HUMAN–COMPUTER INTERACTION, 2013, Volume 28, pp. 518–547

ISSN: 0737-0024 print / 1532-7051 online
DOI: 10.1080/07370024.2013.789348

Do Pedagogical Agents Enhance Software Training?

Hans van der Meij[1]
[1]*University of Twente, The Netherlands*

This study investigates whether a tutorial for software training can be enhanced by adding a pedagogical agent, and whether the type of agent matters (i.e., cognitive, motivational, or mixed). The cognitive agent was designed to stimulate students to process their experiences actively. The motivational agent was designed to increase perceived task relevance and self-efficacy beliefs. A mixed agent combined these features. Process and product data were recorded during and after software training of students from the upper grades of vocational education (M age = 16.2 years). Comparison of scores on performance measures during training revealed a significant advantage of working with the motivational and mixed agents for two important motivational mediators for learning (i.e., strategy systematicity and mood). All students were highly successful during training, improving from an average 30% task completion score on the pretest to a 77% posttest score. On a retention measure 3 weeks later, task completion was still at 66%. Working with the motivational and control agents yielded significantly higher retention scores, whereas working with the motivational and mixed agents led to significantly higher scores on task relevance and self-efficacy beliefs after training. The discussion reflects on the possibilities for improving the internal and external properties of the agents.

1. INTRODUCTION

People generally prefer to be active and to work toward meaningful goals when becoming familiar with software. Therefore, tutorials should be aligned with this tendency; the learner's self-initiated efforts to find meaning in the activities should be obstructed as little as possible (Carroll, 1990, 1998). The user's action-oriented focus may come at the expense of the learning of critical concepts and skills, however. For example, when a tutorial discusses text formatting, users need to become familiar with concepts such as paragraphs, margins, and citations. In addition, users need to

Hans van der Meij is a researcher with an interest in instructional design; he is Senior Lecturer in the Department of Instructional Technology of the University of Twente, the Netherlands.

CONTENTS

learn to master basic formatting skills. Inadequate learning thwarts task progression, forcing the user to consult earlier instructions over and over again. Just as with any instructional design, a tutorial is therefore a trade-off. It assists users with learning relevant knowledge and skills while also supporting their preference for action and for accomplishing meaningful tasks as much as possible.

Realization of this balance has long been considered an important and complex issue in software training (Carroll & Rosson, 1987; Fu & Gray, 2004). To accomplish learning, special measures are needed to enhance user knowledge of specific concepts or tasks. Users should be cognitively stimulated to process the information actively, to make inferences, and to reflect upon their actions. In addition, they should follow up tutorial-directed actions with self-directed actions in order to consolidate learning. The designer must tread carefully, however, as there is very little room for moving away from complete and basic instructions. The tutorial must not tax the user to the point of feeling that it no longer sufficiently supports task execution. Past this tipping point, users will cast the tutorial aside.

Motivation also needs special attention. The tutorial must stimulate users to take on self-directed actions and to try out new goals and new methods (Bhavnani & John, 2000; Van der Meij & Gellevij, 2004; Van Loggem, 2007). In addition, it must assist users to feel confident in overcoming the obstacles and failures that are an inevitable part of software training and use (e.g., Lazonder & Van der Meij, 1995; Norman, 1994; Van der Meij, 2003; Van Loggem, 2007).

The aforementioned considerations suggest that a tutorial would benefit from supplementary support for user cognition and motivation. This study investigates how well a virtual person in the form of a pedagogical agent (henceforth agent) can serve these supportive purposes. First, we discuss how to design such an agent. Next, we present an empirical study on the effect of an added agent in a tutorial. The aim of this study was to investigate whether the effectiveness of a software tutorial on formatting tasks in Microsoft Word is enhanced by including an agent with a particular design: cognitive, motivational, or cognitive and motivational.

2. DESIGNING THE COGNITIVE AGENT

Cognitive agents have been employed, among other uses, to model narrative skills in developing the literacy of preschool children (Ryokai, Vaucelle, & Cassell, 2003), to support concept development about a river ecosystem for fifth-grade students (Holmes, 2007), and to develop the communication skills of human service professionals (Duggan & Adcock, 2007). Such agents have been designed for university students for learning about electric motors (Mayer, Dow, & Mayer, 2003), computer literacy topics (Craig, Sullins, Witherspoon, & Gholson, 2006), concepts from the human circulatory system (Dunsworth & Atkinson, 2007), and solving word problems (Lusk & Atkinson, 2007).

In her review of theoretical and empirical studies on the use of cognitive agents in learning, Moreno (2005) concluded that there is more support for effects of the internal properties of cognitive agents, such as their content, actions, or instructional methods, than there is for external properties such as their visual and auditory presence. This contention is based on studies that examined a self-explanation effect for agents, among other possible effects. In that research, agents were found to affect learning when they stimulated students' active processing of the information and when they stimulated students to reflect after task completion.

In line with these findings, the design of the cognitive agent in this study concentrates on her internal properties, that is, the content of her comments. The comments made by the cognitive agent were based on two views of important learning processes: general processes for active learning, and specific reflective processes.

Contemporary notions of learning emphasize that students should actively process the learning material (Bransford, Brown, & Cocking, 2002). Students who actively engage in meaningful cognitive processes learn more deeply than students who passively process the information that is presented to them. In software training there is the risk that the user will not engage in the thinking necessary for the realization of learning. The production paradox (Carroll & Rosson, 1987; Fu & Gray, 2004) states that efforts to accomplish tasks easily supersede efforts toward comprehension and learning. Supplementary support therefore appears to be necessary to stimulate user engagement in processes that are essential for active learning.

[1]The research reviewed in the sections on designing a cognitive or motivational agent derives from software agents. That is, the agent—which generally is referred to as animated pedagogical agent—is included in the software rather than as a virtual person in a tutorial.

According to Mayer (2001, 2005), selecting, organizing and integrating relevant material are three general processes that are essential for active learning. Selection of relevant material occurs when the user attends to the pertinent elements in the instructions and interface elements presented in the tutorial. Engaging in this process means that the user brings these materials into working memory. Organization of relevant material means that the user connects various components into a coherent model. For a tutorial this means, for example, that the user must process the cause–effect relation between performing an action and perceiving the changes that result on the screen. Integration of relevant material refers to the process of connecting a model with prior knowledge. A tutorial can, for example, explain software jargon in lay terms. The three processes are described and several examples of associated cognitive agent comments are presented in Figure 1.

To determine which specific processes might be particularly effective for learning, we turned to studies of worked examples. Worked examples give detailed and complete information about solution procedures (Reimann & Neubert, 2000), just like the instructions in a tutorial. This research indicates that without additional prompting, students do not engage in active-enough processing of the information presented in the worked example to achieve successful learning (e.g., Atkinson, Derry, Renkl, & Wortham, 2000; Chi, Bassok, Lewis, Reimann, & Glaser, 1989). Researchers in the field of worked examples have studied three design solutions for enhancing learning,

FIGURE 1. General processes for active learning stimulated by the cognitive agent.

Name	*Description*	*Examples*
	General processes for active learning	
Selecting	Paying attention to relevant material: (a) goal information, (b) conceptual information, (c) input devices, and (d) words or pictures from the screen	The agent draws the student's attention to the location of a screen object (e.g., "I need to look carefully to see where I can find that on the screen.") The agent cautions about selecting the right object (e.g., "I need to use only the upper button.")
Organizing	Building internal connections between selected goals, concepts, input devices and interface features to create a coherent model	The agent wonders about the use of a method for a particular task (e.g., "Should I use the Alt-key here too?") The agent checks whether the computer is in the right system state before, during or after an action (e.g., "Is this correct?") The agent tries to fully understand a concept (e.g., "Now, can I do something with the upper and lower margins too?")
Integrating	Building external connections between internal models and prior knowledge	The agent reacts to the presentation of a design problem by wondering whether she also experiences that problem occasionally (e.g., "Do I recognize this?") The agent connects Word's jargon to prior knowledge (e.g., "So, margins and borders are the same.")

namely, fading instructions, including complementary activities, and stimulating reflection (see Atkinson et al., 2000; Atkinson & Renkl, 2007). The first two features are already incorporated in the No-Agent Control tutorial (see Section 4.2). Therefore, the design of the Cognitive agent concentrated on stimulating reflection.

In a seminal study, Renkl (1997) gathered think-aloud protocols from students' spontaneous reflections while processing worked examples. Analysis of these protocols revealed seven types of reflective processes: principle-based explanations, goal-operator combinations, anticipative reasoning, elaboration of problem situation, noticing coherence, monitoring-negative ("I don't understand"), and monitoring-positive ("Oh yeah, I see"). Analyses of the links between these activities and students' learning outcomes revealed that principle-based explanations and anticipative reasoning were most strongly (and significantly) correlated with learning. We therefore decided to attend to these particular reflective processes in designing the cognitive agent. In anticipative reasoning, the student thinks ahead to what needs to be done or what may happen next. This can include predicting the next step in a procedure and checking whether this prediction was correct. In principle-based explaining, the student tries to understand the nature of the task as well as the solution processes.

Several more recent empirical studies have further established the effectiveness of these reflective processes for learning. In addition, this research also suggests that these processes are best taught through prompts rather than with direct instructional methods (see Atkinson & Renkl, 2007; Atkinson, Renkl, & Merrill, 2003; Schworm & Renkl, 2007).

The design of the cognitive agent in our study aligns with these findings. That is, the agent stimulates the learner to engage in anticipative reasoning and principle-based explaining, and the agent does so primarily by prompting. These processes are described and several examples of associated agent comments are presented in Figure 2.

3. DESIGNING THE MOTIVATIONAL AGENT

Motivational agents have been designed to target a diverse set of goals and audiences. Among other uses, they have been developed to help autistic children engage in reciprocal social interactions (Tartaro & Cassell, 2007), to influence the emotions and motivation of elementary school children using software (Van der Meij, 2008), to moderate affective experiences during math for low-achieving high school students (Arroyo, Woolf, Cooper, Burleson, & Muldner, 2011, Woolf et al., 2010), to motivate workers in the footwear industry who are using an on-the-job computer-based training environment (Paiva & Machado, 2002), to model emotions and motivation through interactive pedagogical drama for mothers of pediatric cancer patients (Marsella, Johnson, & LaBore, 2000), and to enhance the adoption of an exercise program for adults (Bickmore, 2003). Agents have been designed for university students to reduce frustration in human-computer interaction systems (Hone, 2006; Klein, Moon, & Picard, 2002); to influence social judgments, interest, and self-efficacy for an e-learning system about instructional design (Kim, Baylor, & Shen, 2007); to model empathetic reasoning and behavior in a gamelike virtual environment (McQuiggan

FIGURE 2. Specific reflective processes for active learning stimulated by the cognitive agent.

Name	*Description*	*Examples*
Specific reflective processes for active learning		
Anticipative reasoning	Thinking ahead to what needs to be done, or what may happen next. This can include predicting the next step in a procedure and checking whether this prediction was correct	The agent says what object(s) should appear on the screen (e.g., "All right. Pay attention. A double arrow will appear.") The agent considers the possible consequences of an action beforehand (e.g., "If I change something in the text, will the table of contents follow suit?") The agent expresses doubts about the necessity of an action (e.g., "Why use the Alt-key here? It works without it too. Or does it?")
Principle-based explaining	Trying to understand the nature of the task and the solution processes	The agent may wonder whether there is a connection with prior knowledge (e.g., "Do I recognize this?") The agent engages in a self-query about understanding a concept (e.g., "Paragraph titles are simple headers, aren't they?") The agent draws a conclusion about a solution process (e.g., "Word changes only the selected paragraph.") The agent draws a conclusion about a concept (e.g., "The chapter is the most important heading. It must be given the Heading 1 Style.")

& Lester, 2007); and to influence female students' attitudes and self-efficacy beliefs regarding engineering (Rosenberg-Kima, Baylor, Plant, & Doerr, 2008).

There are only a handful of empirical studies reporting on the effects of motivational agents. A recurring finding is that these agents make the user experience more engaging. This phenomenon has been called the persona effect. Studies that compare an agent condition (image and voice) with a control condition (no image, and voice replaced by text) suggest that the persona effect derives primarily from dealing with an embodied interface (e.g., Moundridou & Virvou, 2002; Van Mulken, André, & Müller, 1998). Users find the agent attractive because it enlivens their interactions with the computer. The visual and auditory presence of a virtual person renders these interactions more human-like and more social. However, inconsistent effects of motivational agents are reported for direct measures of motivation such as task relevance and self-efficacy, as well as for carry-over effects on learning (e.g., Arroyo et al., 2011; Baylor & Kim, 2003; Bickmore, 2003; Domagk, 2010; Hone, 2006; Kim et al., 2007; Klein et al., 2002; Rickenberg & Reeves, 2000; Rosenberg-Kima et al., 2008; Van der Meij, 2008; Woolf et al., 2010).

Nearly all studies on motivating agents examine the influence of the agent's external features. This study followed a different approach. Just as for the cognitive agent, the design of the motivational agent concentrated on her internal properties, that is, the content of her comments. The motivational agent's focus was derived from an expectancy-value model of achievement motivation (Eccles & Wigfield, 2002). According to this model, the most important motivational predictors of behavior are task value and expectancy for success. Task values have to do with incentives or reasons for task engagement. A commonly used term for the concept that refers to these values is task relevance, which can be defined as a person's valuation of, interest in, and commitment to achieving a particular goal (Pintrich & Schunk, 2002). Expectancies are beliefs about success that affect goal setting, activity choice, and willingness to expend effort and persistence. In this study we focus on the learner's self-efficacy belief, which refers to the student's expectancy for success in tasks with novel or ambiguous elements (Bandura, 1997).

Keller's ARCS model is a widely adopted model for instructional design that gives extensive advice on how to enhance these two types of self-appraisal (Keller, 1987, 1999, 2010; Keller & Kopp, 1987). This model includes four conceptual categories that subsume many facets of motivation. That is, it gives advice on how to enhance attention, relevance, confidence, and satisfaction. We concentrated on the relevance and confidence components because their design guidelines address perceptions of task relevance and self-efficacy beliefs.

One empirical study has used the ARCS model to design a Motivational agent for a tutorial (Van der Meij, 2008). Only a small positive effect on motivation and no effect on learning were found. In that study, a variety of factors may have limited the effects of the agent, such as high initial levels of motivation, low task difficulty perceptions, and a ceiling effect on resulting motivation. Several other empirical studies without agents have reported significant effects of the ARCS design strategies on these motivational constructs (e.g., Feng & Tuan, 2005; Huett, Kalinowski, Moller, & Huett, 2008; Keller & Suzuki, 2004; Loorbach, Karreman, & Steehouder, 2007; Loorbach, Steehouder, & Taal, 2006; Song & Keller, 2001).

The two concepts on which the design of the Motivational agent in this study concentrates are described in Figure 3, and several examples of associated agent comments are presented. The presence and nature of task relevance comments from the agent vary depending on the task or situation. Self-efficacy comments show a growing belief in capacity within and across tasks; the agent expresses stronger beliefs in competence and positive outcomes as task execution progresses.

To make the motivational agent a more convincing model, and to strengthen positive or moderate negative motivational states, her content was extended with comments that conveyed a broad range of emotions and feelings (Baylor & Kim, 2004; Clore & Palmer, 2009; Dehn & Van Mulken, 2000). This was done through the inclusion of additional motivational words and comments. Words such as "boring," "cool," "love to," "handy," "odd," and "cute" were loosely based on a validated list of 500 words from a motivational lexicon (Ortony, Clore, & Foss, 1987). Examples of motivational sentences are "Very annoying" and "I am curious." Occasionally, intensifiers such as exclamation marks (e.g., "Yep!!" and "Made it!!!") were added to further strengthen a comment.

FIGURE 3. The two key motivational concepts addressed by the motivational agent.

Name	*Description*	*Examples*
	Motivational concepts	
Task relevance	Thinking about the present or future value of task engagement. This includes expressing topic interest.	The agent connects the event to the student's experiences (e.g., "This is how I often see it, too") The agent mentions the current or future value of the task ("I can use this," "Very handy for reports and the like.") The agent expresses a need or desire (e.g., "I want to learn how to present my report nicely.") The agent models enthusiasm for the topic (e.g., "Yeah, a great idea," "How nice that I can do this," and "I am curious.")
Self-efficacy belief	Expectations about one's capacity to organize and execute a course of action. These include ascribing positive outcomes to effort or competence.	The agent anticipates being successful (e.g., "Not a problem. I just have to choose the right buttons.") The agent ascribes success to effort rather than easy tasks or luck (e.g., "Sort of difficult, but a success in the end."). The agent compliments herself on task success (e.g., "Great. Now I can adjust margins just how I want to", "Gotcha", and "Almost right.") The agent moderates a negative outcome (e.g., "Pff, that was a tough exercise", and "I'm a bit off.")

The agent's motivational support was expected to affect learning indirectly. According to the cognitive-motivational process model from Vollmeyer and Rheinberg (1999, 2006), the influence of motivation on learning is mediated through strategy systematicity and motivational state. The former term refers to the extent to which a student uses strategies methodically in interacting with the learning material. This can range from not engaging in the task at all to rigorously working through all of the activities. Motivational state is a monitor of the fun, fear, frustration and similar feelings that students experience during training, as well as the confidence they express in their capacity for dealing with the task requirements. In the present study we refer to this motivational state as mood. Vollmeyer and Rheinberg (1999, 2006) reported on several experiments in which they found an effect of motivation on learning outcomes through these mediating processes.

4. RESEARCH QUESTIONS

Four conditions were compared. In the control condition, a tutorial without agent was presented. All three experimental conditions included an agent. In the cognitive and motivational agent conditions, the agent addressed cognition and motivation, respectively. In the mixed agent condition, comments from the cognitive and motivational agent were combined. The study addresses three research questions:

RQ1: Does condition affect learner effectiveness, mood states, and strategy systematicity during training?

Two measures for effectiveness during training are *error rates* and *success rates*. The error rate is the percentage of tasks in which at least one incorrect step is taken during task execution (e.g., making a wrong menu choice). Success rate stands for the percentage of correctly completed training tasks. Both measures reflect skills development, and it was expected that students in the cognitive and mixed agent conditions would do better on these measures than students in the two other conditions.

In line with the research of Vollmeyer and Rheinberg (1999, 2006), the analyses of *mood* concentrated on its valence (i.e., positive, neutral, or negative). It was expected that students in the motivational and mixed agent conditions would report more positive mood states than students in the other two conditions.

Strategy systematicity was measured by looking at *skip rates* and *method adoption*. Skip rate is the percentage of formatting tasks that the student makes no attempt to solve. Method adoption refers to taking up and following the course of action indicated in the tutorial (i.e., formatting the margins of an entire text by starting at the right rather than the left). It was expected that students in the motivational and mixed agent conditions would have lower skip rates and that more of these students would adopt the prescribed method than students in the other two conditions.

RQ2: Does condition affect learning gains?

The student's skill levels were assessed at three points: in a pretest, posttest, and retention test. Each test assesses the student's capacity to complete the six formatting tasks from the training. It was expected that all students would realize substantial *learning gains* (pre–post and preretention). Because the cognitive and mixed agents give more direct support for learning, these conditions were expected to yield higher learning gains than the other two conditions.

RQ3: Does condition affect motivational gains?

Motivation was assessed right before and immediately after training. Each assessment revolved around the two key constructs from expectancy-value theory (Eccles & Wigfield, 2002), namely, perceptions of *task relevance* and *self-efficacy beliefs*. All students were expected to show motivational gains on task relevance and self-efficacy beliefs (pre–post). Because the motivational and mixed agents give more direct support for motivation, these conditions were expected to yield higher motivational gains than the other two conditions.

To control for confounding factors, data were also gathered on training time, cognitive load, and reading frequency and appreciation of the agent. In any study assessing learning, it is important to examine differences in *training time* across conditions. Training time may be influenced by how engaging each condition is, or the extent or difficulty of the learning material. If one condition leads to higher learning outcomes but takes longer, it is possible that this is due to time-on-task rather than to differences in instructional effectiveness. Cognitive load theory cautions designers not to tax the students beyond their capacities. The presence of too much information may result in cognitive overload and reduce learning (Van Merriënboer & Sweller, 2005). To check whether the agent is unduly burdensome, *cognitive load* was measured during training. Students in the experimental conditions could differ in how often they read the agent's comments and how they valued her comments. To check on this factor, students were asked to rate *reading frequency* and to give an *appraisal* of the agent after completing the training.

5. METHOD

5.1. Participants

Participants were 94 students from the upper two grades of vocational education in the Netherlands. Their mean age was 16 years 2 months. Gender was evenly distributed within and across conditions (i.e., 12 boys and 12 girls in both the motivational and the cognitive agent conditions; 12 boys and 11 girls in the mixed agent condition; 11 boys and 12 girls in the no agent control condition). The students came from four classrooms of a training center where they regularly had to use Microsoft Word for writing school reports.

5.2. Instruments

The no-agent control tutorial was the basis for all tutorials involved in the study. The following sections describe the domain of the tutorial, its general design, and the design of the agent.

Domain of the Tutorial

The no-agent tutorial presents instructions on formatting options in Word. The content is useful for, but generally not yet mastered by, the target audience. The first chapter deals with adjusting the right and left margins for an entire document. The second chapter concentrates on formatting paragraphs, citations and lists. The third and final chapter revolves around automatically generating a table of contents. Students are instructed to work with practice files for all tasks in the tutorial.

General Design of the Tutorial

The general design of the tutorial is based on the minimalist approach. That is, the tutorial is action and task oriented, offers support for the handling of mistakes, and facilitates different types of reading (Van der Meij & Carroll, 1998). Special attention was given to the issues of fading of support, stimulating engagement in complementary activities after presentation of instructions, and the choice of style of presentation.

Fading is the gradual removal of instructional support. It is most regularly used in the domains of mathematics and science, where worked examples are commonly studied. Fading has also been employed in software tutorials (Lazonder & Van der Meij, 1995; Leutner, 2000). Whenever possible, fading is used for the relevant tasks in the tutorial.

The first minimalist tutorials supported users with "on-your-own" sections that invited their engagement in complementary activities after processing the instructions (Bannert, 2000; Carroll, 1990). Later studies found that exercises formed a better design alternative. Exercises are prototypical classic tasks after instructions; on-your-own sections have more open goal descriptions (see Figure 4). Wiedenbeck, Zavala, and Nawyn (2000) reported very high compliance for exercises, with users performing nearly all actions required for task completion. In contrast, 7% of the on-your-own sections were ignored completely by their participants, whereas another 34% were explored only partially. Glasbeek (2004) likewise reported better compliance for exercises than for on-your-own sections. Both Wiedenbeck, Zila, and McConnell (1995) and Glasbeek (2004) also measured influence on learning. The former reported significantly higher learning gains for exercises than for invitations, whereas the latter found no difference. Based on these findings, the tutorials in this study always present an exercise immediately following basic instructions on a topic.

The tutorial presents information in a personal style that matches the presence of the agent in the experimental conditions. This style is most clearly evident in the action steps (see Figures 7a and 7b). Instead of the standard, formal format that implies but does not explicitly acknowledge the presence of a user (e.g., "Click Enter"), the instructions personally address the user (e.g., "You click Enter"). Research indicates that this type of personalization significantly enhances learning and slightly raises

FIGURE 4. An exercise and an on-your-own section.

<table>
<tr><th>Exercise: Inserting Text and Data</th><th>On Your Own: Inserting Text and Data</th></tr>
<tr><td>The display below illustrates part of a membership database. List these data in a worksheet and number the names as shown in the example. Save the file under a name of your own choice.

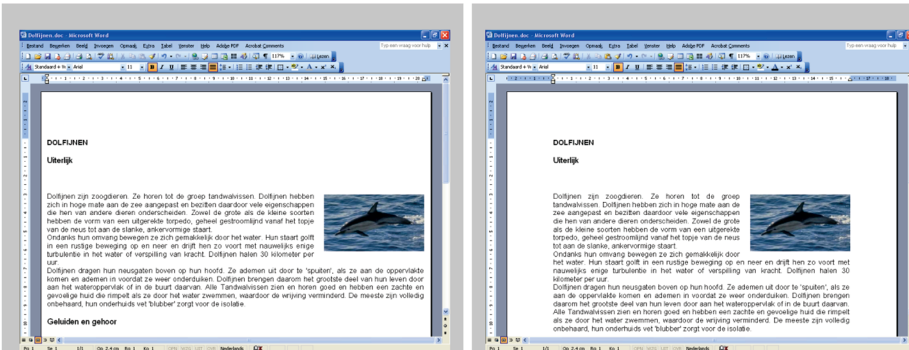
Contribution 2000/2001
<table>
<tr><th>Name</th><th>Paid</th><th>To-be-paid</th></tr>
<tr><td>1. J. Baz</td><td>240</td><td>0</td></tr>
<tr><td>2. M. Smith</td><td>120</td><td>120</td></tr>
<tr><td>3. K. Low</td><td>0</td><td>240</td></tr>
<tr><td>4. E. Karg</td><td>180</td><td>60</td></tr>
</table></td><td>Explore the possibilities for inserting text and data. Save the file under a name of your own choice.</td></tr>
</table>

***Note.* From "*I do what it says, but he does not*" by H. A. Glasbeek, 2001, Ultrecht University. Adapted with permission.**

FIGURE 5. Introduction for a chapter. (Color figure available online.)

1. Adjusting the margins for the whole text

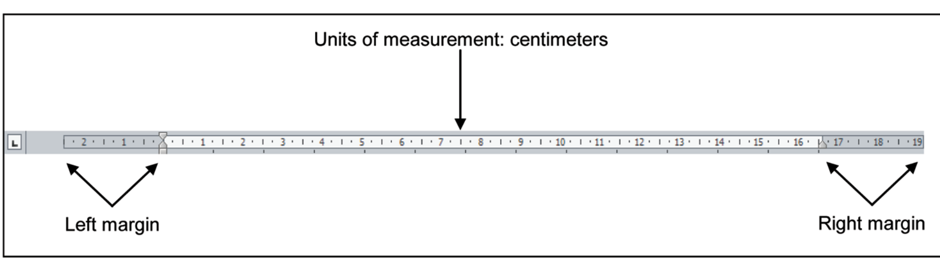
The margin is the space between the side of the paper and the text. Wide margins help give your text a clean look and create space for punch holes.



The problem: Occasionally the text appears too close to the left and right sides of the page.

The solution: You want the text to have plenty of white space on the left and right.

Word speaks of margins when it refers to text borders. You can change the margins of a text by using the buttons on the ruler. You can find the ruler immediately above your page.



interest as compared to a more formal style (Mayer, Fennell, Farmer, & Campbell, 2004; Moreno & Mayer, 2000, 2004).

On a more detailed level (i.e., for the instructions), the tutorial follows the four components model (Van der Meij, Blijleven, & Jansen, 2003; Van der Meij & Gellevij, 2004) that offers specific design guidelines for the treatment of goal information, prerequisites, action & reaction, and unwanted states.

Each chapter starts with a one-page introduction. This prologue discusses the main goal(s), defines the key concept(s), and situates and points out the objects that must be manipulated in the interface. A relevance organizer describes and illustrates the formatting goal (see Figure 5).

Agent Design in the Tutorial

Most pedagogical agents in human–computer interaction systems are animated. A common format is the talking head (e.g., Graesser et al., 2004; Moreno & Flowerday, 2006; Wang et al., 2008). The agent in our study is also a talking head. But rather than being animated and having a voice, our agent is a static, annotated photograph (see Mayer, Hegarty, Mayer, & Campbell, 2005). Our cognitive and motivational agents are both represented by a picture of the same person from the target audience because model-target similarity increases the effectiveness of the model (Bandura, 1997).

The introduction to the tutorial establishes the character of the agent (see Figure 6). The female agent named Lineke presents herself as someone from the target audience. Like the reader, she regularly needs to hand in nicely formatted reports for her school. She also says that she wants to give her reports a nice appearance but finds this difficult. In short, the agent is portrayed as a motivated beginner with little prior knowledge about the content of the tutorial. Kim, Baylor, and PALS Group (2006) found a similar agent profile beneficial for building user confidence.

The agent is easily recognizable as a separate information unit in the tutorial. She consists of a close-up picture, in color, of a girl's face, with comments presented at her side (see Figures 7 and 8). Popular youth magazines, along with pilot testing, helped provide input for creating audience-sensitive words and comments for the agent. The visual display of the agent is identical in all conditions. The photographs show her face and a small part of her upper body, in line with the recommendation to crop an image of a person just below the shoulders to create a pleasing picture (Agrawala, Li, & Berthouzoz, 2011). The agent's face expresses mood states such as pleasure, frustration, challenge, insecurity, cheerfulness, certainty, and surprise.

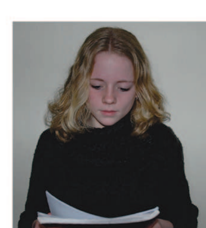
FIGURE 6. The agent's introduction. (Color figure available online.)

Introduction

Hello, I am Lineke. I am 15 years old and I attend secondary school. I regularly have to write school reports for geography and history. I often want to insert texts or pictures from the Internet in these reports. Things often go wrong when I try to do this. Sometimes the text is too wide and does not fit nicely into my page borders. Things also mess up frequently when I copy a list.

My teachers want me to hand in a clean-looking report. I also get a better grade then. My teacher says that it is very easy to do in Word. You merely need to get to know these facts. This tutorial can help us.

I am curious to find out about the tutorial's content. For now, it will take some time to go through it. I hope that it will help me make school reports that look nice.



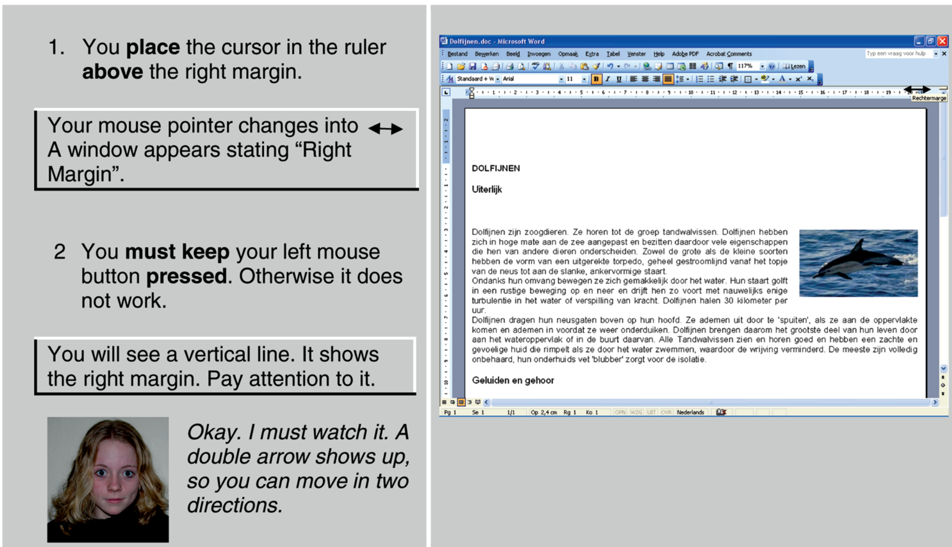


FIGURE 7. Procedural instructions supplemented with a cognitive agent. (Color figure available online.)

Measurement Instruments Before, During, and After Training

The questionnaires in the study all (except mood) use a 10-point Likert scale in which only the anchors are labeled. Depending on the question these endpoints are expressed as *completely agree – completely disagree*, *never – always*, *very easy – very difficult*, and the like. Students answer each question by making a cross at what they decide is the right place on an unmarked 10-cm line. The student's score is determined by using a transparent template that divides the line into 10 equal sections, by which the number

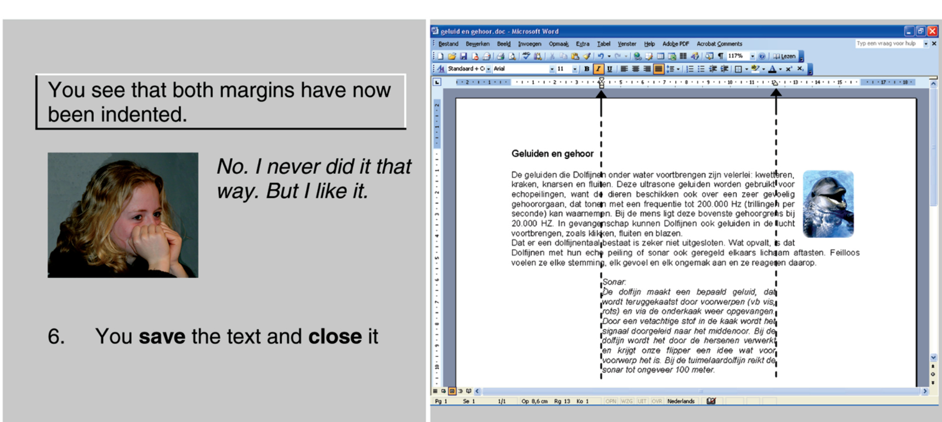


FIGURE 8. Procedural instructions supplemented with a motivational agent. (Color figure available online.)

corresponding to the position of the student's cross can be gauged. Questionnaires were presented in a paper-and-pencil response format.

Pretraining Instruments. The Computer Experience questionnaire asked about participants' experience with using computers. Participants in all conditions gave similar answers to these questions. These findings are therefore not discussed further.

The Prior Motivation and Pretest inventory assesses the students' motivational self-appraisals (i.e., *task relevance* and *self-efficacy belief*) and their prior skill. Six problems are presented, in the form of a screen shot plus explanation. For each, the student is asked to answer an experience question ("Do you ever have this problem?"), two motivation questions, and one skill question. The motivation questions are, "How often do you want to solve this problem?" (task relevance) and "How well do you think you can solve this problem?" (self-efficacy). Satisfactory Cronbach's alpha scores were found for task relevance (0.81) and self-efficacy beliefs (0.74).

In addition, students are asked a skill question that invites them to open a file on the computer and try to solve the problem in Word (prior skill). The inventory covers all subtasks discussed in the tutorial. There are two items for whole-text margins (left and right), three items on indentation (paragraphs, citations, and lists), and one item for creating the table of contents. Students are awarded a score of 0 points for each problem on the pretest that they cannot solve or for which an incorrect method is used (e.g., typing the table of contents). A good solution (and method) yields a score of 1. The maximum score for the pretest is 6. The *skill level* that is reported shows the percentage of successfully completed test tasks.

During-Training Instruments. The Mood, Cognitive Load & Time Questionnaire measures the named dependent variable during training. The mood question asks, "How do you feel after this task?" The student signals his or her mood by selecting the appropriate pictogram plus description (see Read, 2008). Five smileys are presented: happy, sure, neutral, unsure, and angry. Happy and sure smileys are scored as signs of a positive mood. Unsure and angry are seen as signs of a negative mood. The cognitive load question is the widely used one from Paas, Van Merriënboer and Adam (1994) which asks, "How hard was this task for you?" These questions were presented on a separate page. Each time a (sub)task was completed, for a total of 13 times, students would automatically encounter such a page in the tutorial.

All student actions in Word were logged with Camtasia. This program yields a video that shows all the screen states the student has encountered. This log was analyzed for scores on effectiveness (i.e., error rates and success rates) and strategy systematicity (i.e., skip rates and method adoption). The *error rate* is the percentage of all tasks for which a student made one or more mistakes. Errors were operationally defined as making a wrong choice during task execution. Simply hovering over various menus was seen as exploratory behavior and not coded as a mistake. The total number of mistakes per task is not tallied because errors easily entangle (Carroll, 1990), which would inflate the score. The *success rate* shows the percentage of successfully completed training tasks. A distinction is made between success during instructions and on exercises. The *skip rate* is the percentage of training tasks that a student makes no attempt to do. The skip rate signals whether students expend effort on getting to know a task. It is insensitive to success and does not depend on how well the students do at task execution. The measure *method adoption* indicates whether students follow the prescribed sequence for setting the margins of a whole text and stick to this method in the exercise. From previous studies we know that students tend to start at the left when adjusting text margins. In contrast, the tutorial instructs students to begin with the right margin because it is a simpler method (i.e., the hidden object "↔" is easier to find). The data are presented as a percentage, with the number of students in a condition who adopted the prescribed method divided by the total number of students in that condition.

Posttraining Instruments. The Postmotivation questionnaire measures student motivation for the formatting tasks. Seven questions assess *task relevance* (e.g., "I find the ruler very handy for setting the margins."). The Cronbach's alpha of 0.76 for this scale was satisfactory. Nine questions measure *self-efficacy belief* (e.g., "Now I know how to make a nice table of contents."). The Cronbach's alpha of 0.78 for this scale was also satisfactory.

The Agent questionnaire begins with a question about agent *reading frequency* ("How often did you read the agent's comment?"). Next, there are 11 questions on *appraisal* of the agent (e.g., "I agreed very well with Lineke's comments" and "I felt just like Lineke did."). The Cronbach's alpha of 0.88 for this scale was good.

In the posttest, students must complete on the computer six formatting tasks that are similar to those in the tutorial. The test instructions describe each task and display a screen shot of what should be the end result. The retention test is identical to the posttest, except that students are given another Word document to work on. Both tests are scored in exactly the same way as the pretest.

5.3. Procedure

A week before training, students completed the Computer Experience questionnaire and the Prior Motivation and Pretest inventory. Training took place in a computer room, one class at a time. Students were instructed to "work through the whole tutorial, which helps you format reports better." They were to work on their own and ask the experimenter for help only when stuck. They were also instructed to answer the questions about cognitive load, mood, and time that were presented in the tutorial. Students could work with the tutorial for a maximum of 100 min, which piloting had revealed should be sufficient for all or nearly all students. Once they had completed the tutorial or reached the maximum time and after a short break, students answered the Postmotivation questionnaire. Students in the experimental conditions also received the Agent questionnaire. All students then completed the posttest; after a delay of 3 weeks they returned to complete the retention test. Students were not allowed to use the tutorial at any point during testing. The tutorial was also not available in the period between the posttest and retention tests.

5.4. Data Analysis

The study is quasi-experimental with four conditions and assessments before, during, and after training. A chi-square score was computed for method adoption. For frequency counts such as the skip rate, success rate, and all test scores, comparisons were analyzed with analyses of variance (ANOVAs). If an ANOVA was found to be significant for a predicted effect it was followed by planned comparisons. Pretest, posttest, and retention scores were treated as repeated measures. Degrees of freedom in the analyses vary across measures due to missing data. In all analyses the significance level was set at an alpha of 0.05 (two-tailed). Cohen's (1988) *d*-statistic is reported for effect size. These tend to be qualified as small for $d = 0.2$, medium for $d = 0.5$, and large for $d = 0.8$.

6. RESULTS

6.1. Measures Before Training and Control Measures

Comparisons for prior motivation indicated that conditions differed significantly for self-efficacy beliefs, $F(3, 92) = 4.41$, $p < .01$. To assess differences between conditions for these beliefs after training, this difference was taken into account by treating initial state as a covariate. Comparisons for skill level on the pretest showed no statistically significant differences between conditions.

Comparisons for *training time* and *cognitive load* showed no statistically significant differences between conditions. The mean training time of 42 min was well below the preset maximum. For cognitive load, the mean score of 2.4 on the 10-point scale also signals that the students did not feel taxed beyond their limits.

Comparisons for *reading frequency* and *appraisal* of the agent showed no statistically significant differences between the agent conditions. The overall mean reading frequency score of 7.86 on the 10-point scale indicates that students said they often read the agent's comments. With a mean score of 4.62, appraisal was slightly below the scale midpoint of 5.

6.2. Performance Measures During Training

Students had an overall total average *error rate* of 23%. There were comparable error rates across conditions (see Figure 9). A comparison for instructions and exercises indicated that the latter drew more errors, $F(1, 75) = 6.59$, $p < .05$, $d = 0.37$.

Figure 10 shows the findings for *success rate*. Students completed 88% of all instructed tasks successfully. Condition did not influence these scores. There was a 17% decline in success rate for the exercises. All conditions showed a comparable decline, $F(1, 75) = 51.94$, $p < .001$, $d = 0.85$, and there was no interaction effect.

Condition did not affect *skip rates*. There were no statistically significant differences between conditions for skip rates for both instructions and exercises (see Figure 11). Instructions were almost never ignored. Virtually all students in all conditions attempted to work through them. The mean skip rate was higher for exercises,

FIGURE 9. Mean error rates (percentages) during training by type of support.

		Instructions		Exercises	
Agent Type	*N*	*M*	*SD*	*M*	*SD*
Motivational	21	19	14	24	12
Mixed	18	18	9	29	14
Cognitive	18	25	19	25	14
No agent	22	22	14	26	13
Average		21	15	26	13

Note. $N =$ number of students.

FIGURE 10. Mean success rates (percentages) during training by type of support.

		Instructions		Exercises	
Agent Type	*N*	*M*	*SD*	*M*	*SD*
Motivational	21	94	11	75	16
Mixed	18	87	13	72	20
Cognitive	18	85	20	70	19
No agent	22	84	20	73	23
Average		88	17	73	19

Note. $N =$ number of students.

FIGURE 11. Mean skip rates (percentages) during training by type of support.

		Instructions		Exercises	
Agent Type	*N*	*M*	*SD*	*M*	*SD*
Motivational	21	0	0	4	6
Mixed	18	0	0	6	8
Cognitive	18	2	5	12	12
No agent	22	0	0	6	11
Average		0	2	6	10

Note. $N =$ number of students.

$F(1, 75) = 38.6, p < .001, d = 0.94$, with the cognitive agent condition standing out with a rate at least twice as high as any of the others.

Condition significantly affected *method adoption* (see Figure 12). A greater percentage of students who worked with the motivational or mixed agent followed the prescribed sequence in the instructions than did students working with the cognitive agent or no agent tutorial, $\chi^2(1, N = 79) = 4.22, p < .05$. The data for the exercises also showed that more students in the motivational and mixed agent conditions followed the instructed sequence compared with students in the cognitive and no agent conditions, $\chi^2(1, N = 79) = 3.72, p = .05$.

FIGURE 12. Mean method adoption (percentages) during training by type of support.

Agent Type	*N*	Instructions	Exercises
Motivational	21	81	19
Mixed	18	72	28
Cognitive	18	67	6
No agent	22	45	9
Average		66	15

Note. $N =$ number of students.

6.3. Perception Measures During Training

The scores for mood indicate that students predominantly experienced a positive mood state during training. Neutral moods were reported at about five of the thirteen measurement instances. Negative mood states were very infrequent (see Figure 13).

There was a statistically significant difference between conditions for *positive* mood, $F(3, 91) = 3.20$, $p < .05$. Students in the motivational and mixed conditions more often reported having experienced a positive mood than did students in the other conditions, $t(88) = 2.33$, $p < .05$. There was no difference between the motivational and the mixed conditions.

There was also a statistically significant difference between conditions for *neutral* mood, $F(3, 91) = 3.83$, $p < .01$. Students in the motivational and mixed conditions less often reported having experienced a neutral mood than did students in the other conditions, $t(88) = 2.43$, $p < .05$. There was no difference between the motivational and mixed conditions.

Further, there was a statistically significant difference between conditions for *negative* mood, $F(3, 91) = 2.87$, $p < .05$. However, this was not due to a difference between the motivational and mixed conditions, on one hand, and the cognitive and control conditions, on the other. Exploratory post hoc analyses (least significant difference statistic) indicated that students in the cognitive condition more often

FIGURE 13. Mean reports of moods (frequencies) during training, by type of support.

		Positive		Neutral		Negative	
Agent Type	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Motivational	23	9.53	3.37	3.29	3.33	0.17	0.49
Mixed	23	7.74	4.16	4.48	3.85	0.78	1.40
Cognitive	24	7.55	4.08	4.44	3.72	1.01	1.92
No agent	22	5.64	5.12	7.23	5.08	0.14	0.35
Average		7.64	4.37	4.83	4.22	0.53	1.28

Note. $N =$ number of students.

reported negative mood compared to students in the motivational ($p < .05$) or control conditions ($p < .05$).

6.4. Performance Measures After Training

There was no difference between conditions on the pretest. Students began with a mean skill level of 30% successful completion on this test (see Figure 14). On the posttest this had risen to almost 77%. The difference from the pretest scores was both statistically significant and large, $F(1, 79) = 301.0$, $p < .001$, $d = 2.04$. On the retention test these scores decreased to 66%, but this was still well above the pretest score, $F(1, 79) = 164.0$, $p < .001$, $d = 1.53$.

There was no difference between conditions on the posttest, but conditions differed significantly on the retention test, $F(3, 82) = 5.49$, $p < .01$. The retention scores were in the opposite direction from what was expected, however. Students working with tutorials that included the cognitive agent (i.e., cognitive and mixed) scored lower for retention than those using the other tutorials, $t(79) = -3.75$, $p < .001$. The cognitive and mixed conditions did not differ from each other.

6.5. Perception Measures After Training

Task relevance was rated much higher after training (see Figure 15) than before, $F(1, 87) = 212.8$, $p < .001$, $d = 1.94$. An ANOVA showed a marginally significant difference between conditions, $F(3, 90) = 2.46$, $p = .068$. Students in the motivational and mixed conditions rated the formatting tasks as significantly more relevant than did students in the other conditions, $t(87) = 2.67$, $p < .05$. There was no other difference between conditions.

Self-efficacy beliefs rose significantly and substantially following the training, $F(1, 87) = 43.6$, $p < .001$, $d = 0.79$. The analysis of covariance for self-efficacy beliefs showed a statistically significant difference between conditions, $F(3, 89) = 3.09$, $p < .05$. Students in the motivational and mixed condition rated their self-efficacy beliefs higher after training than did students in the other conditions, $F(1, 89) = 9.14$, $p < .01$, $d = 0.66$. There was no other difference between conditions.

FIGURE 14. Skill levels (percentages) on the pretest, posttest, and retention test.

		Pretest		Posttest		Retention Test	
Agent Type	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Motivational	22	39	24	83	14	72	18
Mixed	23	25	25	77	22	62	22
Cognitive	20	23	17	72	23	53	27
No agent	18	34	24	77	24	80	19
Average		30	23	77	21	66	24

Note. N = number of students.

FIGURE 15. Appraisals (means) of task relevance and self-efficacy beliefs before and after training.

		Task Relevance				Self-Efficacy Belief[a]			
		Before		After		Before		After[a]	
Agent Type	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Motivational	23	4.85	2.62	8.48	1.10	7.79	1.67	8.52	1.09
Mixed	23	4.52	2.08	8.60	0.82	7.37	2.06	8.52	1.07
Cognitive	23	4.56	2.12	7.99	1.32	6.27	2.29	7.72	1.07
No agent	22	5.35	2.10	7.85	1.14	5.74	1.94	7.90	1.10
Average		4.81	2.29	8.23	1.13	6.80	2.13	8.17	1.08

Note. Scales run from 1 to 10. A higher score means a more positive appraisal. N = number of students.
[a]Estimated marginal means.

7. DISCUSSION AND CONCLUSION

All tutorials used in this study yielded a significant and substantial effect on learning and motivation. Before training, students mastered 30% of the formatting tasks described in the tutorial. During training, students successfully completed 88% of these tasks when aided by instructions and 73% of the exercises. The lower score for exercises supports a basic contention from researchers in the field of worked examples that holds that instructions should be followed by exercises to consolidate learning (Atkinson et al., 2000; Reimann & Neubert, 2000). An error rate of 23% during training was found, meaning that students made at least one mistake for every five tasks. Performance on the posttest revealed that 77% of the trained tasks were completed successfully. Retention scores, measured 3 weeks later, further indicated that students also retained their skill well. All students substantially increased their appraisals of the task relevance of the formatting tasks encountered in training. In addition, they expressed greater self-efficacy beliefs.

In contrast to what was predicted, the presence of the cognitive and mixed agents did not lead to the highest performance scores during training, or on a test taken immediately after training. The cognitive and mixed agents who were expected to stimulate students to engage in more active processing of information led to the same error rates and task completion as did the no agent and the motivational agent. Likewise, the cognitive and mixed agent conditions did not result in higher performance on the immediate posttest. No effect of the intended cognitive stimulation was found because, to our surprise, the no agent control tutorial had already sufficiently yielded the active processing that needed to take place. Perhaps this was due to the fact that this tutorial incorporated two design measures from worked examples research, namely, some fading and the inclusion of exercises. The latter may have been especially important for learning to occur. Exercises turn a single experiential event into a preview for a second, more mindful attempt. The student can first simply execute the task by following the instructions. Then the exercise serves as a self-test and cue for learning. If the procedure has already become familiar, the user can complete the exercise without help. If not, the user can reread the instructions and apply them to the exercise.

Also in contrast to what was predicted, working with the cognitive and mixed agent led to significantly lower scores than the other conditions on the retention test. Perhaps this outcome is the result of a combination of "distraction" during learning and the moment of testing. The cognition-oriented comments ask students to attend to more issues than strictly needed for task execution, stimulating them to think about concepts, and to predict and reflect upon actions and software reactions, among other things. In an immediate test this information might not yet interfere with performance because the knowledge is still fresh. In a delayed performance test, however, the memory trace could be weaker, and it could be more difficult to retrace earlier steps, especially when one's attention has been partly diverted during initial exposure.

In line with the predictions, the motivational and mixed agent positively affected mood. The students in these two conditions predominantly reported experiencing a positive affective state. In addition, students working with the motivational agent or with the no agent tutorial reported fewer negative moods than students in the cognitive agent condition. Strategy systematicity was found to differ between conditions only for method adoption during instruction. As predicted, the motivational and mixed agent conditions did better than the other conditions on this measure. However, the effect was short-lived. When left to choose their own method (i.e., during the exercises) most students returned to the favored left–right working order. As predicted, the students gave significantly higher self-appraisals for perceived task relevance and self-efficacy beliefs in the motivational and mixed agent conditions than in the other conditions. This finding is in line with other (nonagent) studies in which the ARCS model has effectively enhanced student motivation (e.g., Feng & Tuan, 2005; Huett et al., 2008; Keller & Suzuki, 2004; Song & Keller, 2001).

As reported by Moreno (2005), designers should pay special attention to the internal properties of the agent. This does not mean that the external properties are irrelevant, however. The effectiveness of an agent depends also on satisficing at least minimal demands for appearance (compare Gulz & Haake, 2006). The agent should be a good mix of looks, actions, and words to be sufficiently appealing and credible to affect the student. In the following discussion we reflect on the internal and external properties of the agents we designed and the ways in which these might be improved in future studies.

To further investigate the agents' *internal* properties, two different routes are possible. One approach would involve renewed top-down analyses that could lead to new designs, or other design perspectives than currently taken. The other approach would involve a bottom-up process in which information is gathered from the students as they process the tutorial. This information could then be analyzed to serve as input for redesign. Both approaches are described next.

The poor results for the cognitive agent make that condition a primary candidate for a reappraisal of its basis. The design of the cognitive agent was grounded on Mayer's (2001, 2005) widely adopted and tested distinction between selecting, organizing, and integrating functions. We coupled these functions with suggestions from worked examples research about the specific information processes that can best be supported (i.e., Renkl, 1997). One avenue to take in redesigning this agent could involve examining different ways of expressing these processes. That is, the focus could shift from supporting anticipative reasoning and principle-based explaining to supporting students in reflecting on goal operator combinations, and in using both types of monitoring strategies as distinguished by Renkl.

Another option would be to examine an entirely different set of cognitive processes. Moreno (2005) stated that a cognitive agent could address processes such as cognitive load reduction, external memory expansion, feedback, modeling, and guidance. Clarebout, Elen, Johnson, and Shaw (2002) argued that the six main functions for a cognitive agent are supplanting, testing, coaching, modeling, demonstrating, and scaffolding. Heidig and Clarebout (2011) proposed yet another set of (teaching) functions: motivating, informing, information processing, storing and retrieving, transferring, monitoring, and directing. Research on animated pedagogical agents can perhaps provide insights that are helpful in selecting from these options for redesigning the cognitive agent.

Although it was reasonably successful, the design of the motivational agent also deserves further scrutiny. This agent was based on Eccles and Wigfield's (2002) expectancy-value model of achievement motivation, and we employed Keller's (Keller, 1987, 1999, 2010; Keller & Kopp, 1987) ARCS model to create the agents' supportive comments for the concepts of task relevance and self-efficacy belief.

How can the agent's comments be optimized for these motivational constructs? To answer this question, we could return to theory and conduct an in-depth analysis of research on the two key expectancy-value constructs to find out whether this yields complementary views for the agent's design. An exploration of the literature on self-efficacy beliefs has already yielded a potentially important difference. That is, Keller (Keller, 1987, 1999, 2010; Keller & Kopp, 1987) has mentioned three key strategies: learning requirements (building positive expectations for success), personal control, and success opportunities. Self-efficacy literature mentions these factors as well but also advocates the use of appropriate social modeling in general and the use of coping peer models who overcome learning difficulties in particular (Bandura, 1997). An extensive further study of the literature is needed to find out whether and how our characterization of the motivational agent needs to improve to create a satisfactory expression of such a model. The below-average scores for agent appraisal indicate that there is room for improvement in this respect.

Just as for the cognitive agent, alternative theories can also be chosen for the design of the motivational agent. Here the literature offers little direction, because there are few studies of pedagogical agents whose design is based on a motivational theory (see Heidig & Clarebout, 2011; Schroeder, Adesope, & Gilbert, 2012). One exception is the research from Woolf and her colleagues (2010), who used Weiner's attribution theory to design an effective (animated) pedagogical agent for low-achieving students in math.

In assessing new design options it is important to keep in mind that the primary aim of the software tutorial is to contribute to students' skills development. This sharply contrasts with most of the research on multimedia learning, worked examples, and cognitive (animated) agents, which focuses on supporting conceptual knowledge development. In view of this difference in aim, perhaps totally different cognitive processes than hitherto considered might benefit from scaffolding. Insights about these processes can come from in-depth audience analyses. That is, there should be a close examination of the immediate reactions of the students to the tutorials. This requires a new study in which the students' think-aloud protocols are gathered as they process the tutorial. The students' comments, in combination with their logged actions, can then be classified as cognitive and/or motivational, and further detailed within each category.

These data are preferably gathered from all versions of the tutorials, because each condition is likely to offer complementary views on what benefits or hinders the students. For instance, protocols from the control tutorial can detect key moments at which students encounter obstacles that call for cognitive or motivational support. Likewise, protocols from the cognitive, motivational, and mixed tutorials can indicate whether students are supported or distracted by the agent's comments. In combination with action logs and a record of task completion times, these data might show when a comment stimulates the student to restudy, when it reduces student anxiety, when it confuses the students, or when it is unduly taxing.

With regard to the (re)design of the agents' *external* properties, the central questions are, What are the critical features in presenting an appropriate social model? What are the best design solutions for creating such a model? Three features that we believe to be pertinent are presence, gender, and communication.

A considerable number of studies have investigated whether the presence of animated pedagogical agents influences motivation. By and large, these studies reveal that the agent generally renders the students' experience more entertaining and enjoyable (e.g., Arnott, Hastings, & Allbritton, 2008; Atkinson, Mayer, & Merrill, 2005; Dunsworth & Atkinson 2007; Moreno, Mayer, Spires, & Lester, 2001; Moundridou & Virvou, 2002). Unfortunately, these studies tend to examine a large set of embodiments. Agents appear in the form of a person or an animal, and they are presented as realistic, animated, or cartoon-like. In addition, these images may display the person or animal in full, or show only the upper part of the body. Thus, some positive outcomes notwithstanding, these studies have not yet yielded unequivocal conclusions on what constitutes a motivating presence. In our study we used a photograph of a real person because we believed such a presentation to yield the most appealing and convincing model. However, it is hard to tell whether this was a good choice or whether it even mattered. The agent seemed credible, and the outcomes indicated that she affected the students.

For an agent that is presented as a person, an associated critical question concerns gender. Is it better to present a male or a female agent? An empirical study from Baylor and Kim (2004) shows that this question is not as simple as it seems. Baylor and Kim conducted two consecutive studies with animated pedagogical agents in which they varied gender, among other features. The participants in the first study were preservice teachers who had enrolled in an introductory educational technology class. They received support from an agent when learning about instructional planning. The experiment revealed that the male agent had a significantly more positive effect on students' self-efficiency beliefs than the female agent. The agents in this experiment all performed the same (unspecified) role. In the second study this role was manipulated, along with the agent's gender and ethnicity. The participants in this study were undergraduate students who had enrolled in a computer literacy course. Just as in the first study, the participants received support from an agent when learning about instructional planning. The second study yielded exactly the opposite outcome. The female agent improved students' self-efficiency beliefs significantly more than the male agent. The authors attributed this to an "overall positive student bias toward the male agents" (p. 601) in the first study, whereas in the second study preconceived notions about female agents being less knowledgeable and intelligent helped make students feel more positive about their self-efficacy. A main effect of agent role was also found. Students working with an agent as motivator, or as mentor, increased their self-efficiency beliefs more than those who had received support from the agent who performed the role of expert. What this study reveals, among its other findings, is that gender stereotypes can impact the influence of the agent, as can the agents' role.

Proper design of the agent's communication can lure the students into believing that they are reading about the feelings, thoughts, and actions of a real person of their age. To create an agent that students perceive as one of their own, it is important for the agent to communicate with the right language. Therefore, an issue that we gave special attention to in designing the agents in our study concerned the use of specific words and expressions. For this purpose a pilot was conducted in which we gathered think-aloud protocols from three participants from the target population. These protocols were analyzed, and popular youth magazines were scanned, for words and expressions that could be used by the agents. Before actually using these, the resulting vocabulary was critically evaluated by two participants from the target population.

The experimental studies from Wang et al. (2008; Wang, Lewis Johnson, Rizzo, Shaw, & Mayer, 2005) illustrate the complexity of designing the agent's communication. In these studies the agent's comments were based on a mixture of guidelines for communication (e.g., politeness norms or rules) and education (e.g., stimulate student autonomy and give performance feedback). The equivocal outcomes reported by Wang et al. indicate that considerable fine-tuning of the agent's verbal behavior may be needed to achieve robust results.

To conclude, in designing a pedagogical agent attention is required to both internal and external properties. Even when the focus lies on what the agent says or does, some attention is also needed for the agent's appearance and ways of expressing things. The design choices for these properties are complex, as many routes and options are possible. The present study contributes some insights to the ongoing efforts to create effective pedagogical agents, and to the search for validated principles for the design of such agents.

NOTES

Background. This article is partly based on the MA thesis research from Igo Weber and Marja op de Weegh.

Acknowledgments. I thank Igo Weber and Marja op de Weegh for their help in conducting this study. As so often, I am very grateful for the editorial assistance from Emily Fox.

Author's Present Address. Hans van der Meij, University of Twente, Faculty of Behavioral Science, Department of Instructional Technology (IST), P.O. Box 217, 7500 AE Enschede, the Netherlands. E-mail: H.vanderMeij@utwente.nl

HCI Editorial Record. First manuscript received January 17, 2011. Revisions received May 30, 2012, and February 14, 2013. Final manuscript received March 6, 2013. Accepted by Sharon Oviatt. — *Editor*

REFERENCES

Agrawala, M., Li, W., & Berthouzoz, F. (2011). Design principles for visual communication. *Communications of the ACM, 54*(4), 60–69.

Arnott, E., Hastings, P., & Allbritton, D. (2008). Research methods tutor: Evaluation of a dialogue-based tutoring system in the classroom. *Behavior Research Methods, 40*, 694–698.

Arroyo, I., Woolf, B. P., Cooper, D. G., Burleson, W., & Muldner, K. (2011, July). *The impact of animated pedagogical agents on girls' and boys' emotions, attitudes, behaviors and learning*. Paper presented at the International Conference of Advanced Learning Technologies, Athens, GA.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70*, 181–214.

Atkinson, R., Mayer, R., & Merrill, M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology, 30*, 117–139. doi:10.1016/j.cedpsych.2004.07.001

Atkinson, R. K., & Renkl, A. (2007). Interactive example-based learning environments: Using interactive elements to encourage effective processing of worked examples. *Educational Psychology Review, 19*, 375–386.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology, 95*, 774–783.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Bannert, M. (2000). The effects of training wheels and self-learning materials in software training. *Journal of Computer Assisted Learning, 16*, 336–346.

Baylor, A. L., & Kim, Y. (2003, November). *The role of gender and ethnicity in pedagogical agent perception*. Paper presented at the Proceedings of ELearn (World Conference on ELearning in Corporate, Government, Healthcare, & Higher Education), Phoenix, AZ.

Baylor, A. L., & Kim, Y. (2004). Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. In J. C. Lester, R. M. Vicario, & F. Paraguaçu (Eds.), *Seventh International Conference on Intelligent Tutoring Systems, ITS 2004* (pp. 592–603), Berlin, Germany: Springer Verlag.

Bhavnani, S. K., & John, B. E. (2000). Strategic use of complex computer systems. *Human–Computer Interaction, 15*, 107–137.

Bickmore, T. (2003). *Relational agents: Effecting change through human–computer relationships* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge. Retrieved from http//web.media.mit.edu/∼bickmore/bickmore-thesis.pdf

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2002). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Carroll, J. M. (1990). *The Nurnberg funnel: Designing minimalist instruction for practical computer skill*. Cambridge, MA: MIT Press.

Carroll, J. M. (Ed.). (1998). *Minimalism beyond the Nurnberg Funnel*. Cambridge, MA: MIT Press.

Carroll, J. M., & Rosson, M. B. (1987). The paradox of the active user. In J. M. Carroll (Ed.), *Interfacing thought: Cognitive aspects of human–computer interaction* (pp. 80–111). Cambridge, MA: MIT Press.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145–182.

Clarebout, G., Elen, J., Johnson, W. L., & Shaw, E. (2002). Animated pedagogical agents: An opportunity to be grasped? *Journal of Educational Multimedia and Hypermedia, 11*, 267–286.

Clore, G. L., & Palmer, J. (2009). Affective guidance of intelligent agents: How emotion controls cognition. *Cognitive Systems Research, 10*, 21–30.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Craig, S. D., Sullins, J., Witherspoon, A., & Gholson, B. (2006). The deep-level-reasoning-question effect: The role of dialogue and deep-level-reasoning questions during vicarious learning. *Cognition and Instruction, 24*, 565–591.

Dehn, D. M., & Van Mulken, S. (2000). Impact of animated interface agents: A review of empirical research. *International Journal of Human Computer Studies, 52*, 1–22.

Domagk, S. (2010). Do pedagogical agents facilitate learner motivation and learning outcomes?: The role of the appeal of agent's appearance and voice. *Journal of Media Psychology, 22*(2), 84–97. doi:10.1027/1864-1105/a000011

Duggan, M. H., & Adcock, A. B. (2007). Animated agents teaching helping skills in an online environment: A pilot study. *Journal of Interactive Online Learning, 6*, 56–71.

Dunsworth, Q., & Atkinson, R. K. (2007). Fostering multimedia learning of science: Exploring the role of an animated agent's image. *Computers and Education, 49*, 677–690.

Eccles, J. S., & Wigfield, A. (2002). Motivation beliefs, values, and goals. *Annual Review of Psychology, 53*, 109–132. doi:10.1146/annurev.psych.53.100901.135153

Feng, S. L., & Tuan, H. L. (2005). Using ARCS model to promote 11th graders' motivation and achievement in learning about acids and bases. *International Journal of Science and Mathematics Education, 3*, 463–484.

Fu, W. T., & Gray, W. D. (2004). Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks. *Cognitive Science, 28*, 901–935.

Glasbeek, H. A. (2001). "Ik doe wel wat hier staat, maar hij niet." Een onderzoek naar vaardigheidverwerving door gebruikers van leshandleidingen bij een computerprogramma ("I do what it says, but he does not." A study into skills acquisition by users of software tutorials). Doctoral dissertation, Utrecht University, the Netherlands.

Glasbeek, H. A. (2004). Solving problems on your own: How do exercises in tutorials interact with software learners' level of goal-orientedness? *IEEE Transactions on Professional Communication, 47*, 44–53.

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, and Computers, 36*, 180–192.

Gulz, A., & Haake, M. (2006). Design of animated pedagogical agents—A look at their look. *International Journal of Human Computer Studies, 64*, 322–339.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review, 6*, 27–54.

Holmes, J. (2007). Designing agents to support learning by explaining. *Computers and Education, 48*, 523–547.

Hone, K. (2006). Empathic agents to reduce user frustration: The effects of varying agent characteristics. *Interacting with Computers, 18*, 227–245.

Huett, J. B., Kalinowski, K. E., Moller, L., & Huett, K. C. (2008). Improving the motivation and retention of online students through the use of ARCS-based e-mails. *American Journal of Distance Education, 22*, 159–176. doi:10.1080/08923640802224451

Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development, 10*, 2–10.

Keller, J. M. (1999). Using the ARCS motivational process in computer-based instruction and distance education. *New Directions for Teaching and Learning, 78*, 37–47.

Keller, J. M. (2010). *Motivational design for learning and performance. The ARCS model approach*. New York, NY: Springer Verlag.

Keller, J. M., & Kopp, T. W. (1987). An application of the ARCS model of motivational design. In C. M. Reigeluth (Ed.), *Instructional theories in action: Lessons illustrating selected theories and models* (pp. 289–320). Hillsdale, NJ: Erlbaum.

Keller, J. M., & Suzuki, K. (2004). Learner motivation and E-learning design: A multinationally validated process. *Journal of Educational Media, 29*, 229–239. doi:10.1080/1358165042000283084

Kim, Y., Baylor, A. L., & PALS Group. (2006). Pedagogical agents as learning companions: The role of agent competency and type of interaction. *Educational Technology Research and Development, 54*, 223–243.

Kim, Y., Baylor, A. L., & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning, 23*, 220–234.

Klein, J., Moon, Y., & Picard, R. W. (2002). This computer responds to user frustration: Theory, design and results. *Interacting with Computers, 14*, 119–140.

Lazonder, A. W., & Van der Meij, H. (1995). Error-information in tutorial documentation: Supporting users' errors to facilitate initial skill learning. *International Journal of Human Computer Studies, 42*, 185–206.

Leutner, D. (2000). Double-fading support—A training approach to complex software systems. *Journal of Computer Assisted Learning, 16*, 347–357.

Loorbach, N., Karreman, J., & Steehouder, M. (2007). Adding motivational elements to an instruction manual for seniors: Effects on usability and motivation. *Technical Communication, 54*, 343–358.

Loorbach, N., Steehouder, M., & Taal, E. (2006). The effects of motivational elements in user instructions. *Journal of Business and Technical Communication, 20*, 177–199.

Lusk, M. M., & Atkinson, R. K. (2007). Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples? *Applied Cognitive Psychology, 21*, 747–764.

Marsella, S. C., Johnson, W. L., & LaBore, C. (2000, June). *Interactive pedagogical drama*. Paper presented at the 11th International Conference on Artificial Intelligence in Education, Barcelona, Spain.

Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E. (2005). *Cambridge handbook on multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology, 95*, 806–812. doi:10.1037/0022-0663.95.4.806

Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology, 96*, 389–395.

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied, 11*, 256–265.

McQuiggan, S. W., & Lester, J. C. (2007). Modeling and evaluating empathy in embodied companion agents. *International Journal of Human Computer Studies, 65*, 348–360.

Moreno, R. (2005). Multimedia learning with animated pedagogical agents. In R. Mayer (Ed.), *Cambridge handbook on multimedia learning* (pp. 507–524). New York, NY: Cambridge University Press.

Moreno, R., & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology, 31*, 186–207.

Moreno, R., & Mayer, R. E. (2000). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology, 92*, 724–733.

Moreno, R., & Mayer, R. E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology, 96*, 165–173. doi:10.1037/0022-0663.96.1.165

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 192*, 177–213.

Moundridou, M., & Virvou, M. (2002). Evaluating the persona effect of an interface agent in a tutoring system. *Journal of Computer Assisted Learning, 18*, 253–261.

Norman, D. A. (1994). How might people interact with agents. In *Communications of the ACM* (Vol. 37, pp. 68–71). New York, NY: ACM.

Ortony, A., Clore, G. L., & Foss, M. A. (1987). The referential structure of the affective lexicon. *Cognitive Science, 11*, 341–364.

Paas, F. G., Van Merriënboer, J. J., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*(1 Pt. 2), 419–430.

Paiva, A., & Machado, I. (2002). Lifelong training with Vincent, a web-based pedagogical agent. *Journal of Continuing Engineering Education and Life Long Learning, 12*, 254–266.

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education. Theory, research and applications* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Read, J. C. (2008). Validating the Fun Toolkit: An instrument for measuring children's opinions of technology. *Cognition, Technology and Work, 10*, 119–128.

Reimann, P., & Neubert, C. (2000). The role of self-explanation in learning to use a spreadsheet through examples. *Journal of Computer Assisted Learning, 16*, 316–325.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*, 1–29.

Rickenberg, R., & Reeves, B. (2000, April). *The effects of animated characters on anxiety, task performance, and evaluations of user interfaces*. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, the Hague, the Netherlands.

Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A., & Doerr, C. E. (2008). Interface agents as social models for female students: The effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers in Human Behavior, 24*, 2741–2756.

Ryokai, K., Vaucelle, C., & Cassell, J. (2003). Virtual peers as partners in storytelling and literacy learning. *Journal of Computer Assisted Learning, 19*, 195–208.

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2012, April). *A meta-analysis of the effects of pedagogical agents on learning.* Paper presented at the American Educational Research Association, Vancouver, Canada.

Schworm, S., & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology, 99*, 285–296.

Song, S. H., & Keller, J. M. (2001). Effectiveness of motivationally adaptive computer-assisted instruction on the dynamic aspects of motivation. *Educational Technology Research and Development, 49*(2), 5–22.

Tartaro, A., & Cassell, J. (2007). Authorable virtual peers for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 26*, 169–172.

Van der Meij, H. (2003). Minimalism revisited. *Document Design, 4*, 212–233.

Van der Meij, H. (2008). Designing for user cognition and affect in software instructions. *Learning and Instruction, 18*, 18–29. doi:10.1016/j.learninstruc.2006.08.002

Van der Meij, H., Blijleven, P., & Jansen, L. (2003). What makes up a procedure? In M. J. Albers & B. Mazur (Eds.), *Content & complexity. Information design in technical communication* (pp. 129–286). Mahwah, NJ: Erlbaum.

Van der Meij, H., & Carroll, J. M. (1998). Principles and heuristics for designing minimalist instruction. In J. M. Carroll (Ed.), *Minimalism beyond the Nurnberg funnel* (pp. 19–53). Cambridge, MA: MIT Press.

Van der Meij, H., & Gellevij, M. (2004). The four components of a procedure. *IEEE Transactions on Professional Communication, 47*, 5–14.

Van Loggem, B. (2007). *Paper-based support for computer-mediated activity: A cognitive task-analysis* (Unpublished master's thesis). University of Twente, Enschede, the Netherlands.

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*, 147–177.

Van Mulken, S., André, E., & Müller, J. (1998, September). *The persona effect: How substantial is it?* Paper presented at the People and Computers XIII: HCI'98, Sheffield, UK.

Vollmeyer, R., & Rheinberg, F. (1999). Motivation and metacognition when learning a complex system. *European Journal of Psychology of Education, 14*, 541–554. doi:10.1007/bf03172978

Vollmeyer, R., & Rheinberg, F. (2006). Motivational effects on self-regulated learning with different tasks. *Educational Psychology Review, 18*, 239–253.

Wang, N., Lewis Johnson, W., Mayer, R. E., Rizzo, P., Shaw, E., & Collins, H. (2008). The politeness effect: Pedagogical agents and learning outcomes. *International Journal of Human Computer Studies, 66*, 98–112.

Wang, N., Lewis Johnson, W., Rizzo, P., Shaw, E., & Mayer, R. E. (2005, January). *Experimental evaluation of polite interaction tactics for pedagogical agents.* Paper presented at the International Conference on Intelligent user Interfaces, San Diego, CA.

Wiedenbeck, S., Zavala, J. A., & Nawyn, J. (2000). An activity-based analysis of hands-on practice methods. *Journal of Computer Assisted Learning, 16*, 358–365. doi:10.1046/j.1365-2729.2000.00148.x

Wiedenbeck, S., Zila, J. A., & McConnell, D. S. (1995, May). *End-user training: An empirical study comparing on-line practice methods.* Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, Denver, CO.

Woolf, B. P., Arroyo, I., Muldner, K., Burleson, W., Cooper, D., Dolan, R., & Christopherson, R. M. (2010, June). *The effect of motivational learning companions on low-achieving students and students with learning disabilities.* Paper presented at the International Conference on Intelligent Tutoring Systems, Pittsburgh, PA.