UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

The Time Course of Routine Action

Permalink <https://escholarship.org/uc/item/2kf1g2gs>

Journal Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN 1069-7977

Authors

Cooper, Richard P. Mareschal, Denis Ruh, Nicolas

Publication Date 2005

Peer reviewed

The Time Course of Routine Action

Nicolas Ruh (n.ruh@psychology.bbk.ac.uk)

Richard P. Cooper (r.cooper@bbk.ac.uk)

Denis Mareschal (d.mareschal@bbk.ac.uk)

School of Psychology, Birkbeck, University of London Malet Street, London, WC1E 7HX, UK.

Abstract

Previous studies of action selection in routinized tasks have used error rates as their sole dependent measure (e.g. Reason, 1979; Schwartz et al., 1998). Consequently, conclusions about the underlying mechanisms of correct behavior are necessarily indirect. The present experiment examines the performance of normal subjects in the prototypical coffee task (Botvinick & Plaut, 2004) when carried out in a virtual environment on screen. This has the advantage of (a) constraining the possible errors more tightly than a real world environment, and (b) giving access to latencies as an additional, finer grained measure of performance. We report error data and timing of action selection at the crucial branching points for the production of routinized task sequences both with and without a secondary task. Processing branching points leads to increased latencies. The presence of the secondary task has a greater effect on latencies at branching points than at equivalent non-branching points. Furthermore, error data and latencies dissociate, suggesting that the exact timing is a valid and valuable source of information when trying to understand the processes that govern routine tasks. The results of the experiment are discussed in relation to their implication for computational accounts of routine action selection.

Introduction

A large proportion of the activities that fill our days can be described as *hierarchical sequential routine action*. They are *routine* because we perform these tasks quite often – often enough, in fact, to be able to carry them out without paying much, if any, attention (e.g. brushing our teeth, preparing breakfast, driving to work). They are *sequential* because they require several actions to be performed one after the other. The underlying system(s) not only needs to select the correct action but at the right point in time. Finally, they are *hierarchical* because the sequences in question are best described as basic level actions that are grouped into subsequences that in turn are concatenated in one or the other way to make up longer task sequences. A subsequence is defined as an invariant chain of actions that may appear in different contexts. Thus, one sequence by itself is not hierarchical, but the fact that parts of it can appear in other task sequences as well, or at different places in a task sequence, suggests a hierarchical structuring not unlike the tree structures in generative linguistics.

As in the case of linguistics, it is an open question whether hierarchal structures are an inbuilt feature or an emergent property of the underlying system(s). With respect to this issue, branching points, the steps in a task sequence where a new sub sequence is entered, are of particular importance. Lashley (1951), for example, argued that such points are problematic for simple associative chaining accounts of sequential behavior.

Empirically, routine action has mainly been studied in neurological patients with ADS (action disorder syndrome), a pathology that is usually related to lesions of the prefrontal cortex and leads to severe behavioral breakdown in familiar sequential tasks (e.g. Schwartz et al., 1998; Schwartz, Reed, Montgomery, Palmer, & Mayer, 1991; Sirigu, Zalla, Pillon, Grafman, & Agid, 1995; Zanini, Rumiati, & Shallice, 2002). The errors committed by patients seem to be structurally similar to slips of actions observed in normals (Reason, 1979). Both patient data and action slips and lapses in normals have been interpreted in the light of the single well formulated theory in the field, the SAS (supervisory attentional system) theory of Norman & Shallice (1986), which states that action sequencing involves an executive component (SAS) superimposed upon a basic level system. Its role is to enforce the deliberate selection of an action (sub) sequence in cases where no appropriate schema exists or when the appropriate schema is not likely to be selected given the current context. In this framework, erroneous action selection is explained as a failure of the supervisory process, caused by a lack of attentional resources (distraction) in normals or by malfunction due to brain lesion in patients.

This view is exemplified in a computational model (Cooper & Shallice, 2000) that employs hierarchically organized interactive activation networks (IAN) in which symbolically represented schemas compete for selection. In contrast, Botvinick & Plaut (2004) claim to capture the data within a single embedded SRN (simple recurrent network) that produces the error patterns observed in normals and patients when injected with different amounts of noise. It is important to note that the two major differences between these models (task representation and number of systems) are not necessarily connected.

More data is needed to distinguish between the two approaches. Empirical work on routine action is hampered by several factors, though. One is the heterogeneity of patient behavior following brain lesions. This makes it hard to work with patient groups, while single case studies risk being idiosyncratic. Neurologically unimpaired subjects, on

the other hand, are so good at performing routine tasks that it is all but impossible to observe or induce slips of actions in a controlled environment (but see Humphreys, Forde, & Francis, 2000). This is why the main body of evidence on normals still consists in the extensive diary studies by Reason (1979, 1984). These studies yielded useful insights, but are limited by methodological problems concerning the accuracy and completeness of participant reports. Another especially crucial problem in the study of action errors is the difficulty in producing an objective interpretation of the observed behavior (see Schwartz et al., 1991). Without clear knowledge of the actor's intention, it can be difficult to be sure if an apparent action slip truly was a slip, and if so whether it was due to, for example, an intrusion of another task or mis-selection of an object. In a similar vein, the transfer of error data to the usually correct performance of the underlying system(s) is indirect.

One possibility that overcomes some of the above problems is to test subjects in a virtual environment. This allows for a tighter control of errors by restricting the possible interactions, which makes the classification of errors easier and less prone to misinterpretation. A further benefit is the availability of latencies as an additional and more direct measure of processing difficulties. This paper takes a first step towards the use of "virtual reality" in investigating routine action. In its present form, obviously, the interaction with the objects on screen is very limited and thus nowhere near full scale VR.

We report an experiment that shows a dissociation of latencies and error data, implying that the exact timing is the finer grained measure. Furthermore we found an interaction in latencies between branching points and the presence of a secondary task. This demonstrates the validity of latency data. The results are discussed in the context of the schema representation and dual systems issues described above.

Part 1: Learning to make coffee on screen

The main purpose of the first part of the experiment was to familiarize participants with the virtual coffee task. Participants had to discover the order of steps required to make tea or coffee, subject to constraints imposed by the environment, the instructions and their previous knowledge. For the sake of comparability, the virtual coffee task was held as closely as possible to the task employed in both of the above simulations (see Botvinick & Plaut, 2004). Task sequences were constructed by concatenating a choice out of six invariant sub sequences:

adding coffee grounds (7 steps); adding teabag (6 steps);

adding milk (7 steps); adding sugar from pack (7 steps); adding sugar from bowl (8 steps); drink $(4 \text{ steps})^1$.

Coffee always required adding both milk and sugar, whereas tea was always to be made with sugar only. This leads to four valid coffee sequences:

- c1: grounds sugar from bowl milk drink $(26$ steps)
- c2: grounds milk sugar from bowl drink (26 steps)
- c3: grounds sugar from pack milk drink (25 steps)
- c4: grounds milk sugar from pack drink (25 steps)

and two variations in making tea:

- t1: teabag sugar from pack drink (17 steps)
- t2: teabag sugar from bowl drink (18 steps)

Branching points are of specific interest because the system has to determine the next step by taking into account (a) the context of task sequence (tea or coffee), (b) the history of getting there (sugar already added or not) and (c) the possible choice of valid sub sequences to enter at this point.

Method

Participants & Materials 40 participants (age range: 18 – 59; 21 male) performed both parts of the experiments with an interval of $1 - 2$ weeks between sessions.

Production task: Subjects were faced with a "stage" that showed 11 objects (see Figure 1). The stage had 13 possible object locations, objects were allocated to their respective positions randomly at each trial. In the production task, there were three different conditions with respect to the stage set up. In 50% of the trials, all objects involved in preparing either beverage (cup, teabag, coffee grounds, milk container, sugar packet, sugar bowl, spoon, mouth) were present (unforced cases). In half of the remaining trials the sugar packet was missing, in the other half the sugar bowl was absent (forced trials). The required number of objects was achieved by filling the stage with randomly selected distracters (Nutella jar, tomato, knife, fork, cork screw) in each trial.

Figure 1: A typical stage set-up.

One block of eight trials consisted of two coffee tasks with all objects present, two coffee tasks with one sugar source missing and four tea tasks with the same distribution of setups. The order of trials within a block was randomized.

Subjects were required to make a cup of coffee or tea on screen. This was to be done by manipulating the objects with a standard computer mouse. Clicking on any object led to picking it up (shown by magnifying it by 130%) and

 $\frac{1}{1}$ Sub sequences will be abbreviated by ss_{grounds} , ss_{pack} , etc. As an example, ss_{grounds} consists of the actions: pick-up coffee pack – pull-open coffee pack – pour grounds into cup – put-down coffee pack – pick-up spoon – stir – put-down spoon.

attaching it to the mouse pointer. Clicking once again if an object already was picked up led to putting it down. Clicking once while the held object was over another object led to an interaction of the two, if possible. If, for example, the empty spoon was dragged over the cup and clicked on, it performed a stirring action. Clicking twice (within 250 ms), finally, led to a change of state of the target object $-$ if possible. Double-clicking on the closed sugar pack, e.g., would open it. Production of any of the six valid sequences led to positive feedback in the form of the mouth going "Mmhhh!" after drinking the beverage.

Secondary task: The aim of the secondary task was to divert attention from the production task without interfering in other ways. Therefore, the secondary task was purely auditory, required a response only after completion of the production task and was set up to be as unpredictable as possible to avoid routinization of the secondary task or success by guessing.

While preparing the beverage, subjects would hear different quotes out of several Monty Python movies. They were required to count how often the utterance of the word "Ni" occurred. The "Ni"-sound (lasting 0.5 sec) was chosen with a probability of 50%. Due to the large differences in length of the other 34 quotes $(1.0 - 6.4 \text{ sec})$ it was not possible to predict how often the "Ni" would appear in the variable time span it took a subject to prepare a cup of coffee/tea. The "Ni"-frequency ranged from 0 to13. Subjects were asked to report how many "Ni's" they heard after completion of a production task.

ProcedureSubjects were introduced to the handling of the objects in a guided manner. They learned how to pick up, put down or open an object and how objects can interact. All these steps were explained in detail, subjects were encouraged to try them out with single objects provided for this reason. Subjects were asked to stir after each ingredient was added. Subjects had to find out how to perform the whole task correctly by themselves, they were not shown a complete sequence to prevent inducing any preferences. They were told, however, that their aim should be to make the mouth go "Mmhhh" when drinking the beverage, as this indicated a perfect trial.

The experiment began with 20 trials of making coffee/tea without secondary task ("sect–" trials). The reminder of the experiment consisted of three blocks of the production task plus secondary task ("sect+" trials). This resulted in 44 trials of preparing a beverage.

Each trial started with one of three instructions:

"make a cup of tea with sugar",

"make a cup of coffee with milk and sugar" or

"make a cup of coffee with sugar and milk".

Additionally, subjects were prompted to "count the "Ni's" in the sect+ trials. For each task, actions/answers were recorded as well as latencies of each action.

Results

All subjects learned to produce at least some correct tasks; most ended up with producing most task versions at least once. Out of 44 times each subject was asked to prepare a beverage, on average 22.8 lead to correct task sequences.

Figure 2: Overall task performance in terms of processing time (line) and correct task sequences (bars).

Figure 2 shows how subjects improved in the preparation task. Surprisingly, they reached a rather high level early on (around trial number 14) and did not improve very much subsequently. The introduction of the secondary task from task 20 on influenced the mean latencies more than the error rate, indicating that most participants chose to go more slowly when faced with a more difficult task, while keeping their performance good.

Preferences: Table 1 shows the distribution of sequences produced by all subjects across all trials. Out of the 1760 trials, 394 correct coffee sequences were produced and 520 correct tea sequences.

Table 1: Distribution of correct task versions.

Looking only at the unforced cases (i.e. with free choice of the sugar source), it turns out that subjects developed a strong preference for using the sugar pack (in c3, c4 and t1) rather than using the sugar bowl (c1, c2 and t2). In terms of adding sugar first $(c1 + c3)$ or milk first $(c2 + c4)$, the preference is less evident (209/185). Closer inspection reveals that many subjects followed the order implied by the exact instruction. Thus, if prompted to "make a cup of coffee with sugar and milk" they would add the sugar first.

Secondary task: Subjects performed reasonably well in the secondary task. Out of 960 attempts, the "Ni"-count was correct in 611 cases, and a difference of one between "Ni's"

heard and the subject's count was obtained 291 times. The criterion for having solved the secondary task successfully was set to achieving a difference of one or less.

Latencies: At the first branching point (BP), subjects have the choice of picking up either the sugar pack, the sugar bowl or, in a coffee sequence, the milk container, which leads into the respective sub sequence of adding an ingredient. This action is comparable to picking up the spoon for stirring after the first ingredient has been added (a non-branching point: nBP). Except for the difference in choice availability, the actions are similar.

Only correct trials where the secondary task was solved to criterion were taken into account for further analysis. For each subject the median of all correct trials in the respective condition was calculated. Non-parametric tests were employed to compare latencies at BPs and nBPs because the latencies were not always normally distributed.

Figure 3: Mean latencies at branching points and nonbranching points for the six valid task versions.

The mean latencies in Figure 3 show that, in most task versions, processing BPs takes longer than processing nBPs. The effect size is large when the less favored ss_{bowl} is entered $(Z_{t2}(37) = 4.401$, $p < 0.001$; $Z_{c1}(20) = 3.841$, $p <$ 0.001), of medium size for choosing to add milk $(Z_{c2}(23) =$ 1.900, $p = 0.057$; $Z_{c4}(34) = 3.685$, $p < 0.001$) and very small to lacking for the preferred ss_{pack} (Z_{t1}(39) = 0.081, p = 0.936; $Z_{c3}(34) = 2.850$, $p = 0.004$). Latencies at nonbranching points are relatively invariant across conditions. Obviously, these results have to be treated carefully because firstly, the data are very sparse for some task versions, especially c1 and c2, and secondly they collapse over the process of learning to perform the task – if only over the successful attempts.

Interestingly though, the significance levels remain essentially unchanged when looking at the last 24 tasks only (exceptions where it is sensible to assume that sample sizes were too small). A closer inspection reveals that this pattern is not static: the overall latencies decrease with increasing trial number, the difference between latencies at BP and nBP gets smaller, while the variance also decreases with increasing experience. Thus, although the overall pattern of performance did not change qualitatively, subjects became more efficient and more regular in preparing a beverage on screen.

Discussion

The results of part 1 confirm the special status of branching points. The increase in latencies seems to depend on the subject's preferences. That is, less favored sub sequences are harder to access. Unfortunately, it is unclear how preferences relate to frequency as the calculation of frequencies depends on which contexts are collapsed together, i.e. (a) whether correct sub sequences in incorrect task sequences are taken into account, (b) if ss_{milk} in c3 (milk last) contributes to the frequency of ss_{milk} in c4 (milk first), and (c) if transfer between structurally similar sub sequences is assumed (e.g. ss_{grounds}, ss_{pack} and ss_{milk}).

In summary, the virtual environment approach appears promising. Although not presented with correct examples, subjects learned quickly and produced a good amount of correct task sequences. The majority of errors observed were minor and can be attributed to the special requirements of performing the task on screen, such as confusing the very similar actions of clicking once instead of twice (and vice versa) or trying to put down an object in an area that was already occupied by another object. Latencies seem to reflect some of the underlying complexity of the task; preparing a beverage on screen thus appears to preserve some of the crucial properties of this task in the real world. The secondary task, however, failed to elicit increased error rates. In part 2, we will increase the demands of the secondary task and have a closer look at its effects.

Part 2: The virtual coffee routine

Assuming that correct task representations are in place, we were now able to look at routine behavior in these tasks.

In order to make the secondary task more challenging, the target sound was varied. This time the quotes stemmed from the Star Wars movies and subjects were prompted as to which one of three short sound events was to count in each individual sect+ trial. Each of the possible target sounds occurred with a probability of 25%, giving an observed range of $0 - 15$. Participants were provided with instantaneous feedback on whether they counted correctly or not. Blocks with and without a secondary task were alternated so that performance in these conditions could be compared directly.

A second modification concerned the instructions that had proven to guide subject's preference for order. The instructions at the start of part 2 included the statement that "coffee always requires adding both milk and sugar, whereas tea is always made with sugar only". Subsequently, subjects were simply prompted to make tea or coffee.

Method

Materials and Procedure Materials and procedures were similar to part 1, except for modifications concerning task instructions and the secondary task as mentioned above.

The experiment started with 4 training trials followed by 6 blocks of 8 tasks (block wise alternating sect+/sect–). Blocks were constructed as described in part 1.

Results

Figure 4: Overall performance in part 2, block boundaries are indicated by dotted lines.

Figure 4 shows that subjects started out with performance comparable to the end of part 1, improving only slightly further on. Again, performing the secondary task influenced processing speed, but not the error rate. Ignoring the training trials, 1857 correct task sequences were produced (see Table 2). 124 of these were excluded from further analyses because the secondary task was not solved to criterion (see part 1).

Again, in unforced trials subjects preferred adding sugar from the pack rather than the bowl. The overall distribution of correct task versions was roughly preserved in individual subjects with one exception: most subjects developed an individual preference of order. This is not reflected in the overall numbers, because the two groups cancel each other in their converse preference (17 subjects favored adding sugar first (ratio $> 2:1$), 12 usually added milk first (ratio $>$ 1:2), whereas the remaining 11 subjects did not show a strong preference of order).

Latencies: The pattern of processing times resembled the latencies obtained in part 1. Again, latencies at nBPs were invariant across task versions. Processing the BP took longest when initializing the sugar-from-bowl sub sequence, less so for entering the milk sub sequence, but was not notably prolonged in the preferred versions.

Alternation of sect+ and sect– blocks in the second part of the experiment allowed exploration of the effect of secondary task presence on latencies at branching and nonbranching points. Figure 5 shows an interaction between the

variables when entering either of the sugar sub sequences $(F_{pack}(1,78) = 4.055, p = 0.047; F_{bowl}(1,62) = 8.175, p =$ 0.006), but not for the milk sub sequence $(F_{milk}(1,55) =$ 0.158). The latter is partly explained by the sparse data in this condition. When subjects who contributed only two examples or less are excluded, the interaction is more evident, though not quite statistically significant $(F(1,20) =$ 3.581, $p = 0.073$ due to the small sample size.

Figure 5: Interaction of secondary task and branching point across related sequences.

Error data: Table 3 shows the distribution of error types across trials. Of the 703 erroneous sequences, again, most were due to minor lapses caused by properties of the virtual environment (e.g., trying to put down an object on a region of the screen that was already occupied by another object, or inappropriate single or double clicks). Apart from these inaccuracies, subjects performed at the high level that is to be expected in routinized tasks.

Table 3: Distribution of errors as a function of the presence of the secondary task.

Error type:	minor lapses	stirring omission	sequence wrong errors	task	other total	
sect-	261	49				340
$sect+$	247	57	10	38		363
total	508	106	21	42	26	

At a more conceptual level, omissions of stirring (often realized and rectified after one to four steps) were also common. Only 21 full sequence errors were committed. They included most of the possible incorrect concatenations of sub sequences. The only error that was notably more frequent in the sect+ condition consisted in performing a full, but inappropriate task sequence (usually making coffee when required to make tea). Possibly subjects were distracted by the additional prompt regarding the sound to count and therefore disregarded or forgot the task instruction. The remaining category includes errors that were hard to classify. Some should be attributed to participants exploring their possibilities in the environment (can I stir with the knife?), but about half of them could be interpreted as recovered sequence errors. The distribution of errors did not differ when trials with and without secondary task are compared (Kolmogorov-Smirnov test for equal distributions: $p = 0.958$).

Discussion

Again, branching points prolonged latencies, more so when disfavored sub sequences were entered. However, no effect was observed in the case of the preferred task version without a secondary task. The secondary task induced further processing difficulties at BPs but had no effect on latencies at nBPs, thus suggesting a conflict with resources that are dedicated specifically to help processing at BPs, rather than an overall slowing down of processing speed.

In terms of errors, no qualitatively different behavior was observed in the blocks with or without a secondary task. This dissociation of error data and latencies seems to indicate that the secondary task employed was not disruptive enough to elicit sequence errors, but showed its specific influence in the interaction found in the latency data. The vast majority of errors committed were minor lapses of little theoretical interest. The few sequence errors observed do not clearly differentiate between either way of representing hierarchical task sequences. All of them could be interpreted either as misplaced sub sequence that wrongly won the competition in a basic IAN-type action selection system (Cooper & Shallice, 2000), or as drift into a related task sequence whose internal representation resembles the intended sequence at one point in time (Botvinick & Plaut, 2004). In fact, it is doubtful if it is possible to distinguish between the two representational approaches on the basis of error data alone.

General discussion

The patterns of results (stable over both parts of the experiment) indicate the validity of the use of latencies as a measure in routine tasks and confirm the successful implementation of the new experimental paradigm. The theoretical claim that BPs are harder to process than steps within a sub sequence is supported by prolonged processing times at BPs, while the fact that a secondary task specifically influences BPs but has no effect on nBPs suggests that the observed effects are due to the special properties of branching points and not caused by some confounding variable. Finally, the dissociation with the obtained error data implies that latencies are the finer grained of the two measures.

The two main theoretical results furthermore speak to the issues of task representation and number of systems. The observed interaction strongly supports the two systems view. An additional system seems to influence and facilitate the selection process at branching points only, even in the case of the most preferred task version. This result is in line with the hypothesized SAS and the IAN model (Cooper & Shallice, 2000). However, as currently implemented, this model has difficulties in accounting for the second result, namely the fact that latencies at BPs seem to be influenced by preferences/familiarity. An SRN model might naturally capture this aspect of the data. Unfortunately, Botvinick & Plaut's (2004) model as currently implemented is unable to address this issue directly because it is crucially dependent on a carefully balanced training set that ensures an equal distribution of possibilities at each branching point. With an unbalanced training set, the network is unable to access infrequent task versions because it has no means of overcoming the higher activation of the more frequent task version at a branching point. Enforcing the selection of the non-preferred option, however, is the very function a superimposed executive system would serve.

In conclusion, it seems that a combination of the two existing computational models, namely a familiaritydependent basic system interfaced with a supervisory system (SAS) to bias it at crucial points in a sequence, would be most consistent with our data.

References

- Botvinick, M. M., & Plaut, D. C. (2004). Doing without schema hierarchies: a recurrent connectionist account to normal and impaired routine sequential action. *Psychological Review, 111*(2), 395–429.
- Cooper, R., & Shallice, T. (2000). Contention Scheduling and the control of routine activities. *Cognitive Neuropsychology, 17*(4), 297–338.
- Humphreys, G. W., Forde, E. M. E., & Francis, D. (2000). The Organization of Sequential Action. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes, Attention and Performance 18* (pp. 427–442): MIT Press.
- Lashley, K. S. (1951). The problem of serial order in behaviour. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behaviour*. New York: Wiley.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. Davidson, G. Schwarz & D. Shapiro (Eds.), *Consciousness and Self Regulation* (Vol. 4, pp. 1–18). New York, NY: Plenum.
- Reason, J. (1979). Actions not as planned: the price of automatization. In G. Underwood & R. Stevens (Eds.), *Aspects of Consciousness* (pp. 67–89). London, UK: Academic Press.
- Reason, J. (1984). Lapses of Attention in Everyday Live. In K. Parasuraman & D. Davies (Eds.), *Varieties of Attention* (pp. 515–549): Academic Press.
- Schwartz, M. F., Montgomery, M., Buxbaum, L. J., Lee, S. S., Carew, T. G., Coslett, B. H., et al. (1998). Naturalistic action impairment in Closed Head Injury. *Neuropsychology, 12*, 13–28.
- Schwartz, M. F., Reed, E. F., Montgomery, M., Palmer, C., & Mayer, N. H. (1991). The quantitative description of action disorganisation after brain damage: A Case Study. *Cognitive Neuropsychology, 8*, 381–414.
- Sirigu, A., Zalla, T., Pillon, B., Grafman, J., & Agid, Y. (1995). Selective impairments in managerial knowledge following pre-frontal cortex damage. *Cortex, 31*, 301– 316.
- Zanini, S., Rumiati, R., I., & Shallice, T. (2002). Action Sequencing Deficit Following Frontal Lobe Lesion. *Neurocase, 8*, 88–99.