

 Open access • Journal Article • DOI:10.1037/A0030986

Momentary interruptions can derail the train of thought. — Source link

Erik M. Altmann, J. Gregory Trafton, David Z. Hambrick

Institutions: Michigan State University, United States Naval Research Laboratory

Published on: 01 Feb 2014 - Journal of Experimental Psychology: General (American Psychological Association)

Related papers:

- [Memory for goals: an activation-based model](#)
- [The effect of interruption duration and demand on resuming suspended goals.](#)
- [Preparing to resume an interrupted task: effects of prospective goal encoding and retrospective rehearsal](#)
- [A memory for goals model of sequence errors](#)
- [On the need for attention-aware systems : Measuring effects of interruption on task performance, error rate, and affective state](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/momentary-interruptions-can-derail-the-train-of-thought-37hp6hz875>

Momentary Interruptions Can Derail the Train of Thought

Erik M. Altmann
Michigan State University

J. Gregory Trafton
Naval Research Laboratory, Washington, District of Columbia

David Z. Hambrick
Michigan State University

We investigated the effect of short interruptions on performance of a task that required participants to maintain their place in a sequence of steps each with their own performance requirements. Interruptions averaging 4.4 s long tripled the rate of sequence errors on post-interruption trials relative to baseline trials. Interruptions averaging 2.8 s long—about the time to perform a step in the interrupted task—doubled the rate of sequence errors. Nonsequence errors showed no interruption effects, suggesting that global attentional processes were not disrupted. Response latencies showed smaller interruption effects than sequence errors, a difference we interpret in terms of high levels of interference generated by the primary task. The results are consistent with an account in which activation spreading from the focus of attention allows control processes to navigate task-relevant representations and in which momentary interruptions are disruptive because they shift the focus and thereby cut off the flow.

Keywords: task interruption, procedural error, cognitive control, placekeeping

Understanding the mechanisms of cognitive control is one of the main aims of cognitive psychology, with consequences for understanding human abilities and brain–behavior relationships in a wide range of domains and populations. Cognitive control is expressed in many ways, one of which involves selecting the next thought to focus on when there are multiple options and when interruptions or distractions can intervene. An example is a conversation, which generally follows a coherent thread. If an interruption occurs, such as an interlocutor’s cell phone ringing, the thread may get lost, leading to a “Where were we?” moment (e.g., Trafton, Jacobs, & Harrison, 2012).

An important question for practical and theoretical reasons is just how minimal such an interruption can be and still affect the train of thought. For the cell phone example, the question is whether the interlocutor need only attend to his or her phone for a couple of seconds—to shut it off, for example, rather than to actually take the call. In safety-critical contexts, the question is whether it is safe to interrupt someone even for a few seconds in the middle of a procedure without increasing the chance of error.

To operationalize “train of thought,” we used a task in which subtasks—or what we call *steps*—have to be performed in a particular sequence, and correct performance depends on remembering one’s place in this sequence. In behavioral research, sequential tasks like this arise mainly in two domains, one focused on errors in sequential action selection, and the other on effects of task interruption. In the error domain, the focus has been on “routine” tasks (Norman & Shallice, 1986), like coffee-making (Botvinick & Bylsma, 2005; Botvinick & Plaut, 2004; Cooper & Shallice, 2000, 2006), packing a lunch (Cooper, Schwartz, Yule, & Shallice, 2005), or cleaning the house (Botvinick & Plaut, 2004). However, routine tasks are generally performed with high accuracy, so the error data they generate are sparse. Reason (1990) used diary methods to collect errors in routine tasks over long periods, and others have focused on neurological patients for whom routine tasks are difficult (Schwartz, Reed, Montgomery, Palmer, & Mayer, 1991). In the task interruption domain, sequential tasks with “post-completion” steps (Byrne & Bovair, 1997; Li, Blandford, Cairns, & Young, 2008; Trafton, Altmann, & Ratwani, 2011) generate elevated error rates on those steps, but by design those steps are still a small proportion of the total. In a study of interrupted routine performance, Botvinick and Bylsma (2005) had to have their participants make 50 actual cups of coffee to generate enough errors to analyze.

We therefore saw a need for a procedure that generates rich and interpretable data on sequence errors, from healthy populations performing under controlled laboratory conditions. To meet this need, we developed a sequential task that combines a routine sequencing component with relatively complex individual steps. An important criterion was that every participant should effectively be a trained expert on the step sequence, so that sequence errors could be attributed to on-line control mechanisms rather than lack of knowledge about the sequence. We therefore built the

This article was published Online First January 7, 2013.

Erik M. Altmann, Department of Psychology, Michigan State University; J. Gregory Trafton, Naval Research Laboratory, Washington, District of Columbia; David Z. Hambrick, Department of Psychology, Michigan State University.

This research was supported by grants from the Office of Naval Research to Erik M. Altmann (N000140910093) and J. Gregory Trafton (N0001412RX20082, N0001411WX30014). The authors thank Rick Cooper and Dave Plaut for their comments during the review process, Mark Becker and Amber Markey for comments on earlier drafts, and Amber Markey for collecting the data.

Correspondence concerning this article should be addressed to Erik M. Altmann, Department of Psychology, Michigan State University, East Lansing, MI 48824. E-mail: ema@msu.edu

task around a word functioning as an acronym, with each letter of the word identifying a step and the word itself specifying the step sequence. The operations for each step were mnemonically linked to the letter for that step but were themselves “nonroutine” (Norman & Shallice, 1986) in that the stimulus was randomly constructed and therefore had to be newly parsed each time a step was performed. The task rules and responses also differed for each step, requiring an explicit shift of mental focus in the transition from one step to the next, and making for a task representation with many elements; both sequential focusing and large task representations have been identified as potentially taxing general cognitive ability (Duncan et al., 2008). In all, the task is more like applying a well-learned statistical procedure to new data, for example, than it is to coffee making—with the proviso that all literate adults are in effect trained experts in its sequential structure.

In the remainder of the article, we first describe our task and then present two experiments. Experiment 1 includes a manipulation to examine the influence of the kind of material processed during the interruption, and Experiment 2 tests effects of especially short interruptions; the methods are the same otherwise, so we report the experiments together. In the four sets of analyses that follow we (a) examine effects of interruption on different behavioral measures (sequence errors, nonsequence errors, and response latencies); (b) ask whether these effects interact with differences in the difficulty of individual steps in the step sequence; (c) examine distributions of sequence errors as a function of distance from the correct step; and (d) examine the reliability of sequence errors as a measure of individual differences. In the general discussion, we develop a theoretical perspective on the suite of empirical results.

The Experimental Task

The primary task—the one that was periodically interrupted—is defined by the acronym UNRAVEL. Each letter identifies a step, and the letter sequence defines the order in which the steps are to be performed. The sequence is performed in a cycle, with L followed immediately by U, producing continuous performance that can be interrupted many times.

Figure 1a shows two sample stimuli that illustrate the nature of the task. Each stimulus has various features: two characters (a letter and a digit), a font style applied to one character (underline or italic), a color applied to one character (red or yellow), and one character outside the gray outline box (above or below). (The box appears in the same fixed location on each trial.) Each step of the UNRAVEL sequence requires a two-alternative forced choice applied to one feature. For example, the U step involves choosing whether the font style is underline or italic, and the R step involves choosing whether the color is red or yellow. Not all the steps are this directly perceptual; the N step, for example, requires a choice about whether the letter in the stimulus is near to or far from the start of the alphabet. For each step, its letter in the sequence mnemonically relates to one of the two candidate responses (e.g., “u” for underline, “n” for near to, “r” for red; their opposites are “i” for italic, “f” for far, and “y” for yellow, respectively).

Figure 1b shows the choice rules and responses for all seven steps. The candidate responses are all unique, meaning that a response made by a participant tells us which step he or she thought was correct on that trial, and thus lets us score sequence

errors. We can also score nonsequence errors, meaning trials on which the selected step was correct, but the response was incorrect given the stimulus.

Constructing the stimulus for a given trial involves randomly selecting a letter and a digit and randomly assigning them a left-right order, then randomly and independently assigning a font style, color, and vertical location to one character or the other.¹ This method of construction meant that perceptual search was always required to isolate the correct feature for the current step, because the location of the feature was never predictable, and different features were never correlated. This perceptual search always occurred in a context of interference from features relevant to other steps, because each stimulus afforded performance of all steps.

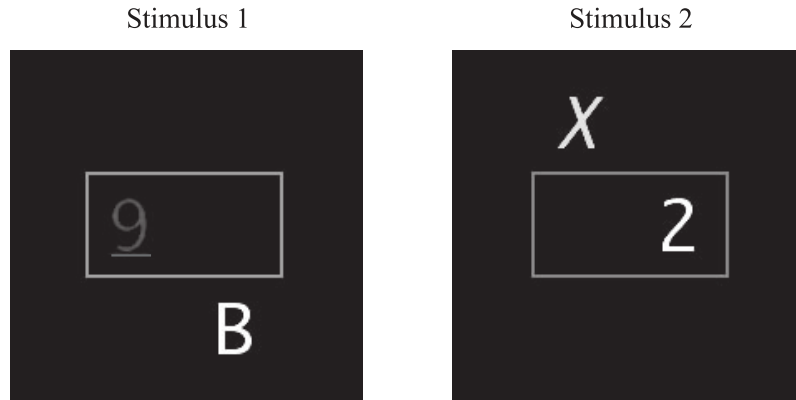
The task imposed various kinds of load. There were several cognitive and perceptual operations involved on each step, including placekeeping operations to select the step to perform, operations to focus on the choice rule for that step, perceptual search for the associated feature, and response selection. Sources of interference included the many stimulus features and responses that were not currently relevant; stale working memory contents, such as placekeeping information and stimulus and response elements from previous trials; and the multiple meanings of task elements. An example of these multiple meanings is that the letter U represents, at various stages and times, a step in the sequence, one of two possible responses for that step, and a stimulus feature requiring an “F” response when the step is N (see Figure 1b). Thus, the letter U in the stimulus could prime several different conflicting elements of the task representation.

Interruptions occurred every six trials on average (a *trial* being a performed step). For a given run of trials following an interruption, the length of the run was computed as the sum of 3 and a rounded sample from an exponential distribution with mean 3. The purpose of the exponential sample was to produce a relatively flat hazard function for the onset of the interruption, starting after the first three trials. The average of six trials per run was one less than the length of the UNRAVEL sequence, so that interruptions would not always occur at the same point in the sequence.

An interruption was triggered by the response to a trial. The interruption began with the interruption stimulus immediately replacing the primary-task stimulus. Figure 1c shows a sample interruption stimulus, made up of a set of characters to type (letters, in this case) and a box to type them into. The participant’s task was simply to type the characters (the “code”) and press the Return key (meaning that the temporal duration of the interruption was self-paced). The number and type of characters in the set (letters or digits) varied between groups, as described

¹ Formally, for purposes of stimulus construction, a stimulus has five *attributes*, each with four candidate *values* (in parentheses): letter (A, B, U, X), digit (1, 2, 8, 9), font style (left-underline, left-italic, right-underline, right-italic), color (left-red, left-yellow, right-red, right-yellow), and height (left-above, left-below, right-above, right-below). For font style, color, and height, the “left” and “right” constituents of the candidate values refer to the left and right character positions, and the “above” and “below” constituents are relative to the box. On a given trial, the left-right order of the letter and digit attributes is randomly determined, then a value is randomly sampled for each attribute, subject to the constraint that a value cannot repeat between trials. For example, Stimulus 1 in Figure 1a has values 9, B, left-underline, left-red, and right-below.

(a) Sample stimuli for UNRAVEL task:



(b) Choice rules and candidate responses for UNRAVEL task, and responses to stimuli in (a):

Step	Candidate responses	Choice rules	Responses to sample stimuli	
			Stimulus 1	Stimulus 2
U	u i	character is Underlined or in Italics	u	i
N	n f	letter is Near to or Far from start of alphabet	n	f
R	r y	character is Red or Yellow	r	y
A	a b	character is Above or Below the box	b	a
V	v c	letter is Vowel or Consonant	c	c
E	e o	digit is Even or Odd	o	e
L	l m	digit is Less than or More than 5	m	l

(c) Sample stimulus for interrupting task:

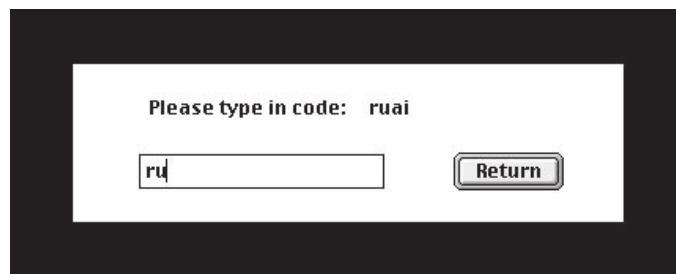


Figure 1. a. Two sample stimuli for the UNRAVEL task (the 9 is red and the X is yellow). b. Response mappings for the UNRAVEL task, and responses for the two sample stimuli in Panel a. c. Sample stimulus for the interrupting task, after two letters have been typed.

below. The cursor began in the box, so mouse movements were not required, and the characters typed into the box remained visible so they could be compared to the presented set. If the participant typed the set correctly, pressing Return ended the interruption and triggered immediate onset of the next trial of the primary task. If the participant made any errors typing the set, no feedback was given after individual errors, but pressing Return cleared the box and displayed a new set of characters for the participant to type.

Experiments 1 and 2

Our two experiments were identical except for the number and type of characters presented for typing during interruptions. In Experiment 1, there were four characters per set, with character type manipulated between participants. For the *letters* group, the characters for each set were sampled randomly without replacement from the 14 responses to the UNRAVEL task (Figure 1b). For the *digits* group, the characters for each set

were selected randomly without replacement from the digits 1–9. We expected that interruption effects would be greater for the letters group because participants would find it more difficult to use external memory for placekeeping. That is, we expected that participants would generally touch type the characters, given that typing is generally automatized in the population we were sampling, and that touch typing would make it difficult to keep a finger on the key corresponding to the most recently performed step (or the correct next step, depending on the hypothetical strategy).

In Experiment 2, we asked if interruptions could be even shorter than four characters and still be disruptive. Accordingly, there were two letters per set, sampled as in Experiment 1.

Method

Participants

Participants were undergraduates from Michigan State University. There were 100 in each group of Experiment 1 and 100 in Experiment 2, for a total of 300. Nine additional participants were replaced because their accuracy was below a threshold, as described below.

Procedure

Participants were tested individually. A session began with a step-by-step introduction to the UNRAVEL step sequence. The introduction emphasized the acronym, showing how each step in turn corresponded to a constituent letter, and then presented a summary screen showing the choice rules for each step and the letters spelling out the word. After this, to ensure that participants understood the task, there were 16 trials during which the computer required the participant to make the correct response on each trial before allowing the participant to move on. This 16-trial sequence was interrupted twice, to illustrate for participants how they were supposed to pick up after an interruption where they had left off. The experimenter remained present during this period to help if necessary. A sheet of paper with the choice rules for the UNRAVEL sequence remained visible to the side of the computer throughout the session.

In preparation for the experimental phase of the session, participants were reminded to “please try to keep your place in the UNRAVEL sequence,” and to “please try to pick up in the sequence where you left off” after an interruption. The experimental phase consisted of four blocks, each with 10 interruptions and thus about 66 trials. A session took about 30 min to complete.

Sequence errors were coded with respect to the previous step. For example, if steps U, R, and A were performed in succession, R would be a sequence error, because N was skipped, but A would be correct, because A follows R in the UNRAVEL sequence.

After each block the participant was given his or her score, computed as the percentage of trials that block for which the step and response were both correct. If the score was above 90%, the participant was asked to go faster. If the score was below 70%, the participant was asked to be more accurate and that block was excluded from analysis (six cases: two in Experiment 1 letters, three in Experiment 1 digits, one in Experiment 2). If a participant scored below 70% on two or more blocks (eight cases: three in

Experiment 1 letters, one in Experiment 1 digits, four in Experiment 2) or did not score significantly above chance on the first trial after an interruption (one additional case, in Experiment 1 letters), he or she was replaced. No feedback was given after individual trials.

Analyses

We analyze the results in four sections, and here give a brief preview of each. Analysis 1 examines interruption effects by comparing performance across successive trials following an interruption (the first five). To preview, interruption effects registered on the first trial after the interruption, with later trials not differing from one another, and the effects registered most strongly for sequence errors, with nonsequence errors showing no effect and response latencies a weak effect.

Analysis 2 examines differences across the steps of the UNRAVEL sequence. Step difficulty varied widely, as reflected in large differences in nonsequence errors and response latencies, but did not interact strongly with interruption effects, suggesting that the latter are reasonably general and not tied to specific choice rules.

Analysis 3 examines sequence errors as a function of distance from the correct step (i.e., the number of steps skipped backward or forward in the sequence). The distributions followed gradients, with fewer errors at greater distances, though the gradients were most clearly defined on trials other than the first after an interruption. The gradients also showed asymmetries that promise to help constrain models of the underlying mechanisms.

Finally, Analysis 4 examines the reliability of sequence errors as a measure of individual differences. Reliability was as high for sequence errors as for other measures of cognitive control, suggesting that our task could be used in conjunction with others to measure sequential control at the latent level, for example, or possibly to predict individual differences in general ability (Duncan et al., 2008).

Analysis 1: Interruption Effects

In this analysis we examined the effects of interruption on three measures of interest. The first measure was sequence errors, defined as the proportion of trials on which the performed step was not the immediate successor in UNRAVEL to the step performed on the previous trial. The second measure was nonsequence errors, defined as the proportion of trials on which the correct step was selected but the incorrect choice was made given the stimulus. The third measure was response latencies, measured on trials on which the step and response were both correct. We did not trim response latencies, and simply took the mean for each participant for each cell of the design (the results were qualitatively unchanged with participant medians). In this analysis we also report mean interruption duration and the effect on this of the group factor in Experiment 1 (letters vs. digits).

In this analysis we separate the data by *position* (1, 2, 3, 4, 5) of a trial after an interruption, where Position 1 is the first trial after the interruption, Position 2 is the second, and so on. Although runs could be longer than five trials, we stop at Position 5 because variability on later positions begins to increase with fewer observations. (Under the method for determining run length that we

noted earlier, more than half of all runs extended to five trials, and fewer than half extended to six trials.)

In this analysis we collapse over the factor *step* (U, N, R, A, V, E, L), because when both position and step are included in the design there are missing cases (i.e., not every Position \times Step \times Participant cell contains trials). Nevertheless, the step factor is meaningful, and we examine its influence in Analysis 2 (to foreshadow, step does not strongly interact with interruption effects).

Results. The data are plotted in Figure 2, with sequence errors in the top panels, nonsequence errors in the middle panels, and response latencies and interruption durations in the bottom panels. Error bars with caps (all except those for interruption duration) are graphical significance tests for pairwise contrasts between Position 1 and the position over which the error bar is located. A contrast is significant if the data point for that position lies outside the bar (which it does if and only if the Position 1 data point also lies outside the bar).² For each measure, we performed an omnibus Position (5) \times Group (2: letter, digit) analysis of variance (ANOVA) for Experiment 1, a one-way ANOVA on position (5) for Experiment 2, and follow-up contrasts as necessary.

For sequence errors, in Experiment 1 the main effect of position was significant, $F(4, 792) = 62.3, p < .001, \eta_p^2 = .239$. A Helmert contrast between Position 1 and the mean of Positions 2–5 was significant, $F(1, 792) = 98.2, p < .001, \eta_p^2 = .332$, whereas Positions 2–5 did not differ significantly ($F < 1$). There was no effect of group ($F < 1$) and no Position \times Group interaction, $F(4, 792) = 1.7, p = .147, \eta_p^2 = .009$. The pattern was similar in Experiment 2: The main effect of position was significant, $F(4, 396) = 8.4, p < .001, \eta_p^2 = .078$, and a contrast between Position 1 and the mean of Positions 2–5 was significant, $F(1, 396) = 12.7, p < .001, \eta_p^2 = .114$, whereas Positions 2–5 did not differ significantly, $F(3, 297) = 1.3, p = .288, \eta_p^2 = .013$. Thus, in both experiments, interruptions increased the rate of sequence errors on the first trial after the interruption. In Experiment 1, the manipulation of character type during the interruption (letters vs. digits) had no effect.

For nonsequence errors, in Experiment 1 there was no effect of position, no effect of group, and no interaction ($F_s < 1$). In Experiment 2, there was again no effect of position, $F(4, 396) = 1.8, p = .122, \eta_p^2 = .018$. Thus, nonsequence errors showed no interruption effects.

For response latencies, in Experiment 1 the main effect of position was significant, $F(4, 792) = 28.7, p < .001, \eta_p^2 = .126$. A Helmert contrast between Position 1 and the mean of Positions 2–5 was significant, $F(1, 792) = 44.3, p < .001, \eta_p^2 = .183$, whereas Positions 2–5 did not differ significantly, $F(3, 594) = 1.8, p = .144, \eta_p^2 = .009$. There was no effect of group and no Position \times Group interaction ($F_s < 1$). In Experiment 2 there was no effect of position, $F(4, 396) = 1.1, p = .351, \eta_p^2 = .011$. Thus, interruptions increased Position 1 latencies slightly but significantly in Experiment 1 and had no effect in Experiment 2.

Finally, interruption duration varied by group in Experiment 1, with faster interruptions for the group that typed letters ($M = 3.96$ s, $SD = 1.07$) than for the group that typed digits ($M = 4.77$ s, $SD = 1.20$), $F(1, 198) = 25.1, p < .001, \eta_p^2 = .112$. This effect may reflect longer typing time for digits, or hand movements to and from the keypad at the right of the keyboard, or some combination.

Discussion. Sequence error rates were significantly higher on Position 1 than on Positions 2–5—by more than a factor of 3 in Experiment 1 (5.8% vs. 1.8%) and more than a factor of 2 in Experiment 2 (4.3% vs. 2.0%). Positions 2–5 did not differ significantly in either experiment, which means that interruption effects were limited to Position 1.

Interruptions themselves were short—4.36 s in Experiment 1 (averaged over groups) and 2.76 s in Experiment 2. Interruption duration in Experiment 2 was numerically slightly faster than the average time to perform a step of the primary task, which was 2.77 s averaged across Positions 1–5. Thus, for interruptions in Experiment 2, their effect on sequence errors cannot be linked to temporal decay alone, which is a theoretical clue that we pursue in the General Discussion.

The type of character typed during the interruption (letter vs. digit, represented by the group factor) did not modulate interruption effects. The implication is either that adopting an external placekeeping strategy was no easier with digits than with letters—or that it was easier with digits, but the strategy had no effect. Either way, this finding suggests that disruptive effects of momentary interruptions are not particularly sensitive to what is processed during the interruption, and instead are mediated by more basic architectural mechanisms.

Nonsequence error rates showed no interruption effects, indicating that interruptions can disrupt placekeeping without disrupting step-specific operations. In the alternative, it seemed possible that interruptions could disrupt global attentional resources or parameters in a way that would affect all kinds of processing afterward. Instead, interruptions affected just those processes that depended on mental context carrying over from before the interruption.

Response latencies showed smaller interruption effects than sequence errors did. In Experiment 1, response latency was significantly higher on Position 1 than on later positions, but the difference was relatively small (3.04 s vs. 2.73 s, a factor of 1.1), and the size of the position main effect was smaller for latencies ($\eta_p^2 = .126$) than for sequence errors ($\eta_p^2 = .239$). In Experiment 2, response latency showed no interruption effect. The small latency effect in Experiment 1 probably reflected additional time spent on placekeeping operations, rather than time spent on step-specific operations, given the null interruption effect for nonsequence errors.

Previous studies of task interruption have usually shown the opposite pattern, with large latency effects and no error effects (Altmann & Trafton, 2007; Grundgeiger, Sanderson, MacDougall, & Venkatesh, 2010; Hodgetts & Jones, 2006a, 2006b; Monk, Trafton, & Boehm-Davis, 2008; Trafton, Altmann, Brock, & Mintz, 2003). However, in many of these studies the task envi-

² The graphical significance tests (GSTs) in Figures 2, 3, and 4 are adaptations of within-subject confidence intervals (Hollands & Jarmasz, 2010; Masson & Loftus, 2003) that visually represent the significance of a difference D between two means. D is statistically significant if and only if $|D| > GST = \sqrt{2CI}$, where $CI = t_{df,\alpha} \sqrt{MSE/n}$ is the within-subject confidence interval for the contrast for D , with df the degrees of freedom for MSE , n the total number of observations contributing to MSE , and significance level $\alpha = .05$ (two-tailed) unless otherwise noted. The GST is centered vertically between the two means, so D is significant if and only if the GST excludes both (and if it excludes one, it also excludes the other).

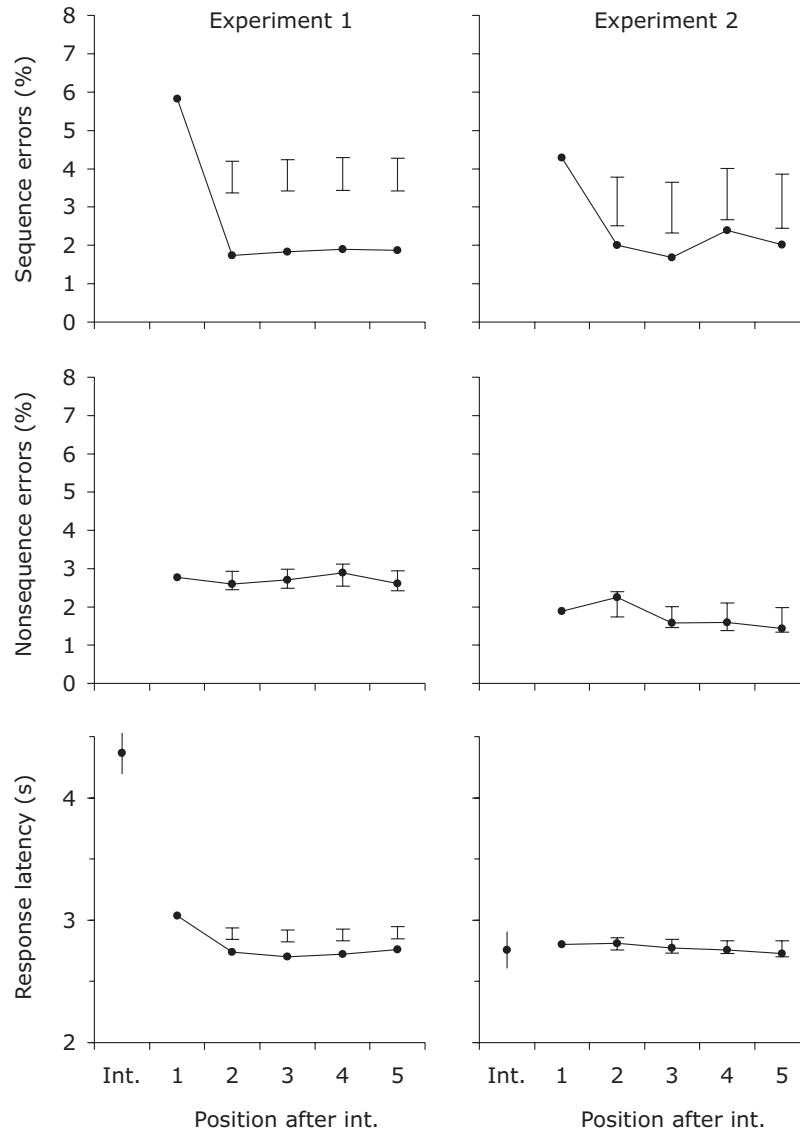


Figure 2. Performance as a function of trial position after interruption. Capped error bars are graphical significance tests for contrasts between Position 1 and the position at which the error bar is located. Uncapped error bars are 95% confidence intervals. Int. = interruption.

ronment had richer perceptual displays, which may have allowed participants to spend time reconstructing their mental context to resume more accurately. Errors also do not always directly measure the representations disrupted by interruptions, as illustrated by the null effect of interruptions here on nonsequence errors. Indeed, interruptions can actually lower error rates, if they trigger changes in arousal (Speier, Vessey, & Valacich, 2003) or performance strategy (Zijlstra, Roe, Leonora, & Krediet, 1999), or if performance would otherwise show a “vigilance decrement” (Ariga & Lleras, 2011). Ariga and Lleras (2011) argued that the vigilance decrement reflects habituation to the primary task and that an interruption forces the primary task to be reactivated when the interruption is over. Thus, the processes reflected in interruption effects vary widely across task environments. In our procedure, the relatively low sensitivity of response latencies to interruption

effects may indicate a high activation of interfering as well as target elements of the task representation, a possibility we explore in the General Discussion.

Analysis 2: Step Effects

In this analysis we examine the effect of step (U, N, R, A, V, E, L) on the three measures examined in Analysis 1. Subjectively the steps are very different in terms of difficulty of applying the choice rules, and one question is whether these subjective differences are borne out behaviorally. A more specific theoretical question is whether any difficulty differences interact with interruption effects. Conceivably, a step that involves more processing and/or more representational elements may leave behind more residual traces in episodic memory

(Altmann, *in press*; Altmann & John, 1999) that, in turn, could make that step easier to remember having performed, thus modulating the frequency of sequence errors.

We evaluate step effects separately from the position effects in Analysis 1, because there are missing cases when both step and position are included in the design. To address this issue here, while still measuring interruption effects, we restructure the position factor from Analysis 1 and refer to it now (and in the rest of the article) as *context*, with levels *post-interruption* and *baseline*. The post-interruption context is simply Position 1, which is where all interruption effects registered in Analysis 1. The baseline context collapses over Positions 2 and later, including those beyond 5, which were excluded in Analysis 1 (where position was a fixed factor). With the data aggregated this way there are only two missing cases (both in the digits

group of Experiment 1, as reflected in the degrees of freedom for Experiment 1 ANOVAs).

Results. The data are plotted in Figure 3, with step on the abscissa and context as separate lines. Sequence errors are again in the top panels, nonsequence errors in the middle panels, and response latencies in the bottom panels. Error bars are again graphical significant tests, here for pairwise contrasts between the post-interruption and baseline contexts at each step. A contrast is significant if the data points at that step lie outside the bar. For each measure, we performed a Step (7) \times Context (2) \times Group (2) ANOVA for Experiment 1 and a Step (2) \times Context (2) ANOVA for Experiment 2.

For sequence errors, in Experiment 1 there were significant main effects of step, $F(6, 1176) = 10.8, p < .001, \eta_p^2 = .052$, and context, $F(1, 196) = 82.3, p < .001, \eta_p^2 = .296$, and a significant

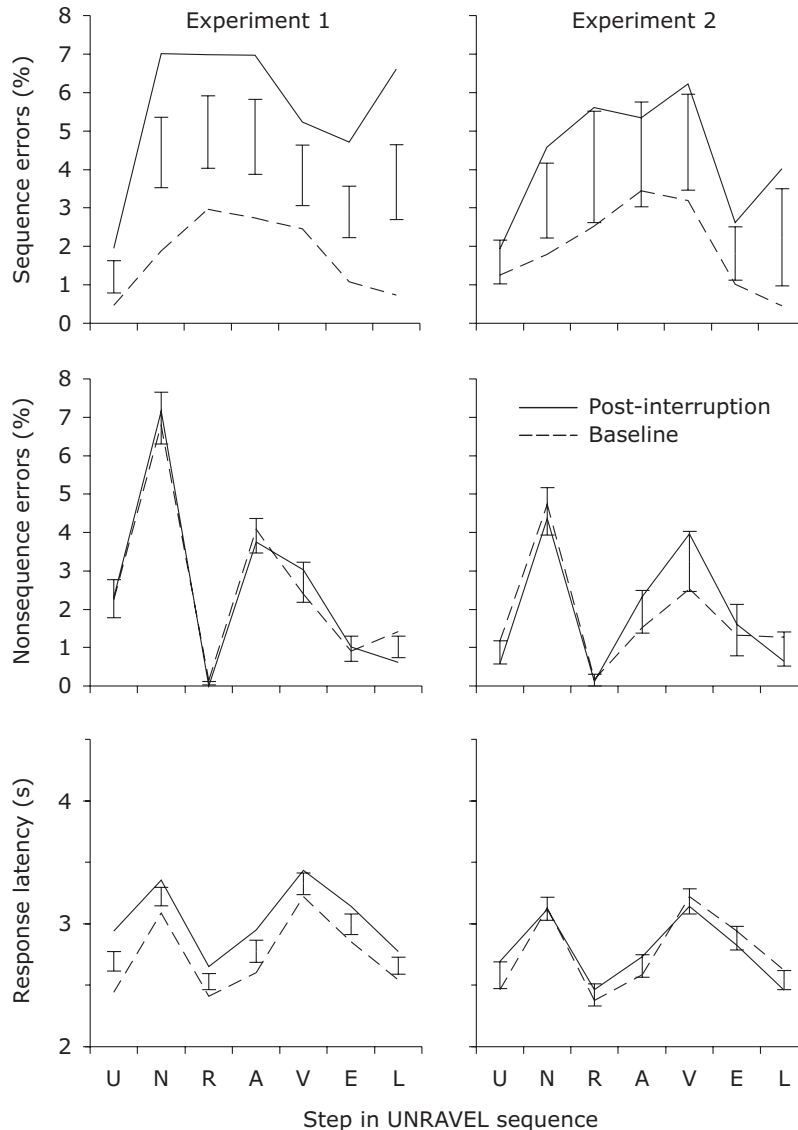


Figure 3. Performance as a function of step within the UNRAVEL sequence. Error bars are graphical significance tests for contrasts between the post-interruption and baseline contexts at each step.

interaction, $F(6, 1176) = 3.4, p = .003, \eta_p^2 = .017$. The main effect of group was not significant, $F(1, 196) = 1.5, p = .229, \eta_p^2 = .007$. There was a marginally significant Step \times Group interaction, $F(6, 1176) = 2.0, p = .068, \eta_p^2 = .010$, which we do not try to interpret. No other interactions involving group were significant ($F_s < 1.2$). In Experiment 2 there were again significant main effects of step, $F(6, 594) = 4.9, p < .001, \eta_p^2 = .047$, and context, $F(1, 99) = 14.7, p < .001, \eta_p^2 = .129$, but here the interaction was not significant ($F < 1$). Thus, in Experiment 2 there was no evidence that interruption effects differed by step. In Experiment 1 there was, but the simple effect of context was highly significant for all steps, as indicated by the graphical significance tests in the top left panel of Figure 3.

For nonsequence errors, in Experiment 1 there was a significant main effect of step, $F(6, 1176) = 15.8, p < .001, \eta_p^2 = .075$, but no main effect of context ($F < 1$) and no interaction, $F(6, 1176) = 1.1, p = .351, \eta_p^2 = .006$. In Experiment 2 there was again a significant main effect of step, $F(6, 594) = 5.7, p < .001, \eta_p^2 = .055$, and no effect of context ($F < 1$). The interaction in Experiment 2 was marginally significant, $F(6, 594) = 2.0, p = .061, \eta_p^2 = .020$, probably reflecting the difference between the contexts at the V step, but this effect had no counterpart in Experiment 1, and we do not try to interpret it. Thus, there was substantial variability in the step factor, but this affected the two contexts in similar ways.

For response latencies, in Experiment 1 there were significant main effects of step, $F(6, 1176) = 58.1, p < .001, \eta_p^2 = .231$, and context, $F(1, 196) = 43.7, p < .001, \eta_p^2 = .184$, and there was a marginally significant interaction, $F(6, 1176) = 2.0, p = .069, \eta_p^2 = .010$. Neither the main effect of group nor its interactions with other factors were significant ($F_s < 1$). In Experiment 2 there was a significant main effect of step, $F(6, 594) = 34.8, p < .001, \eta_p^2 = .264$, no effect of context ($F < 1$), and a significant interaction, $F(6, 594) = 2.7, p = .014, \eta_p^2 = .027$. In both experiments, step effects for latencies were more similar to those for nonsequence errors than to those for sequence errors (Figure 3). Also in both experiments, context effects for latencies mirrored those in Analysis 1, with the post-interruption context (Position 1) showing an interruption effect in Experiment 1 but not Experiment 2.

Discussion. Step had a large effect on performance, but the nature of the effect differed by dependent measure. For sequence errors, step effects seem to reflect effects of the word boundaries—middle steps in the sequence showed higher error rates than end steps, as seen clearly in the baseline contexts (Figure 3, top panels). This pattern may indicate a role for hierarchical representations in place-keeping, with a full pass through the sequence—as defined by the word, here—serving as a task boundary (Botvinick & Bylisma, 2005).

For nonsequence errors and response latencies, step effects seemed to track differences in the difficulty of individual steps. The N and R steps best illustrate the pattern. N requires judging whether a letter is near to or far from the start of the alphabet, which involves search through a mental representation of the alphabet, and somewhat opaque names for the relevant response categories. In contrast, R requires a simple color judgment, and the color names (red and yellow) map transparently to responses (“r” and “y”). Nonsequence errors and response latencies were correspondingly high for N and low for R (Figure 3, middle and bottom

panels). However, the two steps had similar rates of sequence errors (Figure 3, top panels), suggesting that the amount of processing involved in performing a step does not bear much on the accuracy of remembering whether that step was performed.

There were Step \times Context interactions for sequence errors in Experiment 1 (with a similar trend in Experiment 2) and response latencies in Experiment 2 (with a similar trend in Experiment 1). The effects were relatively weak and can be at least partially attributed to a speed–accuracy tradeoff specifically on the U step. There, compared to other steps, the context effect was numerically smallest for sequence errors and numerically largest for response latencies (Figure 3, top vs. bottom panels). What mechanism would drive a speed–accuracy tradeoff on this step is unclear, though it may be related to U being the first step of the sequence. Regardless, the simple effect of context on sequence errors was significant on all steps in Experiment 1 (as indicated by the graphical significance tests), and in Experiment 2 the Step \times Context interaction was not significant. Thus, all steps were still subject to an interruption effect on sequence errors, even if the effect was smaller on some steps than others.

In the end, although step difficulty varied widely, this factor did not interact strongly with interruption effects. Together with the null effect of interruptions on nonsequence errors, this pattern suggests that sequential control is a modular process that is relatively independent of step-specific operations. This finding suggests that different steps could be substituted into the sequence—or that the sequence itself could be replaced—and produce similar results, making the overall architecture of our procedure potentially adaptable to different applications.

Analysis 3: Sequence Error Distributions

In this analysis we examine the distribution of sequence errors over *distance* (−3, −2, −1, +1, +2, +3), the number of steps repeated or skipped within the sequence when a sequence error occurs. For example, a U trial after a U trial would be a −1 error, and an R trial after a U trial would be a +1 error. In keeping with previous work (Cooper & Shallice, 2000; Trafton et al., 2011), we refer to “−” errors as *perseverations* and “+” errors as *anticipations*. When we refer to a distance without a sign, this denotes a particular number of steps from the correct step in either *direction* (perseveration, anticipation).

Sequence error distributions are likely to be an important theoretical constraint, but, as we noted earlier, sequence errors have been difficult to generate in numbers large enough to assess distributions, and only one study to date has done so (Trafton et al., 2011). That study found evidence for a gradient distribution (error rate decreasing as distance increased) for perseveration errors on trials after interruptions. However, there were not enough errors to establish whether there was also an anticipation gradient after interruptions, or whether there were gradients in either direction during baseline performance. Thus, the aim here is to evaluate the evidence for gradients in both directions and in both contexts.

Results. The data are plotted in Figure 4, with post-interruption distributions in the upper panels and baseline distributions in the lower panels (we reuse the context factor from Analysis 2). Error bars are graphical significance tests (with significance level $\alpha = .025$; see below) for pairwise contrasts be-

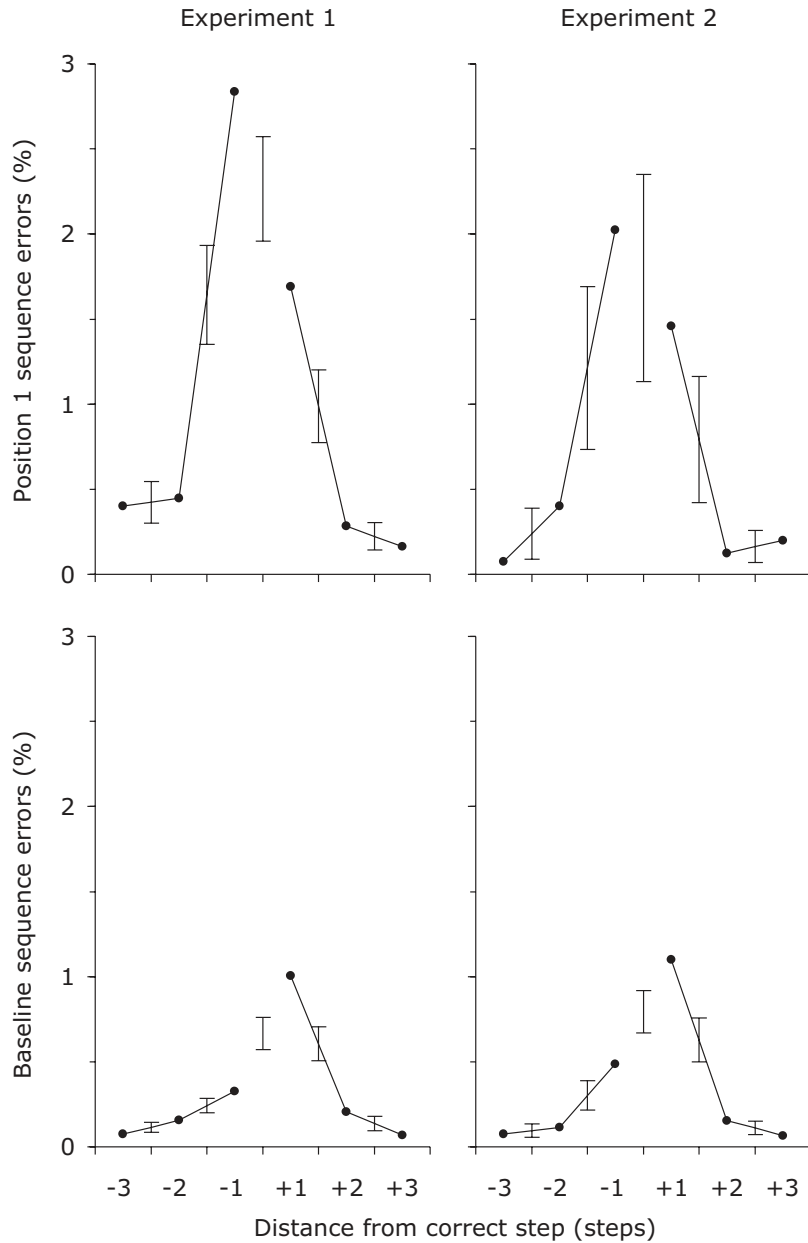


Figure 4. Sequence errors as a function of distance from the correct step (e.g., a U that follows a U would be a -1 error, and an R that follows a U would be a $+1$ error). Error bars are graphical significance tests for contrasts between neighboring distances.

tween neighboring distances on the abscissa. A contrast is significant if the data points to the immediate left and right of the error bar lie (vertically) outside the bar.

We first performed omnibus Distance (6) \times Context (2) ANOVAs for both experiments. (We included group as a factor in the Experiment 1 design, but few contrasts involving group were significant, and we detected no systematic pattern, so for clarity we do not report them.) In Experiment 1, there were significant main effects of distance, $F(5, 990) = 76.0, p < .001, \eta_p^2 = .277$, and context, $F(1, 198) = 102.9, p < .001, \eta_p^2 = .342$, and there was a

significant interaction, $F(5, 990) = 49.2, p < .001, \eta_p^2 = .199$. In Experiment 2, there were again significant main effects of distance, $F(5, 495) = 21.5, p < .001, \eta_p^2 = .179$, and context, $F(1, 99) = 13.6, p < .001, \eta_p^2 = .121$, and a significant interaction, $F(5, 495) = 6.6, p < .001, \eta_p^2 = .063$.

To follow up the Distance \times Context interactions, we performed separate ANOVAs for each context. The main effects of distance were significant for each context in Experiment 1, $F(5, 990) = 63.2, p < .001, \eta_p^2 = .242$ for post-interruption and $F(5, 990) = 73.8, p < .001, \eta_p^2 = .271$ for baseline, and in Experiment 2,

Table 1
Inferential Statistics for Pair-Wise Contrasts Between Neighboring Levels of Distance

Distances	Experiment 1			Experiment 2		
	<i>F</i> (1, 198)	<i>p</i>	η_p^2	<i>F</i> (1, 99)	<i>p</i>	η_p^2
Post-interruption context						
-3, -2	0.2	.670	.001	6.1	.015	.058
-2, -1	86.2	<.001	.303	14.9	<.001	.131
-1, +1	17.8	<.001	.082	1.1	.292	.011
+1, +2	54.9	<.001	.217	16.8	<.001	.145
+2, +3	2.9	.091	.014	0.8	.368	.008
Baseline context						
-3, -2	9.5	.002	.046	1.3	.259	.013
-2, -1	19.8	<.001	.091	23.8	<.001	.194
-1, +1	67.1	<.001	.253	32.0	<.001	.245
+1, +2	81.9	<.001	.293	69.7	<.001	.413
+2, +3	13.4	<.001	.063	6.8	.011	.064

Note. Boldface values are significant at Bonferroni-corrected $\alpha = .025$.

$F(5, 495) = 12.5, p < .001, \eta_p^2 = .112$ for post-interruption and $F(5, 495) = 51.8, p < .001, \eta_p^2 = .344$ for baseline.

Finally, to follow up the main effects of distance, and to assess whether the distributions followed gradients, we performed pairwise contrasts between neighboring distances within each context. These contrasts are represented by the graphical significance tests in Figure 4 and are reported in full in Table 1. We used a Bonferroni-corrected significance level of $\alpha = .025$, based on $n = 2$ nonorthogonal contrasts per gradient. Sequence errors were significantly less frequent at Distance 2 than at Distance 1 in all eight cases (2 Directions \times 2 Contexts \times 2 Experiments). Sequence errors were significantly less frequent at Distance 3 than at Distance 2 in some cases. In the post-interruption context, the -3 versus -2 contrast was significant in Experiment 2. In the baseline context, the situation was clearer, with the -3 versus -2 contrast significant in Experiment 1 and the +2 versus +3 contrast significant in both experiments. In no cases were sequence errors at the greater distance significantly more frequent than those at the smaller distance.

As a simple test of whether anticipation and perseveration gradients were symmetrical, we included -1 versus +1 in the set of pairwise contrasts (Figure 4 and Table 1). These were significant in three of four cases (all but the post-interruption context of Experiment 2). However, the direction differed by context, with -1 errors numerically more frequent than +1 errors in the post-interruption contexts and less frequent in the baseline contexts. The interaction of distance (-1, +1) and context (post-interruption, baseline) was significant for Experiment 1, $F(1, 198) = 44.1, p < .001, \eta_p^2 = .182$, and Experiment 2, $F(1, 99) = 4.9, p = .029, \eta_p^2 = .047$.

Discussion. There was evidence for gradient-shaped distributions at least in the baseline context, where the frequency of sequence errors decreased through to the maximum distance of three steps in three of four possible cases (2 Directions \times 2 Experiments). Moreover, the pairwise comparisons we used to evaluate gradients were a conservative test, in that pairwise comparisons do not factor in trends across more than two levels. Considering just the numerical trends, errors decreased as distance increased in seven of eight possible cases (including both contexts).

A consideration in evaluating these gradients is that the step sequence was performed iteratively and did not involve that many

steps per iteration. Thus, the perseveration and anticipation gradients would have overlapped behaviorally, even if they reflected different underlying mechanisms. For example, a -3 error would have been behaviorally indistinguishable from a +4 error, even if the two had different origins. In previous work (Trafton et al., 2011), post-interruption perseveration gradients were more clearly delineated than they were here, probably because the step sequence was both longer and more differentiated between iterations.

The asymmetry between perseveration and anticipation gradients, reflected in the differences between -1 and +1 errors and how these interacted with context, suggests that they may not be generated by the same mechanisms, a point we reprise in the General Discussion.

Analysis 4: Reliability of Sequence Errors

In this analysis we examine the potential of our procedure for individual-difference research. Our analyses so far suggest that placekeeping operations in our task are relatively independent of the operations required for individual steps, and thus may represent general mechanisms that capture individual differences in sequential performance across a range of tasks. However, a prerequisite for predicting individual differences is reliability.

We assessed internal consistency reliability of sequence errors using Cronbach's alpha, with Blocks 1-4 as indicators. The reliability estimates were $\alpha = .54$ in Experiment 1 ($N = 195$) and $\alpha = .72$ in Experiment 2 ($N = 99$). (Six cases were excluded from this analysis because accuracy in one block was below threshold, as discussed in the Method section.) These estimates are in line with tests of working memory capacity and executive function (e.g., Friedman & Miyake, 2004). Ranges of error proportions were 0-.18 in Experiment 1 and 0-.12 in Experiment 2, which are large for a cognitive control task.

Our procedure is thus a candidate for assessing individual differences in sequential control, potentially at the latent level in conjunction with other tasks. A further use would be to test the hypothesized link between sequential control and general intelligence (Duncan, 2010; Duncan et al., 2008). A related use would be in testing whether our measure of sequential control and complex

span measures typically used as indices of executive attention (e.g., operation span; Engle, 2002; Kane & Engle, 2002) capture the same latent factor representing interference control. Finally, a simplified version of our procedure may help to characterize developmental and neuropsychological deficits in sequential control, such as action disorganization syndrome, more precisely than has been possible with existing measures (Cooper et al., 2005).

General Discussion

Short interruptions increased the rate of sequence errors on the post-interruption trial (Figure 2, top panels). Interruptions averaging 4.4 s (Experiment 1) increased the error rate by more than a factor of 3 relative to baseline, and interruptions averaging 2.8 s (Experiment 2) increased the error rate more than a factor of 2 relative to baseline. Nonsequence errors showed no interruption effects (Figure 2, middle panels), indicating that interruptions selectively affected the mental representations used in sequential control and did not affect performance on individual steps. Response latencies showed interruption effects only in Experiment 1 (Figure 2, lower panels), and these effects were smaller than those on sequence errors.

Here we address three theoretical questions: why very short interruptions should increase the rate of sequence errors, what mechanisms might lead to the gradient distributions in Figure 4, and why sequence errors were more sensitive to interruption effects than response latencies were.

Theoretical accounts of interruption effects have often rested on decay of task-relevant representations during the interruption (Altmann & Trafton, 2002; Hodgetts & Jones, 2006a, 2006b). Decay is arguably a pervasive architectural process (Altmann & Gray, 2008; Altmann & Schunn, 2012) and can be quite rapid (Campoy, 2012; Muter, 1980). Here, however, decay must have interacted with other factors to cause the interruption effects we observed, given that interruptions in Experiment 2 took no more time, on average, than a step of the primary task. One candidate factor is simply that processing during the interruption was focused on information not related to the primary task.

A second candidate factor is that the primary task involved substantial interference from irrelevant representations competing for a role in any given operation. Under such conditions, the challenge for executive processes is to focus on the correct element for each operation, and to navigate to the next element when that becomes relevant (see, e.g., Kane & Engle, 2002). We suggest that this navigation is supported by activation spreading from the focus of attention through links connecting the elements of the task-relevant representation—a mechanism that plays a core role in architectural models of working memory (e.g., Anderson & Lebiere, 1998; Cowan, 1997). To account for sequential control specifically, we assume that activation spreads from the current step in the sequence to its successor when the current step is in the focus of attention. When an interruption displaces the current step from the focus of attention, the flow of activation spreading to the next step is cut off, allowing the next step to decay.

Spreading activation could account for anticipation gradients, if it spreads to some extent beyond a single step forward in the sequence (Altmann & Trafton, 2007). A pending step would then be primed by its predecessors as its turn approaches, gradually becoming more active and thus more likely to intrude on the step selections leading up to it. A different mechanism may account for

perseveration gradients. In a model developed by Trafton et al. (2011), the system selects the next step based on memory for the last step it performed, and perseveration gradients arise because memories for performed steps decay and thus intrude less often the older they are. The possibility that different mechanisms are responsible for anticipation and perseveration gradients is consistent with the various asymmetries in their empirical forms in Figure 4.

In contrast, a single mechanism usually accounts for error gradients in serial recall and related tasks. The most common mechanism is a “context signal” that fluctuates such that neighboring items occasionally swap positions (e.g., Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999). Another mechanism is the blurring of item representations that occurs with a recurrent network (Botvinick & Plaut, 2006). For any single mechanism to account for our data, the challenge might be whether it can reproduce the different asymmetries evident in Figure 4.

A remaining question is why response latencies were less sensitive to interruption effects than sequence errors were. A similar dissociation arises in task-switching procedures, where “switch cost” registers in terms of errors but not latencies under conditions in which the current task must be retrieved from memory (Altmann, 2004; Altmann & Gray, 2008). In a model that reproduces this effect (Altmann & Gray, 2008), retrieval latency depends on the activation of the most active item, whether or not this is the target, whereas retrieval accuracy depends on the activation difference between the target and distractors. Thus, under conditions of high interference, with many active distractors, retrieval latency is relatively insensitive to a decrease in activation of just the target. On this account, when all steps in the sequence are active—due to recent use, priming from overlapping meanings of task elements, and perhaps other factors—and an interruption cuts off priming to the correct step, retrieval accuracy will be the main casualty.

This theoretical account of our data is largely assembled from existing constructs, including attentional focus, spreading activation, and decay and interference. However, the proposed configuration of these components is new and has many moving parts to potentially interact, and our empirical data offer some strong constraints. Thus, an important step in future work will be to develop an integrated representation of this account and test its sufficiency to reproduce behavioral measures through computational and/or formal modeling.

In conclusion, our core empirical finding is that when someone is momentarily interrupted or distracted and then returns to their task, they may do so without obvious hesitation, but with an increased chance of resuming at a different point in their train of thought than they might have otherwise. This contextual jitter—being taken out of the moment and landed back in a slightly different place—may be why even momentary interruptions can seem jarring when they occur during a cognitively engaging activity.

References

- Altmann, E. M. (2004). Advance preparation in task switching: What work is being done? *Psychological Science*, *15*, 616–622. doi:10.1111/j.0956-7976.2004.00729.x
- Altmann, E. M. (in press). Fine-grain episodic memory processes in cognitive control. *Zeitschrift für Psychologie*.
- Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, *115*, 602–639. doi: 10.1037/0033-295X.115.3.602

- Altmann, E. M., & John, B. E. (1999). Episodic indexing: A model of memory for attention events. *Cognitive Science*, 23, 117–156. doi:10.1207/s15516709cog2302_1
- Altmann, E. M., & Schunn, C. D. (2012). Decay versus interference: A new look at an old interaction. *Psychological Science*, 23, 1435–1437. doi:10.1177/0956797612446027
- Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science*, 26, 39–83. doi:10.1207/s15516709cog2601_2
- Altmann, E. M., & Trafton, J. G. (2007). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin & Review*, 14, 1079–1084. doi:10.3758/BF03193094
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Erlbaum.
- Ariga, A., & Lleras, A. (2011). Brief and rare mental “breaks” keep you focused: Deactivation and reactivation of task goals preempt vigilance decrements. *Cognition*, 118, 439–443. doi:10.1016/j.cognition.2010.12.007
- Botvinick, M. M., & Bylsma, L. M. (2005). Distraction and action slips in an everyday task: Evidence for a dynamic representation of task context. *Psychonomic Bulletin & Review*, 12, 1011–1017. doi:10.3758/BF03206436
- Botvinick, M. M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, 111, 395–429. doi:10.1037/0033-295X.111.2.395
- Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review*, 113, 201–233. doi:10.1037/0033-295X.113.2.201
- Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, 107, 127–181. doi:10.1037/0033-295X.107.1.127
- Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, 106, 551–581. doi:10.1037/0033-295X.106.3.551
- Byrne, M. D., & Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science*, 21, 31–61. doi:10.1207/s15516709cog2101_2
- Campoy, G. (2012). Evidence for decay in verbal short-term memory: A commentary on Berman, Jonides, and Lewis (2009). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1129–1136. doi:10.1037/a0026934
- Cooper, R. P., Schwartz, M. F., Yule, P., & Shallice, T. (2005). The simulation of action disorganisation in complex activities of daily living. *Cognitive Neuropsychology*, 22, 959–1004. doi:10.1080/02643290442000419
- Cooper, R. P., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, 17, 297–338. doi:10.1080/026432900380427
- Cooper, R. P., & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review*, 113, 887–916. doi:10.1037/0033-295X.113.4.887
- Cowan, N. (1997). *Attention and memory: An integrated framework*. Oxford, England: Oxford University Press.
- Duncan, J. (2010). *How intelligence happens*. New Haven, CT: Yale University Press.
- Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., . . . Nimmo-Smith, I. (2008). Goal neglect and Spearman’s g: Competing parts of a complex task. *Journal of Experimental Psychology: General*, 137, 131–148. doi:10.1037/0096-3445.137.1.131
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11, 19–23. doi:10.1111/1467-8721.00160
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, 133, 101–135. doi:10.1037/0096-3445.133.1.101
- Grundgeiger, T., Sanderson, P. M., MacDougall, H. G., & Venkatesh, B. (2010). Interruption management in the intensive care unit: Predicting resumption times and assessing distributed support. *Journal of Experimental Psychology: Applied*, 16, 317–334. doi:10.1037/a0021912
- Hodgetts, H. M., & Jones, D. M. (2006a). Contextual cues aid recovery from interruption: The role of associative activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1120–1132. doi:10.1037/0278-7393.32.5.1120
- Hodgetts, H. M., & Jones, D. M. (2006b). Interruption of the Tower of London task: Support for a goal activation approach. *Journal of Experimental Psychology: General*, 135, 103–115. doi:10.1037/0096-3445.135.1.103
- Hollands, J. G., & Jarmasz, J. (2010). Revisiting confidence intervals for repeated measures designs. *Psychonomic Bulletin & Review*, 17, 135–138. doi:10.3758/PBR.17.1.135
- Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, 9, 637–671. doi:10.3758/BF03196323
- Li, S. Y. W., Blandford, A., Cairns, P., & Young, R. M. (2008). The effect of interruptions on postcompletion and other procedural errors: An account based on the activation-based goal memory model. *Journal of Experimental Psychology: Applied*, 14, 314–328. doi:10.1037/a0014397
- Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, 57, 203–220. doi:10.1037/h0087426
- Monk, C. A., Trafton, J. G., & Boehm-Davis, D. A. (2008). The effect of interruption duration and demand on resuming suspended goals. *Journal of Experimental Psychology: Applied*, 14, 299–313. doi:10.1037/a0014402
- Muter, P. (1980). Very rapid forgetting. *Memory & Cognition*, 8, 174–179. doi:10.3758/BF03213420
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1–18). New York, NY: Plenum Press.
- Reason, J. (1990). *Human error*. New York, NY: Cambridge University Press. doi:10.1017/CBO9781139062367
- Schwartz, M. F., Reed, E. S., Montgomery, M. W., Palmer, C., & Mayer, N. H. (1991). The quantitative description of action disorganisation after brain damage: A case study. *Cognitive Neuropsychology*, 8, 381–414.
- Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences*, 34, 771–797. doi:10.1111/j.1540-5414.2003.02292.x
- Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies*, 58, 583–603. doi:10.1016/S1071-5819(03)00023-5
- Trafton, J. G., Altmann, E. M., & Ratwani, R. (2011). A memory for goals model of sequence errors. *Cognitive Systems Research*, 12, 134–143. doi:10.1016/j.cogsys.2010.07.010
- Trafton, J. G., Jacobs, A., & Harrison, A. M. (2012). Building and verifying a predictive model of interruption resumption. *Proceedings of the IEEE*, 100, 648–659. doi:10.1109/JPROC.2011.2175149
- Zijlstra, F. R. H., Roe, R. A., Leonora, A. B., & Krediet, I. (1999). Temporal factors in mental work: Effects of interrupted activities. *Journal of Occupational and Organizational Psychology*, 72, 163–185. doi:10.1348/096317999166581

Received April 4, 2012

Revision received October 26, 2012

Accepted October 26, 2012 ■