

## Other Articles

# SCRIPTKELL: A tool for measuring cognitive effort and time processing in writing and other complex cognitive activities

ANNIE PIOLAT, THIERRY OLIVE, JEAN-YVES ROUSSEY,  
OLIVIER THUNIN, and JOHANNES C. ZIEGLER

*Centre PsyCLE and Université de Provence, Aix-en-Provence, France*

We present SCRIPTKELL, a computer-assisted experimental tool that makes it possible to measure the time and cognitive effort allocated to the subprocesses of writing and other cognitive activities. SCRIPTKELL was designed to easily use and modulate Kellogg's (1986) triple-task procedure, which consists of a combination of three tasks: a writing task (or another task), a reaction time task (auditory signal detection), and a directed retrospection task (after each signal detection during writing). We demonstrate how this tool can be used to address several novel empirical and theoretical issues. In sum, SCRIPTKELL should facilitate the flexible realization of experimental designs and the investigation of critical issues concerning the functional characteristics of complex cognitive activities.

To capture on-line covert processes involved in reading, writing, learning, problem solving, reasoning, and decision making, investigators often use concurrent thinking-aloud protocols and/or retrospective reports (e.g., Ericsson & Simon, 1993; Smagorinsky, 1994). In addition, the amount of cognitive effort associated with these mental activities can be evaluated by using reaction time (RT) tasks (Fisk, Derrick, & Schneider, 1986-1987; Tyler, Hertel, McCallum, & Ellis, 1979). As concerns research on writing, such methods helped researchers to identify a number of basic subprocesses involved in this activity, such as planning, translating, and reviewing. However, it still remains an important research goal to measure the degree of effort and mental load associated with each subprocess and the way in which writers successively engage in these different subprocesses (time processing).

For this purpose, Levy and Ransdell (1994) proposed a methodology and computer software that allowed the investigation of on-line processes involved in writing with a word processor. Their software makes it possible to relate an RT measure to different writing subprocesses (captured with concurrent verbal protocols). Thus, Levy and Ransdell (1994) provided a powerful tool for com-

binning verbal and think-aloud protocols "to better understand the time course and effort of writing processes itself" (p. 222). In addition, their methodology allows researchers to evaluate the extent to which other concurrent tasks have an effect on writing (e.g., irrelevant speech, a visuospatial task; for a review, see Levy, 1997). Finally, their software gives researchers information about the speed of typing as well as the frequency, duration, and location of pauses (Janssen, van Waes, & van den Bergh, 1996; Ransdell, 1990). In fact, their methodology has proven to be extremely useful and efficient for the study of writing. However, it is limited to the study of writing, and it cannot be easily extended to the study of other cognitive activities mentioned above.

Our software, SCRIPTKELL, 1 was designed to measure the time and cognitive effort allocated to the subprocesses of writing or other cognitive activities. In particular, SCRIPTKELL "implements" Kellogg's (1988) triple-task procedure, which consists of a combination of three tasks: a writing task (or any other task), an RT task (auditory signal detection), and a directed retrospection task (after each signal detection during writing). SCRIPTKELL differs from the above described tools in three ways: (1) Because it is not necessarily connected to a word processor, its use is not restricted to the analysis of writing. Instead, it can also be used to study cognitive effort and time processing in other complex cognitive activities, such as reading, learning, problem solving, and decision making, (2) Because SCRIPTKELL does not rely on using a computer keyboard, our program can be used with populations that have little or no experience with typing on a computer keyboard (e.g., children or older participants), (3) The parameters of both added tasks (i.e., signal de-

The research reported in this article was partially supported by a Postdoctoral Fellowship of the Fondation Fyssen to J. C. Ziegler and a Doctoral Fellowship of the Ministère de l'Enseignement Supérieur et de la Recherche to T. Olive. The authors would like to thank R. T. Kellogg, D. Janssen, P. Lemaire, J. Vauclair, and an anonymous reviewer for their comments on the manuscript. Correspondence should be addressed to A. Piolat, Centre PsyCLE, Université de Provence, 29 Av., Robert Schuman, F-13621 Aix-en-Provence, Cedex, France ([annie.piolat@up.univ-mrs.fr](mailto:annie.piolat@up.univ-mrs.fr))

tection and directed retrospection) can be easily modified as a function of the questions and problems to be studied. In the following section, we describe in greater detail Kellogg's (1988) triple-task procedure.

### KELLOGG'S TRIPLE-TASK PROCEDURE

Kellogg (1986, 1987a, 1988) proposed an experimental procedure in which writers were asked to perform three tasks: a primary task that required the composition of a text, an RT task that required the detection of auditory signals, and a directed retrospection. In the context of writing, Kellogg's triple-task procedure proceeds in the following way: During the primary task (text composition), participants hear an auditory signal about every 30 sec. They are asked to say "stop" as quickly as possible whenever they detect such a signal (RT task). After each signal detection, they are requested to perform a directed retrospection (i.e., to identify the mental process in which they were engaged when they perceived the auditory signal). For this purpose, they choose among four response categories (planning, translating, reviewing, or other) by selecting one of four response keys on the computer keyboard. The time it takes to detect the auditory signal and to say "stop" can be taken as a measure of the degree of effort associated with the writing process that was interrupted by the signal. Note that the final RT measure is computed by subtracting the participant's RTs from the mean baseline RTs that were measured for each participant in a session in which the participant performed the auditory signal detection task only.

The directed retrospection makes it possible to find out which mental process was interrupted by the auditory signal. That is, the writer could have been planning, translating, reviewing, or thinking of something completely different (other) when the auditory signal occurred. Thus, by analyzing the mean frequency of each mental process, the retrospective task aims at describing time processing. As concerns the validity of the retrospection task, one should note that participants are trained to perform categorizations in this task before an experimental session. Kellogg (1987b, pp. 277-278) provided some empirical evidence on the validity of these categorizations. As concerns the important critique that retrospections are susceptible to omissions and reconstruction (e.g., Levy, Marek, & Lea, 1996), one should, however, note that each directed retrospection happens, on average, 2 or 3 sec after the writing was interrupted by the auditory signal. Therefore, the retrospection task differs from a retrospective report that would only occur once the whole task has been completed. Thus, some of the shortcomings of retrospective protocols, such as omission and reconstruction, are less likely to affect the present version of the retrospection task.

Using this triple-task procedure, Kellogg (1986, 1994) investigated the degree of cognitive effort allocated to different subprocesses as a function of the primary task (e.g., writing, reading, chess playing, learning), the type of text planning (e.g., rough or elaborated first draft), the topic (familiar or unfamiliar content), the type of text (e.g.,

story, description, argumentation), and, finally, the writing tool (e.g., paper/pencil or word processor). A comprehensive review of his results can be found in Kellogg (1994).

In our studies, we extended Kellogg's research efforts. First, we investigated the extent to which certain parameters (e.g., interval between auditory signals, training of the retrospection task) have an effect on cognitive load and time processing (Olive, 1997; Piolat, Farioli, Olive, & Roussey, 1996; Piolat, Roussey, Olive, & Farioli, 1996). Second, similarly to Kellogg, we evaluated the extent to which participant-specific or situation-specific factors affect writing (e.g., the role of knowledge, writing expertise, writing tool; Barbier, Piolat, & Roussey, 1998; Olive & Piolat, 1997; Olive, Piolat, & Polge, 1997; Olive, Piolat, & Roussey, 1996, 1997; Piolat, Barbier, & Roussey, 1996; Piolat, Roussey, & Rous, 1996). After our first experiments conducted with a prototypical software, we decided to build a versatile software for future experiments within this methodology.

### Empirical Issues

In our empirical studies mentioned above, we used a simple prototypical version of SCRIPTKELL. Some of the critical issues and software limitations observed in these studies led us to develop the present software, which allows researchers to easily modulate Kellogg's procedure by specifying the essential parameters and the global configuration of the secondary tasks (e.g., duration of the training phase, interval between auditory signals, number and nature of response categories). SCRIPTKELL should therefore facilitate the realization of experimental designs that may allow us to address new questions concerning functional characteristics of writing and other complex cognitive activities.

In the following sections, we describe some of our own experiments that focused on four crucial aspects of Kellogg's original triple-task procedure. The first aspect concerns the global design of the triple-task procedure, the second aspect concerns the training of the directed retrospection, the third aspect concerns the design of the directed retrospection, and the fourth aspect concerns the design of the RT task. We show how SCRIPTKELL can be used to address these issues.

### Design of the Triple-Task Procedure

The secondary-task RT technique (Kahneman, 1973; Kerr, 1973; Power, 1986) is based on the principle that participants have to simultaneously perform two tasks: a primary task and a secondary task. Performance in one task is typically in conflict with performance in the other task. As a consequence, performance in the secondary task or performance in both the primary task and the secondary task will decrease. It is assumed that the secondary-task RT technique works because the capacity of the working memory is limited and available cognitive resources are difficult to distribute across different tasks.

Kellogg extended the secondary-task RT technique by adding the directed retrospection. In fact, the strength of Kellogg's approach results from the combination of three tasks: a writing task, an RT task (detection of an auditory

signal), and the directed retrospection. This experimental procedure makes it possible to measure the cognitive effort related to each subprocess involved in text composition (e.g., planning, translating, reviewing) and to measure the frequency with which each subprocess is used. Note that Kellogg's technique "samples" the processes in question only about every 30 sec, whereas the procedure by Levy and Ransdell (1994, 1995, 1996) allows a much finer temporal resolution (e.g., every 1 sec). However, the temporal resolution of Kellogg's procedure was highly sufficient for our purpose.

One should note that Kellogg (1994) does not argue in terms of competition between the primary and secondary tasks in the sense that the "spare capacity not required by the primary task such as writing will be directed to the secondary reaction time (RT) task" (p. 54). Furthermore, Kellogg showed that performing a secondary task, such as directed retrospection, does not affect the writers' efficiency or the quality of their text composition (Kellogg, 1987a, p. 266, note 5; Kellogg, 1987b, pp. 277-278). However, Kellogg did not explicitly compare the RTs and quality of the written composition in a situation in which writers perform only the primary task, as opposed to a situation in which writers perform the primary and the secondary tasks.

To address these issues, we included an option in SCRIPTKELL that allows one to choose between a dual-task situation and a triple-task situation. The dual-task situation consists of either the writing task plus the RT task or the writing task plus the directed retrospection. The triple-task situation includes all three tasks. This option was used to test the validity of Kellogg's procedure by comparing the RTs of a group of 20 participants who performed the triple-task procedure with those of a group who performed the writing and the RT tasks only (Piolat, Farioli, et al., 1996; Piolat, Roussey, & Rous, 1996). Theme of the text and experimental conditions (e.g., number of the auditory signals, interval between auditory signals, instructions) were similar to those of Kellogg (1988). Results showed that there was no significant difference between the two groups in terms of cognitive effort [triple task, mean RT = 393 msec; double task, mean RT = 350 msec;  $F(1,33) = 1.109$ , n.s.; cf. Piolat, Roussey, & Rous, 1996, p. 345]. Therefore, at this point, we have no empirical evidence that performing the directed retrospection immediately after the RT task influences the mental load associated with the writing task.

In a second line of research, Barbier et al. (1998) and Piolat, Barbier, & Roussey (1996) observed that high-school students (about 15 years old) were slower in the RT task when writing in a foreign language than when writing in their native language. Therefore, Kellogg's procedure seems to be sensitive to measure variations in cognitive effort when participants are asked to write in a language other than their native language. In addition, we noticed that students, like many adults in previous research, experienced some difficulties for labeling their mental processes in the directed retrospection (they fre-

quently used the label *other*). The exaggerated use of this category seems to reflect participants' difficulty in identifying the specific process they were engaged in when being interrupted by the auditory signal. Using our option of a dual-task procedure (writing task plus RT task) would be an appropriate way to investigate whether the language-dependent RT differences mentioned above persist when participants do not have to perform the apparently more difficult directed retrospection.

In sum, it seems necessary that researchers who use the triple-task procedure have the possibility to control for the effect of each additional task. We are currently investigating the impact of the RT task on the writing process by comparing a triple-task situation with a dual-task situation that only includes the writing task and the directed retrospection (i.e., when the beep is given, the participant does not perform the speeded auditory detection task but simply selects a key corresponding to the interrupted process before going back to the primary writing task). Finally, SCRIPTKELL may provide an appropriate tool for studying the acquisition of expertise of novice writers for whom the conscious collection of the mental processes presents a major challenge (Scardamalia & Bereiter, 1991).

#### Training of the Directed Retrospection

Kellogg (1986, 1987a, 1988, 1994) gives little information about how writers were trained to perform the directed retrospection task. Typically, the training phase lasts about 20 min and contains

- (a) instructions that define and give examples of the writing processes, (b) sample situations where a fictitious businessman thinks aloud while writing letters and his thoughts (in written form) are categorized by the experimenter as planning, translating, reviewing, or unrelated, and (c) a test situation where the participant makes such categorizations and is provided feedback from the experimenter. The unrelated category was for all thoughts that did not fit planning, translating, or reviewing (e.g., day-dreaming about school). (Kellogg, 1988, p. 356)

Apparently, Kellogg's (1987b, 1994) adult participants had no difficulty categorizing a set of think-aloud protocols and performing a directed retrospection. Since our writers frequently showed some difficulties in labeling their mental activity in the directed retrospection, we investigated the role of training for successful performance in this task. Two groups of 20 participants were tested. The first group received Kellogg's original training. The second group received two training sessions. The first session was a group discussion (lasting 3 h) in which participants were confronted with their usual way of writing and learned to label their writing processes according to Kellogg's categories. In the second session, 1 week later, they obtained Kellogg's original training (Piolat, Farioli, et al., 1996; Piolat, Roussey, & Rous, 1996).

The results showed that each group significantly differed in the use of different writing subprocesses [Kellogg's training group, planning = 29%, translating = 45%,

reviewing = 19%,  $F(2,36) = 11.6, p < .001$ ; high-training group, planning = 25%, translating = 46%, reviewing = 22%,  $F(2,36) = 16.293, p < .0001$ ; cf. Piolat, Roussey, & Rous, 1996, p. 346]. However, the two groups did not significantly differ in the use of the different writing subprocesses. However, writers who received the more elaborated training were slower in the RT task (624 msec) than were writers who were trained according to Kellogg's original procedure [393 msec;  $F(1,33) = 12.86, p < .002$ ; cf. Piolat, Roussey, & Rous, 1996, p. 345]. This result suggests that the more elaborated training increases participants' effort in writing, which translates into longer RTs for this group. It is also interesting to look at a second RT measure that is automatically recorded by SCRIPT KELL: the time it takes for participants to choose among the response categories (writing, planning, reviewing, or other). Although there is some systematic change in this variable, Kellogg never considered this potential source of information. In our study, writers who received the more elaborated training were significantly faster in choosing the labels in the directed retrospection than were writers who were less extensively trained [1.2 vs. 1.85 sec;  $F(1,33) = 16.09, p < .0003$ ]. This result is important because it suggests that focusing on a particular task (here, the directed retrospection with increased training) may change the cognitive effort devoted to the other tasks (for similar objections, see Fisk et al., 1986-1987). In such cases, the participant may have only a small margin in allocating his or her resources to each of the three tasks and does so on the basis of a subjective evaluation of the experimenter's demands (cf. the notion of "cognitive balance" in Glynn, Britton, Muth, & Dogan, 1982).

#### Design of the Directed Retrospection

Concurrent and retrospective protocols have provided a major tool for uncovering processes in complex cognitive activities. The debate over the relative advantages and inconveniences of each type of protocol is far from being resolved (cf. Ericsson & Simon, 1993; Smagorinsky, 1994). Recent work (Janssen et al., 1996; Levy et al., 1996) has criticized the directed retrospection method for being more intrusive and reactive than concurrent think-aloud protocols and for being limited to the a priori theoretically defined categories. However, it should be acknowledged that, within the framework of the triple-task procedure, retrospection is performed very quickly (1 or 2 sec after the speeded auditory signal detection by selecting and pressing one of the responses keys). Therefore, it can be assumed that retrospection occurs during writing and not once the writing process is completed. It therefore seems that forgetting and reconstruction are less likely to bias the retrospection in the present setup.

Although we agree with some of the criticism concerning retrospective methods, two points should be made in their defense. First, even if the retrospection is subject to forgetting and omission, this would still affect the different experimental conditions in very much the same way. However, because we are interested in the role of specific

experimental manipulations, a systematic bias that affects all experimental conditions is less worrisome. Second, being less time consuming and requiring less technical equipment, the directed retrospection is probably more economical than Levy and Ransdell's technique.

An important design feature of the present implementation of the retrospection task is the possibility to modify the number and the content of response categories depending on the cognitive activity under investigation and the particular research question. In fact, some studies used up to 15 response categories (Levy et al., 1996; Schumacher, Klare, Cronin, & Moses, 1984). Moreover, to support the development of more fine-grained theoretical descriptions of the processes involved in writing (e.g., Brown, McDonald, Brown, & Carr, 1988; Hayes, 1996; Kellogg, 1996), researchers should not be restricted to using only the same three or four response categories as in Kellogg's original retrospection. In the following sections, we give some examples about how changing some characteristics of the directed retrospection method allows researchers to address a number of critical issues.

In an unpublished study (Piolat & Roussey, 1993), writers were asked to correct an unfinished text that contained either local surface errors or more global content errors. Before the experimental session, participants were trained to perform a directed retrospection using six labels (*reading to comprehend text*, *reading for problem detection*, *detecting a problem*, *diagnosing a problem*, *searching for a correction*, and *revising*). These labels characterize the revision, process according to Hayes, Flower, Schriver, Stratman, and Carey (1987). Thus, the number and the nature of response categories were adapted to match the primary task (i.e., improving the surface or the semantic coherence of a text). The results showed that the cognitive effort allocated to improving the semantic coherence of the text was larger than the one allocated for improving the surface of the text [respectively, 1,097 and 751 msec;  $F(1,36) = 7.4, p < .01$ ]. In addition, the processes related to text comprehension (reading for comprehension, reading for problem detection, detecting and diagnosing a problem) were more demanding than those related to production (searching and making the correction). Moreover, writers engaged more frequently in comprehension than in production processes. It seems that investigations of this kind can greatly improve our understanding of the processes involved in writing.

In yet another line of research (Barbier et al., 1998; Piolat, Barbier, & Roussey, 1996), students composed text in either their native or a foreign language. After each detection of an auditory signal, they were asked to label their mental activity by choosing one of six response categories: (1) *I was paying attention to spelling*, (2) *I was paying attention to word choice*, (3) *I was paying attention to grammar*, (4) *I was paying attention to ideas*, (5) *I was paying attention to text organization*, or (6) *I was paying attention to other things*. Although the students had a difficult time performing the directed retrospection (see above), results showed that the label *I was paying attention*

to word choice was more frequently chosen in the foreign language than in the native language. In contrast, the option 1 was paving attention to ideas was more frequently chosen in the native language than in the foreign language. These results indicate that language expertise and familiarity have a direct effect on the extent to which various writing processes are used. As in the research described above, the possibility of modifying the number and nature of response categories seems to open interesting venues for future research.

A final issue concerning the retrospection task is related to the questionable assumption according to which the different writing subprocesses are activated in a serial and recursive way. Hayes and Flower (1980) argued that writers engage recursively in processes, such as planning, translating, and reviewing. Research on patterns of transitional probability matrices, also referred to as *writing signatures*, can also be taken as further evidence for the recursive nature of writing processes (see Levy, 1997; Levy & Ransdell, 1994). Kellogg (1987a, 1988) indicated in the description of his experimental procedure that writers should choose, among all mental processes, the one that appears the most dominant at the time the auditory signal occurs. Kellogg (1994, 1996) admitted that some processes may be fairly complex and difficult to classify into only three categories. This complexity may stem from writers' capacity to engage in several subprocesses in parallel. For example, Levy et al. (1996, p. 548) showed that 8% of concomitant protocols concerned at least two processes. SCRIPTKELL allows researchers to test some of these issues by providing them with the possibility to freely change and extend the number and the nature of the retrospective options.

#### Design of the RT Task

A further constraint of Kellogg's procedure is the choice of the time interval with which the auditory signal interrupts the writing process. Kellogg (1987a, 1988) did not justify the choice of a particular interval. For example, in Kellogg's (1988) Experiment 1, the interval between two auditory signals, in which the participant was also asked to make the retrospection, was 60 sec. However, in Experiment 2 of the same study, it was reduced to an average of 30 sec (ranging from 15 to 45 sec to avoid an anticipation of the auditory signal).

In recent studies (Olive, 1997; Piolat, Farioli, et al., 1996; Piolat, Roussey, & Rous, 1996), we addressed the question of whether the choice of a particular interruption interval affects the cognitive effort and the activation of different subprocesses involved in text composition. The design included three groups: (1) short interval (auditory signal every 15 sec varying between 10 and 20 sec), (2) average interval (every 30 sec varying between 15 and 45 sec), and (3) long interval (every 60 sec varying between 45 and 75 sec). Note that the average interval was the one used in the studies by Kellogg.

Results showed that the use of different writing subprocesses was not significantly affected by variations of

the interruption interval (slow cadence, planning = 33%, translating = 41%, reviewing = 20%; Kellogg's cadence, planning = 29%, translating = 45%, reviewing = 19%; fast cadence, planning = 36%, translating = 43%, reviewing = 17%; cf. Piolat, Roussey, Olive, & Farioli, 1996, p. 345). In contrast, RTs in the detection task were significantly longer in both the short-interval group and the long-interval group than in the average-interval group [slow cadence, 525 msec; Kellogg's cadence, 393 msec; fast cadence, 502 msec;  $F(2,54) = 3.98, p < .025$ ]. In the long-interval group, a rating after the actual experiment also revealed three interesting patterns: (1) Writers in this group were little disturbed by the auditory signal; (2) they remained focused on the primary task; and (3) they reacted more slowly to the auditory signal. There seemed to be little competition between the primary and secondary tasks. However, in the short-interval group, writers indicated that they were very disturbed by the frequent interruptions and the competition between the primary and secondary tasks. They seemed to neglect the secondary task (long RTs) in favor of the primary writing task. Although ratings in the average-interval condition were almost identical to those in the short-interval condition, RTs in the secondary task were shortest in the average-interval condition. It seems that, in the average-interval condition, writers were able to perform simultaneously and efficiently both the primary task and the secondary task. Note that the three experimental groups did not significantly differ in terms of the quality (in the sense of Kellogg) of the written composition. Thus, it seems that varying the interval between auditory signals affects the distribution of attentional resources across the different tasks.

Without interpreting these results any further (cf. Olive, 1997; Piolat, Farioli, et al., 1996; Piolat, Roussey, & Rous, 1996), it seems that the choice of the interruption interval provides a major constraint of Kellogg's procedure. It is clear that the choice of the interruption interval may affect writers' performance as a function of writing expertise: The younger or less experienced the writer, the higher the probability that critical interruption intervals produce what is often referred to as the *cognitive overload* during writing (Glynn et al., 1982; Hayes & Flower, 1980). Together, the above-mentioned studies underline the importance of being able to freely modify the interruption interval in SCRIPTKELL.

#### HOW TO USE SCRIPTKELL

The program SCRIPTKELL runs on a Macintosh (system 7.0 or higher). It is programmed in HyperTalk (English version) in a HyperCard 2.1 environment. The SCRIPTKELL folder contains the program and two additional folders: (1) Parameters and (2) Experimental Data.

#### Configuration of the Experimental Procedure

When SCRIPTKELL is started, the program displays a window that allows the experimenter to specify the entire configuration (e.g., tasks, parameters, response keys, au-

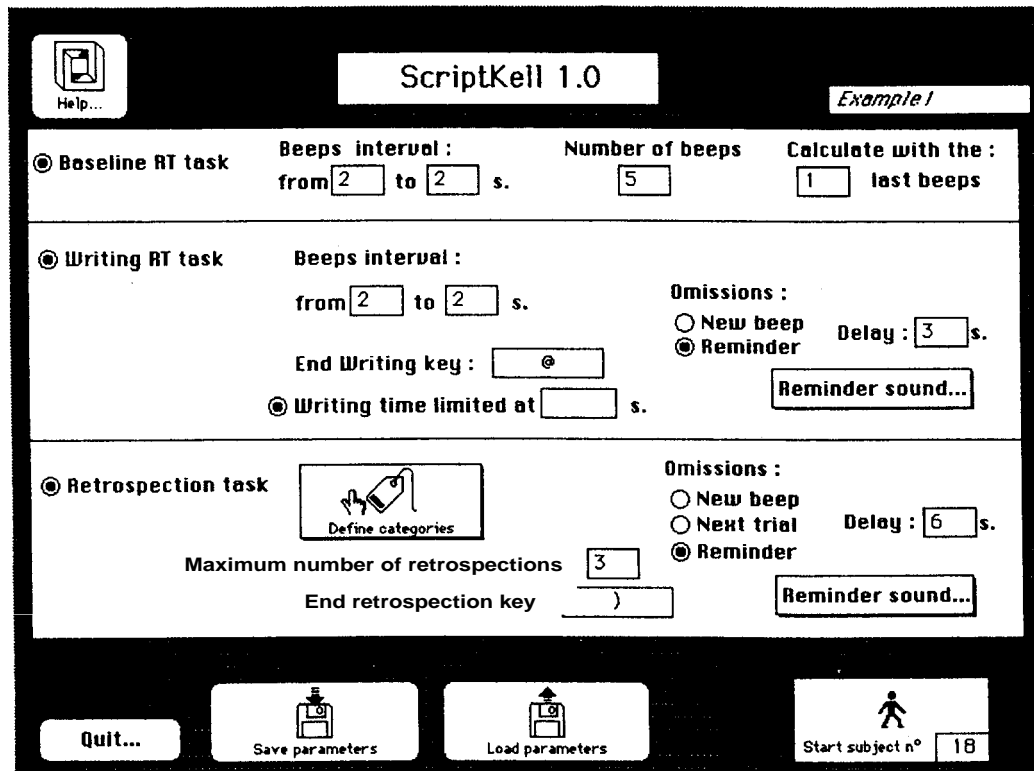


Figure 1. Main window of SCRIPTKELL that allows the experimenter to specify the global configuration of an experiment.

ditary signals, writing time) for any given experiment. By default, the program gives the values of the parameters that had been used in the previous session. This configuration window is presented in Figure 1.

As can be seen in Figure 1, the configuration window allows the experimenter to choose and specify the secondary tasks, the baseline RT task, the writing RT task, and the retrospection. In the baseline RT task, participants perform the auditory signal detection only. In the writing RT task, they perform the auditory signal detection during a writing session. In the retrospection, they are asked to label the subprocesses they engaged in when being interrupted. The selected parameters are saved in a file that is located in the Parameters folder. Parameters can be reactivated and reused at any time (see Save and Load parameters). In the upper left corner of the configuration window, the program displays a HELP function that contains a detailed description of all possible choices. In the upper right corner, the program displays the selected experimental procedure. In the lower right corner, the experiment can be started after having specified the participant's identification number. The following sections describe the major choices.

**Baseline RT task.** In the baseline RT task, participants are presented only with a number of auditory signals (beeps) that they have to detect as rapidly as possible. This task measures the RT baseline of each participant against

which the RTs in the double- or triple-task procedure can be compared. As can be seen in Figure 1, the interval between beeps needs to be specified. For this purpose, the experimenter enters the longest interval and the shortest interval between two beeps (Kellogg, 1987a, 1988, often used an interval between 5 and 15 sec). Within this interval, the program distributes the beeps in a pseudorandom fashion. The experimenter also needs to specify the total number of beeps (Kellogg typically used 30) and the number of final beeps that is used to calculate the baseline RT (Kellogg typically used the latest 25). Speeded responses are given by pushing the button of the computer mouse. Us are measured to the nearest millisecond.

**Writing RT task.** Two options are available for determining writing time. Either the experimenter selects the option *Limited writing time* and then specifies the duration of the writing task or the experimenter selects the option *Free writing time*. In this case, the participants can compose as long as they want. In both cases, the experimenter also defines a key with which participants terminate the program when they have completed the primary writing task ("@" in Figure 1).

Participants sometimes fail to detect a beep (omissions). To handle such omissions, the experimenter can activate one of two options: *New beep* or *Reminder*. If the option *New beep* is selected, the experimenter chooses a delay interval (in seconds), after which the program emits a

Categories (15 car. max)	Associated keys
1 Activity "a"	q
2 Activity "b"	w
3 Activity "c"	e
4 Other	r

Number of categories :  (max. 12)

Careful: é=è=e=E

OK

Figure 2. Window that allows the experimenter to specify the number, labels, and response categories of the retrospection task.

new beep in case the participant fails to respond. In this case, the RT corresponding to the new beep, but not the one corresponding to the omission, is taken into consideration. If the option *Reminder* is selected, the experimenter also chooses a delay. If the participant has not reacted until the delay has elapsed, the program emits a reminder (warning beep), and the actual trial continues (no new beep is emitted). The warning beep itself can be modified by clicking on the option *Reminder sound*.

**Directed retrospection.** In the directed retrospection, participants are asked to choose one or several categories that best reflect what they were doing when being interrupted by the beep. By clicking on the option *Define labels*, the experimenter has access to a second window that is presented in Figure 2.

In this window, the experimenter specifies the number of response categories (maximum of 12). In addition, the response categories are labeled (e.g., *planning, translating, revising, other*) and specific keys are assigned to each category.

Back in the configuration window (Figure 1), the experimenter specifies the key with which the participant terminates the retrospection. If participants can choose more than one response category at any given trial, then the experimenter indicates the number of maximum re-

sponses in the corresponding field. In the case of multiple response choices, participants push the *End* key to indicate that they finished the retrospection and that they continue composing.

Occasionally, participants may forget to perform the directed retrospection. For such cases, the program offers several options: (1) If the option *New beep* is activated and participants fail to perform the retrospection before a deadline (specified by the experimenter) has elapsed, then the program initiates a new beep/retrospection trial. The trial that includes the omitted retrospection is not taken into consideration. (2) If the option *Next trial* is activated and participants fail to perform the retrospection before the deadline, then the program initiates the next beep/retrospection trial. The incomplete trial is neither repeated nor taken into consideration. (3) If the option *Reminder* is activated, the program emits a warning beep, and the actual trial continues.

#### Data Recording and Analysis

SCRIPTKELL automatically records and analyzes the different dependent variables of Kellogg's procedure (e.g., number of reactions, frequency of category choices, mean RTs, mean difference scores). These data are saved in a text file that is located in the Experimental Data folder.

When this text file is opened with Microsoft Excel, the results are organized into different cells of a spreadsheet with corresponding mean values and standard deviations.

In the first section of the spreadsheet, summary information about the participant and the experiment is given, such as the participant's name and identification and the date and time at which the experiment started. In the second section, the parameters and the results of the baseline task are given (i.e., the interval between beeps, the number of beeps that was used for the computation of the RTs, individual RTs, and the mean baseline RT). The third section deals with the overall results. The program gives the relevant parameters and the total number of responses. All responses are then listed in chronological order along with their corresponding measures: the time since the beginning of the composition, the time since the last beep, the RT corresponding to the auditory signal detection, the difference score (each RT minus the mean baseline RT), the subprocess identified during the directed retrospection task, and the designation time. In addition, mean RTs, mean difference scores, and mean designation times are given along with their standard deviations. In a fourth section, the data are broken down as a function of the categories that were selected in the retrospection (i.e., as a function of the identified subprocesses). Finally, in order to investigate changes in performance over time, the total writing time is broken down into three and four equally long parts, and all analyses are provided separately for each part.

## CONCLUSIONS

SCRIPTKELL is a computer-assisted tool that allows researchers to efficiently use Kellogg's triple-task procedure to measure the time and effort allocated to the subprocesses of writing other complex cognitive activities. As we have shown, some of the theoretical and methodological issues in the field of writing may be addressed by modulating Kellogg's original procedure (e.g., the global configuration, the training phase, the interval between auditory signals, the number and nature of the categories used in the directed retrospection).

Because the number and nature of retrospective categories can be freely selected as a function of the cognitive activity under investigation and the particular research question, SCRIPTKELL is not limited to writing tasks; it can also be used to study other cognitive tasks in different areas of investigation. For example, it could be possible to use SCRIPTKELL for analyzing the cognitive effort and the sequence of mental processes when participants are confronted with text reading, problem solving, reasoning, or decision making (e.g., Britton & Tesser, 1982). Therefore, SCRIPTKELL may offer interesting new ways for studying how individual characteristics (skills, knowledge, working memory span, age, etc.) affect mental processes involved in several higher level tasks.

## Availability

A copy of SCRIPTKELL can be obtained without charge from the first author ([annie.piolat@up.univ-mrs.fr](mailto:annie.piolat@up.univ-mrs.fr))

## REFERENCES

- BARBIER, M. L., PIOLAT, A., & ROUSSEY, J.-Y. (1998). Effet du traitement de texte et des correcteurs orthographiques sur la maîtrise de l'orthographe et de la grammaire en langue seconde [The effect of word processors and spell checkers on the acquisition of spelling and grammatical skills in second language]. *Revue Française de Pédagogie*, 122, 83-98.
- BRITTON, B. K., & TESSER, A. (1982). Effects of prior knowledge on use of cognitive capacity in three complex cognitive tasks. *Journal of Verbal Learning & Verbal Behavior*, 21, 421-436.
- BROWN, J. S., McDONALD, J. L., BROWN, T., & CARR, T. H. (1988). Adapting to processing demands in discourse production: The case of handwriting. *Journal of Experimental Psychology: Human Perception & Performance*, 14, 45-59.
- ERICSSON, K. A., & SIMON, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.
- FISK, A. D., DERRICK, W. L., & SCHNEIDER, W. (1986-1987). A methodological assessment and evaluation of dual-task paradigms. *Current Psychological Research & Reviews*, 5, 315-327.
- GLYNN, S. M., BRITTON, B. K., MUTH, D., & DOGAN, N. (1982). Writing and revising persuasive documents: Cognitive demands. *Journal of Educational Psychology*, 74, 557-567.
- HAYES, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 1-27). Mahwah, NJ: Erlbaum.
- HAYES, J. R., & FLOWER, L. S. (1980). Identifying the organization of writing process. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach* (pp. 3-30). Hillsdale, NJ: Erlbaum.
- HAYES, J. R., FLOWER, L. S., SCHRIEVER, K., STRATMAN, J., & CAREY, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics: Vol. 2. Reading, writing, and language processing* (pp. 176-240). Cambridge: Cambridge University Press.
- JANSEN, D., VAN WAES, L., & VAN DEN BERGH, H. (1996). Effects of thinking aloud on writing processes. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 233-250). Mahwah, NJ: Erlbaum.
- KAHNEMAN, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- KELLOGG, R. T. (1986). Designing idea processors for document composition. *Behavior Research Methods, Instruments, & Computers*, 18, 118-128.
- KELLOGG, R. T. (1987a). Effects of topic knowledge on the allocation of processing time and cognitive effort to writing processes. *Memory & Cognition*, 15, 256-266.
- KELLOGG, R. T. (1987b). Writing performance: Effect of cognitive strategies. *Written Communication*, 4, 269-298.
- KELLOGG, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 14, 355-365.
- KELLOGG, R. T. (1994). *The psychology of writing*. New York: Oxford University Press.
- KELLOGG, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 57-71). Mahwah, NJ: Erlbaum.
- KERR, B. (1973). Processing demands during mental operations. *Acta Psychologica*, 1, 101-112.
- LEVY, C. M. (1997). The "R" that psychology forgot: Research on writ-

- ing processes. *Behavior Research Methods, Instruments, & Computers*, 29, 137-145.
- LEVY, C. M., MAREK, I. P., & LEA, J. (1996). Concurrent and retrospective protocols in writing research. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn (Eds.), *Writing research: Theories, models and methodology* (pp. 542-556). Amsterdam: Amsterdam University Press.
- LEVY, C. M., & RANSDELL, S. (1994). Computer-aided protocol analysis of writing processes. *Behavior Research Methods, Instruments, & Computers*, 26, 219-223.
- LEVY, C. M., & RANSDELL, S. (1995). Is writing as difficult as it seems? *Memory & Cognition*, 23, 767-779.
- LEVY, C. M., & RANSDELL, S. (1996). Writing signatures. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 149-161). Mahwah, NJ: Erlbaum.
- OLIVE, T. (1997). Validité de la méthode de TR et de rétrospection dirigée pour étudier la production de texte [Validity of the method of reaction time with directed retrospection for the study of writing]. *Informations in Cognito*, 9, 19-30.
- OLIVE, T., & PIOLAT, A. (1997, May). Existe-t-il un lien entre l'effort cognitif développé par des rédacteurs et la qualité du texte? [Is there a relationship between a writer's cognitive effort and the quality of written text?] Paper presented at the Congrès Annuel de la Société Française de Psychologie, Nice.
- OLIVE, T., PIOLAT, A., & POLGE, M. (1997, August). *An evaluation of mental effort of professional and experienced writers during the use of a word processor*. Paper presented at the 7th European Conference for Research on Learning and Instruction, Athens.
- OLIVE, T., PIOLAT, A., & ROUSSEY, J.-Y. (1996, October). *The effect of the degree of expertise on mental effort associated with writing processes*. Paper presented at the European Writing Conference, Barcelona.
- OLIVE, T., PIOLAT, A., & ROUSSEY, J.-Y. (1997). Effort cognitif et mobilisation des processus: Effet de l'habileté rédactionnelle et du niveau de connaissances [Cognitive effort and processing time: Effect of writing ability and level of knowledge]. In D. Mellier & A. Vam Hoffe (Eds.), *Attention et contrôle cognitif: Mécanismes, développement, habiletés et pathologies* (pp. 71-85). Rouen: P.U.R.
- PIOLAT, A., BARBIER, M. L., & ROUSSEY, J.-Y. (1996). Performance of grade students using a speller while learning a second language (Abstracts of the International Congress of Psychology). *International Journal of Psychology*, 31, 284-86.
- PIOLAT, A., FARIOLI, F., OLIVE, T., & ROUSSEY, J.-Y. (1996). Test of Kellogg's cognitive effort method in a writing task. (Abstracts of the International Congress of Psychology). *International Journal of Psychology*.
- PIOLAT, A., & ROUSSEY, J.-Y. (1993). *La charge cognitive des processus de révision d'un texte en fonction des tapes d'erreurs* [Cognitive load of the processes involved in text revision]. Unpublished manuscript, University of Provence.
- PIOLAT, A., ROUSSEY, J.-Y., OLIVE, T., & FARIOLI, F. (1996). Charge mentale et mobilisation des processus rédactionnels: Examen de la procédure de Kellogg [Cognitive effort and time processing: Test of Kellogg's writing task]. *Psychologie Française*, 41, 339-354.
- PIOLAT, A., ROUSSEY, J.-Y., & ROUS, P. (1996, October). *Effects of attention directed to different texts on the mental effort related to writing processes*. Paper presented at the European Writing Conference, Barcelona.
- POWER, M. J. (1986). A technique for measuring processing load during speech production. *Journal of Psycholinguistic Research*, 15, 371-382.
- RANSDELL, S. E. (1990). Using a real-time replay of students' word processing to understand and promote better writing. *Behavior Research Methods, Instruments, & Computers*, 22, 142-144.
- SCARDAMALIA, M., & BEREITER, C. (1991). Literate expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 172-194). Cambridge: Cambridge University Press.
- SCHUMACHER, G. M., KLARE, G. K., CRONIN, F. C., & MOSES, J. D. (1984). Cognitive activities of beginning and advanced college writers: A pausal analysis. *Research in the Teaching of English*, 18, 169-187.
- SMAGORINSKY, P. (ED.) (1994). *Speaking about writing: Reflections on research methodology*. Thousand Oaks, CA: Sage.
- TYLER, S. W., HERTEL, P. T., MCCALLUM, M. C., & ELLIS, H. C. (1979). Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning & Memory*, 5, 607-617.

## NOTES

1. SCRIPT is used in French in the sense of "scenario" (i.e., a sequence of actions); KELL refers to "Kellogg."
2. Note that HyperCard uses the point as decimal separator. It should be verified that EXCEL also uses the point as separator. If not, this mode can be changed in the number file of the control panel.

(Manuscript received April 24, 1997;  
revision accepted for publication January 29, 1998.)