0
WTA 50
PTA 3

LAB "BL" "EX" 19 indicates the start of an item (exercise, EX) 19. The good answer button is 2 (A2). It is an item for the first grade (G1) and belongs to subscale OSC. CLR clears the screen, and QWP displays the item on the screen. SHL displays the 4-point color scale with message group 2 (defined with MSL, as was described above). (The extra arguments are of no importance for this example.) WTA waits until the child presses a button; 50 indicates that, after 5 sec, a warning message will be displayed to

encourage the child to press a button. Finally, PTA puts the answer away in the second field of the answer array for this item.

When the test is finished, the results for the different subscales, as described above, are displayed and can be printed. Four commands are used for describing the output of the test on the screen. The first command (RUL) describes how the test results are to be summed in the output. The answer on one particular test item, in terms of good or wrong, can be combined with the answer on the 4-point color rating scale. The second command (CAL) is a filter for selecting one or more subscales for which the results are to be summed. The third command (SAV) is used for saving results. The last command (PRN) is the actual print command that contains the message to be displayed. Only the first string in the PRN is used here. The other four strings are used for data collection and will be described later. For example,

LAB "CA" "input of information" 0 "" "" "" ""
RUL 2 "EL1" "AG"
CAL 2 "PO"
SAV "R" "a"
SAV "T" "f"
LAB "PR" "Input of information" 0 "" "" "" ""
PRN "Input of information $a/$f" "IOF"
"($a/$f*100)" "IOFT" "($f)"

If, in the evaluation part (EL), the first button of the 4-point color scale (EL1) was pressed and the child's an-

swer was correct (AG), then add 2 points (the first argument) to the total. Select only "PO" items (PO is the posttest, one of the subscales), The first argument of CAL indicates that the selector for summing up is the subscale. If this argument is 1, a grade is selected. SAV indicates that the result is saved in the "a" variable and the total number of items in the "f" variable. In the SAV command, a maximum of six variables (a–f) can be used. These variables are then used in the printout commands (PRN), as indicated in the next PRN command. In the PRN commands, the variables are preceded with a $.

There is a data collection option built into the instrument. On demand, a number of collecting data can be prepared for direct statistical processing with SPSS. For this purpose, the last four arguments of the PRN command are used. Arguments 2 and 4 (IOF and IOFT in the preceding example) describe the name of a particular column in the collected output. The formulas in Arguments 4 and 5 represent the corresponding data.

The Visual Front-End Editor

After loading a script file, a scrollable text window appears (Figure 3). The window contains a list with one line for each block in the script file. As was mentioned, each block consists of a number of lines, starting with a LAB command that is followed by other commands until the next LAB command. The user can click on one of these lines and can add, remove, or modify each block. The blocks are displayed as a whole on the screen, which
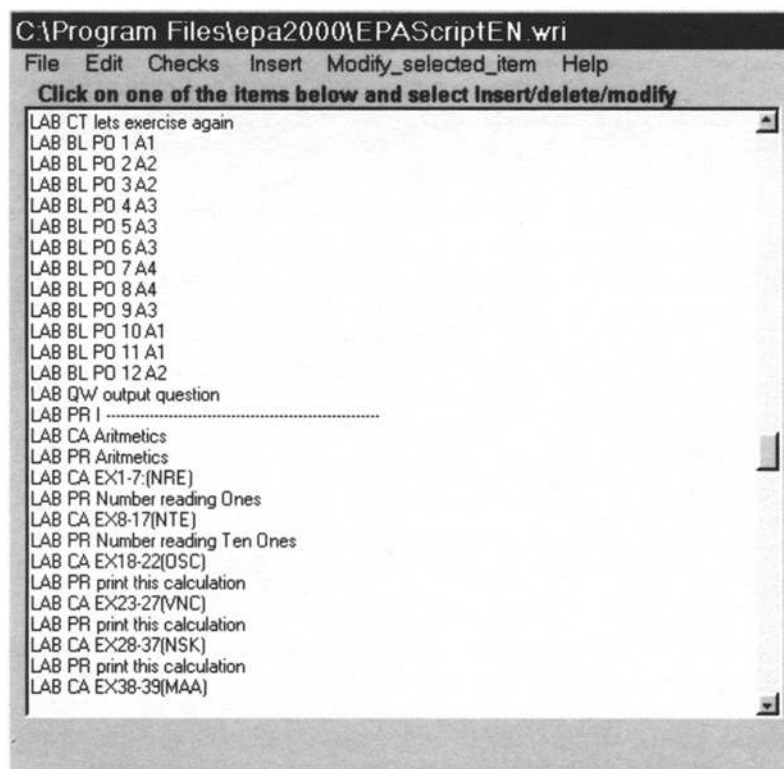


```
C:\Program Files\epa2000\EPAScriptEN.wri
File   Edit   Checks   Insert   Modify_selected_item   Help
Click on one of the items below and select Insert/delete/modify
LAB CT lets exercise again
LAB BL PO 1 A1
LAB BL PO 2 A2
LAB BL PO 3 A2
LAB BL PO 4 A3
LAB BL PO 5 A3
LAB BL PO 6 A3
LAB BL PO 7 A4
LAB BL PO 8 A4
LAB BL PO 9 A3
LAB BL PO 10 A1
LAB BL PO 11 A1
LAB BL PO 12 A2
LAB QW output question
LAB PR I ------------------------------------------------
LAB CA Aritmetics
LAB PR Aritmetics
LAB CA EX1-7:(NRE)
LAB PR Number reading Ones
LAB CA EX8-17(NTE)
LAB PR Number reading Ten Ones
LAB CA EX18-22(OSC)
LAB PR print this calculation
LAB CA EX23-27(VNC)
LAB PR print this calculation
LAB CA EX28-37(NSK)
LAB PR print this calculation
LAB CA EX38-39(MAA)
```

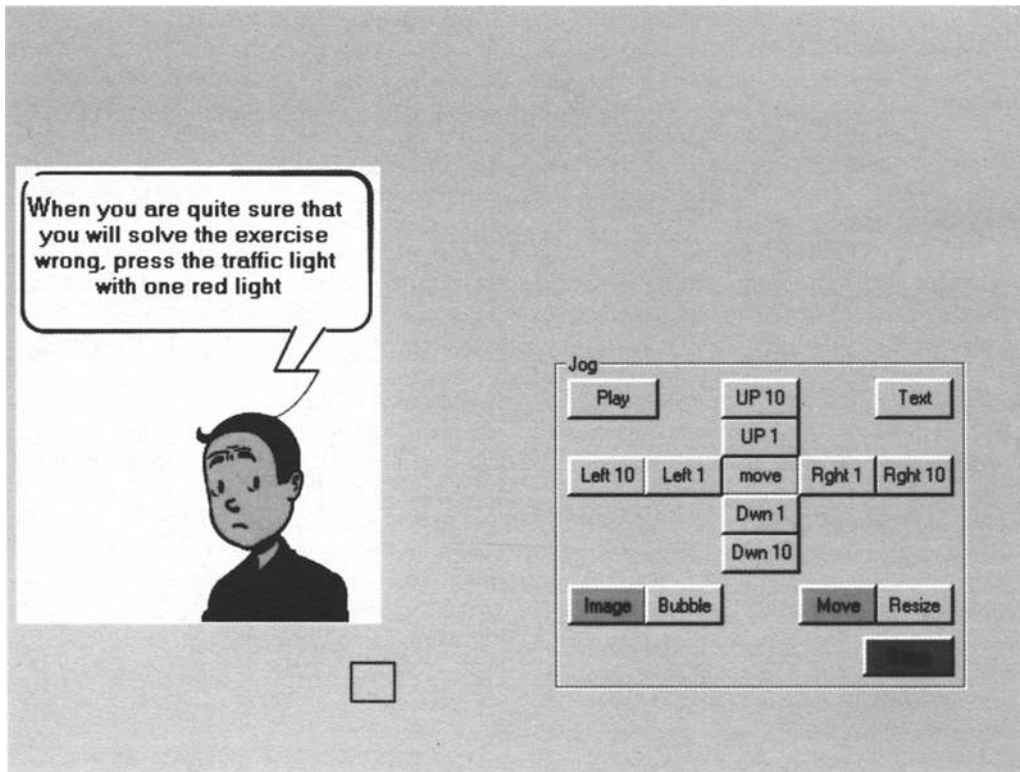Figure 3. Start screen of the front-end editor.

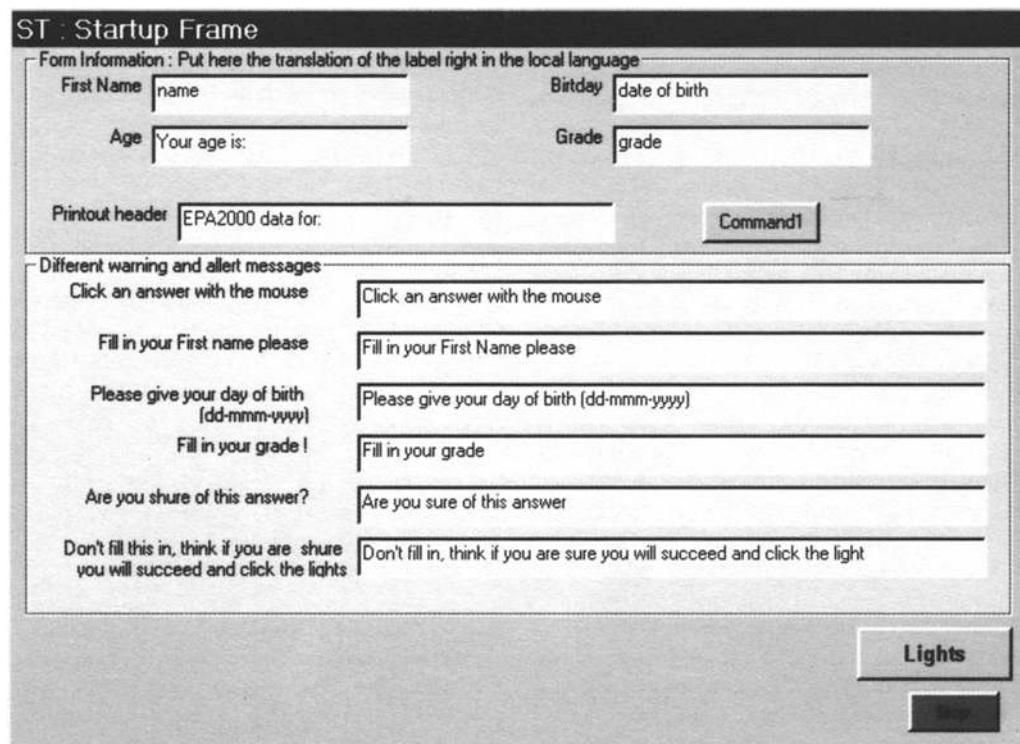**Figure 4. The editor screen for altering cartoons.**

Figure 5. One of the screens for modifying the button messages.

contains all the information that can be modified in the corresponding script lines. For every message on the screen, there is a small help text in English. The user can then type the corresponding text in his own language. The layout of every exercise is the same as that in the actual test and can be modified. Answer buttons can be added or removed. Also, exercises can be added or removed, and there is a consistency check when the items are saved. Figures 4 and 5 are two examples of display screens. The first is the screen for altering cartoons. With the control panel, it is possible to move the text and the picture. Another display for altering the file name of the picture and the audio file can be obtained by pushing the "text" button. The second is one of the screens for modifying the button messages. If one exercise is modified, the second, corresponding exercise is updated automatically.

DISCUSSION

This paper has offered some insights into the relationships between metacognitive skills in young children. Off-line metacognition seemed to be important in arithmetic. These findings are consistent with the research of Verschaffel (1999), stressing the importance of metacognition during the initial (prediction) and final (evaluation) stages of problems. Furthermore, the EPA2000 seemed to be useful in assessing the metacognitive differences in young children.

Previous versions of the EPA2000 have been tested in several studies between 1998 and 2000, in order to determine its usefulness for the age group and its sensitivity in measuring metacognitive differences. The psychometric data on reliability and validity seem promising.

Many commonly used paper-and-pencil instruments have been adapted for computerized administration (Schulenberg & Yutrzenka, 1999). They include the MMPI-2 (Pinsoneault, 1996), the Rotter Incomplete Sentence Blank (RISB; Rasulis, Schuldberg, & Murtagh, 1996), and many others. Although most of the instruments were tested with adult subjects, a number of comparisons were made for children, such as those for the Peabody Picture Vocabulary Test (PPVT; Elwood & Clark, 1978), the Wisconsin Card Sorting Test (WCST; Artiola i Fortuny & Heaton, 1996; Van Schijndel, 1994), the State–Trait Anxiety Inventory for Children (Katz & Dalby, 1981), and the Piers–Harris Children's Self Concept Scale (Simola & Holden, 1992). Most of the studies have suggested a trend that largely supports the equivalence of conventional and computerized instruments (Schulenberg & Yutrzenka, 1999). Still there are some aspects that need to be accounted for, such as the impact of computer aversion (Schulenberg & Yutrzenka, 1999; Tseng, Tiplady, Macleod, & Wright, 1998). In our studies, no computer aversion was observed.

Although the number of subjects was small, the test–retest correlations between EPA paper-and-pencil version and EPA2000 computer version ($r = .80$) were very good.

A program package, running under Microsoft Windows, for implementing the computer version of metacognitive assessment has been presented, with the contents of the test being easily modified and translated by a psychologist without computer knowledge. Since children with mathematical LDs can have problems in different numeral processes, different types of arithmetic tasks (see the introduction) were included (in subscales) in a pretest–intervention–posttest design. Furthermore, since it was the aim to use the computerized arithmetic and metacognitive assessment with other populations as well, it was necessary that the test could be adapted for different situations. Highly intelligent children with LDs differ from normally intelligent children with LDs. Subjects with combined mathematical and reading LDs (mathematical LD plus) differ from peers with specific mathematical LDs (mathematical LD), and so forth. Therefore, the program had to be sufficiently flexible to allow modifications of nearly all the parameters of the test without reprogramming the computer program. In addition, we wanted to create a multilingual test, to allow for cross-cultural assessment.

CONCLUSION

This paper has described the construction and use of a computerized assessment of arithmetic and metacognitive skills in young children.

Dividing the computer program into two separate parts—one part that executes a limited set of elementary commands, and a second part that generates/edits the command script—has a number of advantages. First, using a limited set of elementary commands keeps the system stable throughout the development stages; second, transporting the test to another environment (e.g., Mac) requires only a redesign of the script engine; third, using a separated editor makes the system sufficiently user friendly for nonprogrammers.

All the software, including the source, is free and can be downloaded (http://twiprof1.rug.ac.be/epa2000). A demo of the script file and of the cartoons is included. A complete version of the instrument and the cartoon pictures can be obtained by contacting A. Desoete (e-mail: anne.desoete@rug.ac.be).

REFERENCES

ACKERMANN, P. T., ANHALT, J. M., & DYKMAN, R. A. (1986). Arithmetic automatization failure in children with attention and reading disorders: Associations and sequelae. *Journal of Learning Disabilities*, **19**, 222-231.

AMERICAN PSYCHIATRIC ASSOCIATION (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

ARTIOLA I FORTUNY, L., & HEATON, R. K. (1996). Standard versus computerized administration of the Wisconsin Card Sorting Test performance in patients with multiple sclerosis. *Clinical Neuropsychologist*, **44**, 420-425.

ASHCRAFT, M. A., FIERMAN, B. A., & BARTOLOTTA, R. (1984). The production and verification tasks in mental addition: An empirical comparison. *Developmental Review*, **4**, 157-170.

BORKOWSKI, J. G. (1992). Metacognitive theory: A framework for teaching literacy, writing, and math skills. *Journal of Learning Disabilities, 25*, 253-257.

BROWN, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. Reiner & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65-116). Hillsdale, NJ: Erlbaum.

BUYENS, J. (1997). *Running Microsoft FrontPage 98.* Redmond, WA: Microsoft Press.

CAMPBELL, J. I. D. (1998). Linguistic influences in cognitive arithmetic: Comment on Noel, Fias and Brysbaert (1997). *Cognition, 67*, 353-364.

CARR, M., & JESSUP, D. L. (1995). Cognitive and metacognitive predictors of mathematics strategy use. *Learning & Individual Differences, 7*, 235-247.

CATTEEUW, G., & GHESKIERE, P. (1987). *GRIPA.* Kortrijk, Belgium: Vrij PMS-centrum3.

CRACCO, J. BAUDONCK, M., DEBUSSCHERE, A., DEWULF, B., SAMYN, F., & VERCAEMST, V. (1995). *Kortrijkse Rekentest* [Kortrijk arithmetic test]. Kortrijk, Belgium: Revalidatiecentrum Overleie.

DESOETE, A., ROEYERS, H., & BUYSSE, A. (2000). *Metacognition and arithmetic performances in primary school.* Manuscript submitted for publication.

DESOETE, A., ROEYERS, H., & DE CLERCQ, A. (2000). *Assessment of metacognitive skills in young children with arithmetic disabilities.* Paper presented at the Regional Conference Program IACE, Leiden.

DUDAL, P. (1985). *Vraagstukken 2de 3de leerjaar* [Word problems, Grades 2 and 3]. Torhout: Mach.

EFKLIDES, A. (1999). Training domain-specific abilities: The case of experiential structuralism. In J. H. M. Hamers, J. E. H. Van Luit, & B. Csapo (Eds.), *Teaching and learning thinking skills. Contexts of learning* (pp. 105-130). Lisse: Swets & Zeitlinger.

ELWOOD, D., & CLARK, C. L. (1978). Computer administration of the Peabody Picture Vocabulary Test to young children. *Behavior Research Methods, Instruments, & Computers, 10*, 43-46.

FLAVELL, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-236). Hillsdale, NJ: Erlbaum.

GEARY, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*, 345-362.

KATZ, L., & DALBY, J. T. (1981). Computer-assisted and traditional psychological assessment of elementary-school-aged children. *Contemporary Educational Psychology, 6*, 314-322.

LIDZ, C. S. (1997). Dynamic assessment approaches. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (pp. 281-296). New York: Guilford.

LUCANGELI, D., & CORNOLDI, C. (1997). Mathematics and metacognition: What is the nature of the relationship? *Mathematical Cognition, 3*, 121-139.

MASUI, C., & DE CORTE, E. (1999). Enhancing learning and problem solving skills: Orienting and self-judging, two powerful and trainable learning tools. *Learning & Instruction, 9*, 517-542.

MCCLOSKEY, M., & MACARUSO, P. (1995). Representing and using numerical information. *American Psychologist, 50*, 351-363.

METCALFE, J. (1998). Cognitive optimism: Self-deception or memory-base processing heuristics? *Personality & Social Psychology Review, 2*, 100-110.

MONTAGUE, M. (1998). Research on metacognition in special education. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities* (pp. 151-183). Greenwich: JAI.

PINSONEAULT, T. B. (1996). Equivalency of computer-assisted and paper-and-pencil administered versions of the Minnesota Multiphasic Personality Inventory–2. *Computers in Human Behavior, 12*, 291-300.

RASULIS, R., SCHULDBERG, D., & MURTAGH, M. (1996). Computer-administered testing with the Rotter Incomplete Sentences Blank. *Computers in Human Behavior, 12*, 497-513.

RUTLAND, A. (1995). The validity of dynamic assessment methods for children with learning difficulties and nondisabled children. *Journal of Cognitive Education, 5*, 81-94.

SCHNEIDER, W. (1998, November). *E-Prime: A cross platform experiment generator studio for computerized behavioral research.* Paper printed at the Commercial Symposium of the Twenty-Eighth Annual Conference of the Society for Computers in Psychology, Dallas.

SCHULENBERG, S. E., & YUTRZENKA, B. A. (1999). The equivalence of computerized and paper-and-pencil psychological instruments: Implications for measures of negative affect. *Behavior Research Methods, Instruments, & Computers, 31*, 315-321.

SHALEV, R., MANOR, O., AUERBACH, J., & GROSS-TSUR, V. (1998). Persistence of developmental dyscalculia: What counts? Results of a 3-year prospective follow-up study. *Journal of Pediatrics, 133*, 358-362.

SIMOLA, S. K., & HOLDEN, R. R. (1992). Equivalence of computerized and standard administration of the Piers–Harris Children's Self-Concept Scale. *Journal of Personality Assessment, 58*, 287-294.

TOBIAS, S., & EVERSON, H. T. (1996). Assessing metacognitive knowledge monitoring. In K. Hagtvet (Ed.), *Advances in test anxiety research* (Vol. 7, pp. 18-31). Hillsdale, NJ: Erlbaum.

TOWERS, J. T. (1999). *Dreamweaver 2 for Windows and Macintosh: Visual QuickStart Guide.* Berkeley, CA: Peachpit Press.

TSENG, H. M., TIPLADY, B., MACLEOD, H. A., & WRIGHT, P. (1998). Computer anxiety: A comparison of pen-based personal digital assistants, conventional computer and paper assessment of mood and performance. *British Journal of Psychology, 89*, 599-610.

VAN SCHIJNDEL, F. (1994). *Winsconsin Card Sorting Test: Computerversie. Handleiding* [Computer version. Manual]. Lisse: Swets & Zeitlinger.

VERMEER, H. (1997). *Sixth-grade students' mathematical problem-solving behavior: Motivational variables and gender differences.* Leiden University.

VERMUNT, J. D. H. M. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies: A phenomenographic analysis. *Higher Education, 31*, 25-50.

VERSCHAFFEL, L. (1999). Realistic mathematical modelling and problem solving in the upper elementary school: Analysis and improvement. In J. H. M. Hamers, J. E. H. Van Luit, & B. Csapo (Eds.), *Teaching and learning thinking skills: Contexts of learning* (pp. 215-240). Lisse: Swets & Zeitlinger.