Programming of action timing cannot be completed until immediately prior to initiation of the response to be controlled

Stuart T. Klapp¹ · Dana Maslovat²

Published online: 8 June 2020
© The Psychonomic Society, Inc. 2020

Abstract

We consider, in depth, findings from across 6 decades regarding generating a motor response in a simple reaction-time (i.e., delayed response) paradigm. The early results robustly show a response complexity effect whereby the more response elements, the longer the simple reaction time (RT). This effect is puzzling because it indicates that preparation of some aspect of a response cannot be completed prior to the simple RT interval even though a precue had identified the response in advance. Results obtained during the past 10 years indicate that this finding reflects a restriction in advance preparation of the code that controls timing of the onsets for elements making up the response. This code cannot be prepared prior to the simple RT interval even though the parameters to be achieved by the response can be established earlier so that the difficulty of their selection does not influence RT. The general principle that emerges from these studies is that the abstract goal for a response can be determined in advance, but programming of action timing cannot be completed until immediately prior to initiation of the response to be controlled. We suggest that this restriction may arise because limitations in working memory prevent temporary storage of the timing code even though the abstract response goal can be stored until needed. Regardless of its origin, the limitation regarding preparation of response timing represents a fundamental feature in control of motor action.

Keywords Motor control · Action timing · Response preparation · Response initiation · Simple reaction time · Cerebellum

A motor program is a “set of muscle commands that are structured before a movement sequence begins” (Keele, 1968, p. 387). Here, we consider evidence for postponement of generating the component of a motor program, which directly controls timing of onsets of the elements (chunks) of an action sequence. Although this aspect of programming can occur prior to the response, it is delayed until just before the response begins. This implies that the programmed representation of temporal structure is volatile—once it is generated it must be used immediately. This limitation is a constraint that must

be fulfilled by any viable theory of the manner in which timing is controlled.

This conclusion involves a distinction that was not included in earlier analyses (e.g., Klapp, 1995, 1996). This distinction contrasts selection and planning of the abstract goals for the action, including specification of timing parameters, versus generating the code that implements timing. Recent evidence reviewed in this article indicates that these processes can occur in separate stages. The stages are analogous in some ways to specifying a software source code in a programming language followed by compiling to generate a machine language (.exe) code. A major theme of our analysis concerns the compile function. The evidence indicates that, unlike typical software programs, the compiling function for a timing code cannot be completed until immediately prior to the moment that the code is to be used. In keeping with the computer analogy, one can think of this as akin to a “*just-in-time*” compiler, which compiles the required execution code at run time rather than in advance (see Aycock, 2003, for a more detailed description).

The temporal intervals within an action that are considered in our analysis are on the order of a few hundred milliseconds

✉ Stuart T. Klapp
Stuart.Klapp@csueastbay.edu

Dana Maslovat
dmaslovat@langara.ca

¹ Department of Psychology, California State University, East Bay, Hayward, CA, USA

² School of Kinesiology, University of British Columbia, War Memorial Gymnasium 210-6081 University Boulevard, Vancouver, British Columbia V6T 1Z1, Canada

(ms); intervals of this length are known to be controlled by the cerebellum (Ivry & Keele, 1989). Injury to the cerebellum does not influence timing of continuous movements, but it does disrupt timing of the onsets of action chunks (Spencer, Zelaznik, Diedrichsen, & Ivry, 2003). Our analysis corresponds to this timing function; it examines timing of the onsets of action components making up a sequence. This approach is consistent with the principle that complex actions are composed of a sequence of units or chunks (Klapp & Jagacinski, 2011). Evidence concerning timing of the beginning of each chunk in a sequence is emphasized with only secondary consideration of timing the details of kinematics and muscle activation within individual chunks. For example, in rapid speech articulation, the primary analysis focuses on timing of the onset of the words rather than on timing of articulatory events such as syllables within each word. Although the emphasis is on robust data with respect to this level of analysis, preliminary findings suggesting that the conclusions may generalize to features within the chunks are also considered.

Control of timing at the millisecond range considered here is often critical for successful action. For example, batting a baseball, playing the piano, and vehicular control require that action components must occur at the right moment. Tasks differ from each other with respect to the complexity of timing demands and the studies reviewed here include investigation of effects of varying this complexity as it relates to the time required to program the response.

Behavioral evidence that compiling the timing code is postponed until response initiation

The early studies considered in this section were not originally interpreted in relation to timing of the onsets of response chunks. However, as this research continued (1960 through 2019), it gradually became clear that these studies revealed a limitation with respect to advance compiling of the timing code.

Reaction-time paradigms

In reaction-time (RT) experiments, participants are required to respond to a “go” stimulus as quickly as possible by producing a specified response. The time from onset of this stimulus until the beginning of the corresponding response is identified as RT. Many of the experiments reviewed here measure RT as a function of parameters of the response. This contrasts with the more typical RT experiments, which vary difficulty of processes such as perception and decision-making.

A critical distinction involved in this research is between choice RT and simple RT paradigms (see Fig. 1). In the *choice*

RT paradigm, the response is specified by the stimulus that also marks the beginning of RT, the Response Cue/“Go” signal (see Fig. 1a). All processing must occur during RT after this stimulus. Our analysis focuses on a different situation, the *simple RT* paradigm (see Fig. 1b). Typically, the required response is varied randomly from trial to trial. The required response is specified during each trial by a precue; this is followed by a delay, called the foreperiod, and then a go stimulus, which signals that the participant is to initiate the response that had been precued. In this simple RT paradigm, some processing might occur after the precue is presented, during the foreperiod, rather than during RT. Many of findings to be considered relate to identification of those processes that can occur during the foreperiod and those that cannot, and thus must occur during simple RT.

Classical RT findings

Henry and Rogers (1960) Our analysis of RT results begins with a simple but frequently cited experiment (Henry & Rogers, 1960) that yielded a finding that is puzzling when considered carefully because it was observed in a simple RT paradigm; for a review of 50 years of related research, see Klapp (2010). Participants performed one of three response sequences—(A) lift a finger; (B) lift finger, then reach forward to contact a ball; or (C) lift finger, then do a series of movements involving two balls and a button. Simple RT, the time from the start of an auditory go stimulus until lift-off from the home position at the beginning of the response, increased as a function of response complexity (i.e., RT for A < RT for B < RT for C). This result indicates that aspects of action sequences can be planned (programmed) during RT immediately before response initiation. The authors’ description of this programming process was a rather vague analogy to computer technology circa 1960.

Sternberg et al. (1978) Comprehensive results that extend this type of finding to speech articulation and typing were reported by Sternberg, Monsell, Knoll, and Wright (1978). Simple RT from stimulus onset until the start of speech articulation increased with a slope of approximately 12 ms per word over a range of 1–5 words to be articulated. This effect was observed for pronunciation of single-digit numbers, days of the week (in correct sequence, randomly ordered, and repetition of the same word), and unrelated words. For a typing task, simple RT prior to typing the first letter in a sequence increased as a function of the number of letters in the sequence. However, there are some rather complex details in the findings for typing (see Fig. 5 of Sternberg et al., 1978).

Whereas Henry and Rogers (1960) systematically rotated through or blocked the response types, Sternberg et al. (1978) presented response types at random from trial to trial so that participants could not predict the required movement until the

Fig. 1 Timeline of events in the choice RT (**a**) and simple RT (**b**) paradigms. In choice RT (**a**), a stimulus, labeled Response Cue/"Go," identifies the required response and also indicates that this response should be made as soon as possible. The present analysis is focused on the simple RT paradigm (**b**) in which the response is identified by the response precue occurring at the beginning of a foreperiod; the response is delayed until the "go" stimulus appears. Because of the precue some processing might occur during the foreperiod rather than after the go

stimulus. Processing that occurs during the foreperiod would not influence simple RT. Our interpretation, also represented in **b**, is that response selection, including selection of the abstract timing goal, can occur during the foreperiod. However, timing preparation involving generation of the compiled timing code must be postponed to occur after the "go" signal and thus during RT. Hypothetical stages of processing which are not relevant to this interpretation are omitted from the figure

precue appeared on each trial. It is noteworthy also that Sternberg et al. (1978) observed the RT effect across the modalities of speech articulation and the manual action of typing; this effect also occurs for hand and arm movements (Henry & Rogers, 1960). Thus, the result that simple RT increases with response complexity is not restricted to a particular methodology or response modality.

An important feature of the speech articulation data from Sternberg et al. (1978) was that the slope of the relation between number of words to be articulated and simple RT did not differ between one-syllable and two-syllable words (see their Fig. 3). Thus, the RT effect reflects processing related to words (action chunks) taken as wholes; internal details such as the number of syllables per word are encapsulated inside the chunks and do not influence the amount of increase in RT per chunk.

The first action in the sequence in some of these experiments was constant across number of words to articulate or letters to type. This overcomes a potential confound in the Henry and Rogers (1960) experiment, in which the initial lift-off action may have differed with respect to force profile across the response sequences, so that the effect of response complexity on simple RT might have been due to intensity of the initial action rather than to sequence length (Anson, 1982; Fischman, Christina, & Anson, 2008).

Sternberg et al. (1978) attributed the increase in simple RT as a function of the number of words or letters to scanning through the sequence to locate the first item in the sequence. One problem with this interpretation is that the RT effect was observed with respect to repetition of the same item; in this case, scanning would not be required, because the chunks are interchangeable. Neither this scanning model nor the very different interpretation proposed by Henry and Rogers (1960) are involved in the present analysis because, as

explained next, these models fail to consider implications of using the simple RT paradigm.

These findings are puzzling

These results occur for simple RT Both of the classical papers (Henry & Rogers, 1960; Sternberg et al., 1978) and some of the subsequent discussions of these articles failed to emphasize the consequences of the fact that these experiments used the simple RT paradigm (see Fig. 1b). Because of the response precue in this paradigm, the processes these authors proposed to account for their findings would be expected to occur during the foreperiod rather than during RT (Klapp, 2010). In that case, simple RT would not depend on the time required to complete the processing, and thus RT would not depend on the complexity of the response. But it does! This result challenges the early interpretations and creates a puzzle.

This puzzle can be stated in more general terms. A relation between response complexity and simple RT implies that an aspect of programming the response must be postponed to be completed immediately prior to initiation of the response and thus during RT. It is puzzling that this processing is not completed earlier, during the foreperiod. It would seem that preprogramming could occur during the foreperiod because the response is identified by a precue appearing early in each trial. However, the result that simple RT depends on the complexity of the response suggests that this earlier preprogramming is not completed during the foreperiod. One goal of the present analysis is to understand why this does not happen.

Replications Although puzzling, the finding that simple RT increases as a function of the number of elements in the subsequent response has been replicated for speech articulation

(Klapp, 2003) and for key pressing (Canic & Franks, 1989; Immink & Wright, 2001, Experiment 2; Klapp, 1995; Magnuson, Wright, & Verwey, 2004; Wright, Black, Immink, Brueckner, & Magnuson, 2004). This relationship even occurs in a partial precue paradigm in which, on each trial, the same action element was repeated N times. The number of elements, N , was precued in advance (as in a simple RT paradigm), but the response to be repeated was specified by the go stimulus (Klapp, 2003, Experiment 5; Rosenbaum, Saltzman, & Kingman, 1984). Shin and Proctor (2018, Experiment 3) explored the difference in simple RT between one-press and three-press movements across foreperiod intervals (response cue to go signal) ranging from 0 to 2,000 ms and found that the effect of requiring one versus three presses on simple RT was present throughout this range. Another compelling replication involved examination of response complexity effects using a loud (>120 dB) startling stimulus as the go signal on a small proportion of the trials. For both startle and nonstartle trials, simple RT increased between one-press to three-press movements even though overall RT was much shorter on startle trials (Maslovat, Klapp, Jagacinski, & Franks, 2014, Experiment 1). This experiment also demonstrated that, whereas RT is typically measured as the time from the start of stimulus to the beginning of the overt response, similar results also occur for premotor RT measured from the start of the stimulus to initial EMG activity. Other recent studies that include replication of this effect are considered in subsequent sections. The result that simple RT increases as a function of the number of elements in a motor response, first reported by Henry and Rogers (1960), is a robust and reliable finding.

Participants are not ignoring the precue The puzzle regarding simple RT might be resolved by assuming that participants may have ignored the precue and avoided preparation during the foreperiod even though they could have done so. Instead, they might opt to wait until the go signal appears at the beginning of the RT interval and then prepare to respond during that interval. This strategy of ignoring the precue appears to be highly unlikely for a variety of reasons. Not only is simple RT shorter than choice RT, studies have also shown that RT decreases as more information about the upcoming response is precued (Maslovat, Chua, Klapp, & Franks, 2018; Rosenbaum, 1980). The most logical explanation for this result is that more precued information allows for more advance preparation during the foreperiod, thereby reducing the need for processing during RT. Another relevant finding is that increasing the predictability of the time of occurrence of the go signal results in an overall decrease in RT, a result attributed to greater levels of advance preparation, yet response complexity effects still occur and are of similar magnitude to those seen with an unpredictable go signal (Maslovat, Klapp, Forgaard, Chua, & Franks, 2019). Findings discussed in the

next section demonstrate that participants typically prepare for one aspect of the response during the foreperiod in the simple RT paradigm; this also indicates that the precue stimulus is not ignored.

Distribution of processing between foreperiod and simple RT

We now begin a review of RT data bearing on resolving the puzzle of why does simple RT increase with response complexity even though the response has been precued. These data enable a much more detailed interpretation than in previous analyses (e.g., Klapp, 1995, 1996). They indicate that only one aspect of response complexity is responsible for the relation between response complexity and simple RT. This conclusion implies that although some response programming must occur during simple RT, programming of other aspects can be completed during the foreperiod so that it does not influence simple RT (see Fig. 1b). The data considered in this section do not identify which response aspects can and cannot be planned in advance; that issue is addressed in the next section.

This conclusion that processing is distributed between foreperiod and simple RT is based on findings using two different approaches. One of these is to vary the foreperiod duration (time between the response precue and go signal), which in turn affects the amount of time available for advance processing. Using this approach, Alouche, Sant'Anna, Biagioni, and Ribeiro-do-Valle (2012) found that when the foreperiod duration was very short (300 ms), simple RT increased strongly as a function of response complexity. This strong effect can be attributed to the lack of sufficient time in the foreperiod to enable processing in advance so that all programming occurred during RT. However, as foreperiod duration was increased to 1,000 ms or 2,000 ms, response complexity effects on simple RT were weakened, but still present. This result is evidence that with a longer foreperiod, some preprogramming occurred during the foreperiod rather than during RT (thus, the reduction in the effect of response complexity on RT), yet some aspect of preparation did not occur in advance and instead occurred during the simple RT interval (thus, response complexity effects on RT, although reduced, were still present).

This conclusion is also supported by results from a complementary paradigm in which the participants rather than the experimenter determined the foreperiod duration; participants signaled an end to the foreperiod when they were ready to proceed. This self-determined foreperiod, the so-called study time, is assumed to reflect the time required to complete advance processing for some aspect of responding. This interval increased in duration when there were more response chunks (Magnuson, Robin, & Wright, 2008), indicating that additional time is required to prepare for more complex responses.

Although this finding suggests that preparation of some aspect of the response occurs during the foreperiod, simple RT nevertheless also increased as a function of response complexity even though participants were given as much time as they needed to prepare for responding. These results for participant-determined foreperiod duration converge with those for experimenter-determined duration to indicate that one aspect of advance planning can occur during the foreperiod and yet other another aspect is not completed until during RT. The next section reviews findings that indicate which type of processing occurs in each interval.

Response selection during the foreperiod and compiling the timing code during simple RT

We next consider evidence that response selection is the process that can be completed during the foreperiod prior to the RT interval. Variables that influence the time required for response selection include stimulus–response compatibility (Fitts & Seeger, 1953) and the number of alternative responses (Hick, 1952). In Experiments 1 and 2 of Shin and Proctor (2018), the effect of these variables on simple RT, which was present if the foreperiod was short, vanished if the foreperiod delay was 500 ms or longer. Specifically, their Fig. 2 indicates that the effect of stimulus–response compatibility, whether the stimulus signaled the response location directly by the location of the stimulus or indirectly by the mapping of stimulus color to response location, was strong with a short foreperiod, but vanished with a longer foreperiod. A converging result is that their Fig. 3 shows that the difficulty of response selection, here manipulated in terms of number of alternative responses, produced a strong effect on simple RT when the foreperiod was short; this vanished for long foreperiods. These results indicate that although response selection must occur during simple RT if the foreperiod is too short to permit it to occur earlier, selection can occur in the foreperiod rather than during RT if the foreperiod is longer. By contrast to these findings regarding response selection, their Fig. 5 shows that the effect of the number of chunks in the response (response complexity) on simple RT was robust even at a foreperiod of 2,000 ms. Taken together, these findings indicate that, whereas response selection can be completed during the foreperiod, an aspect of preparation related to the number of chunks in the response cannot, and thus must occur during simple RT. But how can we characterize the aspect of processing that is restricted in this way?

The next study to be considered reinforces the conclusion that response selection can occur during the foreperiod. It also leads to a very critical conclusion related to compiling the timing code—a process described in the Introduction. The conclusion is that compiling is the process that is restricted so that it must occur during simple RT. Maslovat, Chua, Klapp, and Franks (2016) examined self-determined

foreperiod duration (study time) and simple RT using a button-pressing task. For a sequence of three-press responses, selection difficulty was either easy (all presses using the same hand) or difficult (different hand effectors across the presses). Timing difficulty was either easy (same delay between Presses 1 and 2 as between Presses 2 and 3) or difficult (different delays); total movement time was equivalent. Increasing the difficulty of response selection increased the overall participant-determined foreperiod (study time), but had little effect on simple RT. This finding converges with the results from Experiments 1 and 2 of Shin and Proctor (2018) to indicate that response selection can occur during a sufficiently long foreperiod. Conversely, increasing the difficulty of timing of press onsets increased simple RT, but not study time. This finding indicates that compiling the timing code does not occur in the foreperiod, but instead must be delayed until just prior to response initiation so that compiling occurs during simple RT.

This conclusion regarding the necessity of compiling the timing code during simple RT is in line with earlier findings that simple RT is shorter prior to a less complex equal-interval (isochronous) pattern of finger tapping compared with a more complex unequal-interval (nonisochronous) pattern (van Donkelaar & Franks, 1991) and that simple RT increases as a function of the number of taps only if stringent demands for temporal accuracy are imposed (Franks & van Donkelaar, 1990).

Does the principle that response selection can occur during the foreperiod apply to selection of the timing goal even though the timing code is not compiled until during RT? The answer seems to be yes. Maslovat et al. (2018) found that a precue of the timing structure reduced overall RT (indicating selection of the abstract timing goal during the foreperiod), but that RT still increased with the complexity of the timing structure (indicating that compiling of the selected timing occurs later, during RT).

Compiling the timing code is tied to response initiation

Two versions of the conclusion that compiling the timing code must be delayed until RT in the simple RT paradigm can be considered. Although compiling is typically delayed until the go stimulus, perhaps it could be completed in anticipation of this signal if its time of occurrence is predictable. If this occurs, the effect of timing complexity on simple RT would be eliminated. On the other hand, compiling may be possible only in conjunction with response initiation. In that case, the complexity of timing would influence simple RT even if the time of occurrence of the go stimulus is predictable. To distinguish between these alternatives, Maslovat et al. (2019) manipulated the predictability of the duration of the foreperiod prior to occurrence of the go stimulus. This foreperiod was

nonpredictable (random 2,500–3,500 ms), constant and relatively long (3,000 ms), constant and shorter (1,000 ms), or constant at 3,000 ms with countdown stimuli prior to the go stimulus. In all conditions, simple RT was shortest prior to a one-press action compared with three presses and, comparing the three-press actions, shorter prior to the three-press action with simplest timing of press onsets. Thus, making the time of occurrence of the go stimulus more predictable did not eliminate the effect of timing complexity on simple RT. This result suggests that compiling the timing code can occur only in conjunction with initiation of the response.

Summary of findings for simple RT prior to multichunk actions

The data reviewed above indicate that preparation of multichunk actions occurs in two phases. Phase 1 involves selection and planning the abstract goals of the response; this includes selection of response chunks and of timing parameters. This phase can be completed during the foreperiod of the simple RT paradigm (rather than during RT) and the result retained until it is needed. Phase 2 is compiling the code that implements the previously selected goal for response timing. This processing must be postponed to be completed immediately prior to producing the response. Thus, compiling the timing code is completed during simple RT and thereby becomes part of the initiation process.¹ Future directions for research may include use of neurological measures to confirm the two-phase timeline of this model and to provide additional insight into the brain structures associated with each phase.

Findings for choice RT prior to multichunk actions

The primary analysis considered in the present article, as summarized above, involves the simple RT paradigm for multichunk responses. Findings for choice RT, although not the focus of this article, have presented a challenge for interpretation. Given that simple RT is known to increase as a function of the number of chunks in the response, it is rather surprising that choice RT often does not also show this relationship. An earlier hypothesis to explain this finding was that when the response is not known in advance (choice RT paradigm), processing with respect to the sequence of chunks

occurs in parallel with other processes. If these occur at a slower rate, any effect of the number of chunks on choice RT could be obscured (Klapp, 1996, p. 30). Evidence for this hypothesis is limited; recent findings suggest a different view. Rather than due to parallel processing, the lack of an increase in choice RT when the number of chunks is increased may be attributed to some of this processing occurring online during the response rather than during RT. This interpretation is supported by a finding that the between-chunk intervals are longer in the choice RT paradigm compared with the simple RT paradigm (Maslovat et al., 2014). This indicates online control during intervals between chunks in the choice RT paradigm.

Hierarchy of control

Motor actions are typically organized as a sequence of chunks (Klapp & Jagacinski, 2011). Control is in the form of a hierarchy in which the upper level controls the chunks taken as units and the lower level controls features internal to each chunk.² At each level, both simple RT and choice RT have been measured as a function of the parameters of the response that follows (for a review, see Klapp, 2010). With respect to the upper level, simple RT depends on the complexity of the response to follow (e.g., number of chunks in the response), but choice RT is relatively independent of response complexity. These findings were described above. By contrast, at the lower level of the hierarchy, this RT pattern becomes reversed; choice RT depends on the complexity of a single chunk (e.g., the number of syllables to be articulated for a single word), but simple RT is relatively independent of this variable.

This pattern of RT findings with respect to the lower level of the hierarchy has been a component of the analysis of motor programming for many years (e.g., Klapp, Anderson, & Berrian, 1973; see Klapp, 2010, for a review). It is interpreted by assuming that, whereas choice RT includes the time to complete most stages of within-chunk programming, this can occur during the foreperiod prior to RT in the simple RT paradigm because the response is specified by a precue. Thus, the effect of increasing within-chunk complexity is stronger for choice RT than for simple RT. This hypothesis of preprogramming during the foreperiod is consistent with the finding that, in the simple RT paradigm, self-determined foreperiod (study time) is longer prior to a more complex chunk (Immink & Wright, 2001). However, the analysis to be considered below indicates that, contrary to earlier

¹ This interpretation of the relationship between simple RT and the complexity of the response to follow is the result of empirical findings developed in the last 10 years. As reviewed in the text, these findings lead to the new distinction between Phases 1 and 2 and to the conclusion that the effect of number of chunks on simple RT is due to timing of the onsets of the chunks during Phase 2. Previously, the possibility that response timing might, in some way, be responsible for this relation was only briefly mentioned as a hypothesis to be tested (Klapp, 2010).

² The terms SEQ and INT have been used to identify the hierarchical levels of control (e.g., Klapp, 1995, 1996, 2003). SEQ referred to sequencing chunks at the upper level and INT referred to processing related to internal features of individual chunks at the lower level. Although the notion of hierarchical control is preserved in the present analysis, these terms are not used because they have often been identified with older theoretical interpretations that do not fit the recent empirical findings.

interpretations, simple RT may be influenced (to a small extent) by the time required to compile timing for events within a chunk.

The role of practice and training

The organization of a response sequence into chunks is flexible and can become altered by practice or training. After practice, an action typically becomes represented as fewer but longer chunks. This change in organization would reduce the burden of control at the upper level of the hierarchy. Evidence in support of this role of practice is based on the general principle that simple RT depends on the number of chunks in the response. For example, in a button-pressing task, at early stages of practice each press appears to be represented as a separate chunk because simple RT increases as a function of the number of presses (which is the same as the number of chunks). After practice, the chunks often become longer so that each chunk includes more than one press. Thus, the slope of the relation between simple RT and number of presses is reduced with practice (Klapp, 1995; Wright et al., 2004). A similar result has been reported for arm movements (Maslovat, Hodges, Chua, & Franks, 2011). Whereas these results indicate that chunks can become longer with practice, the exact pattern of segmentation of very complex actions often varies across participants (Verwey & Eikelboom, 2003). Converging evidence for flexibility of organization of an action sequence is that training can be used to influence chunk size. This was demonstrated in a task requiring articulation of pseudowords that conform to English usage but are not actual words (Klapp, 2003). After training that emphasized separate articulation of each syllable, simple RT increased as a function of the number of syllables as expected if each chunk contains a single syllable. However, after training that encouraged combining several syllables into one pseudoword, simple RT was relatively independent of the number of syllables; this is expected if the syllables are integrated into one word so that there is only one chunk.

The conclusion about timing chunk onsets may generalize to timing within-chunk events

The conclusion that compiling the timing code must be postponed until immediately prior to responding was based on findings for multichunk actions. This represents the upper level of the hierarchy of control. In this section, we consider the possibility that this conclusion may apply also at the lower level of the hierarchy and thus to within-chunk kinematics.

Previous findings reconsidered

We now consider the possibility that, just as timing of onsets of the chunks that form a sequence must be compiled immediately prior to response initiation, timing of the onsets of internal features within a chunk may also need to be compiled just prior to initiation of a single chunk. If so, simple RT should increase as a function of the complexity of timing of events within a chunk. As indicated above, previous analyses suggested that simple RT does not vary in this way. However, closer examination reveals that a small nonsignificant trend in the direction of longer simple RT prior to more complex single chunks, although not noticed earlier, was present in most of the relevant experiments. We have now identified 17 comparisons in which simple RT was studied as a function of the complexity of timing of events within a chunk (see Table 1). These comparisons involved the number of syllables in a single word to be articulated or whether a single button-press response did or did not include a delay component between press and release. Of these 17 comparisons, 14 showed a trend toward longer simple RT prior to a single chunk for which the internal timing was more complex. A result that occurs in 14 out of 17 comparison is significant by sign test, $p = .006$; the average difference in simple RT was 4 ms. However, only one of the comparisons, that reported by Sternberg et al. (1978) for speech articulation, was statistically significant considered in isolation. Thus, in contrast to the robust and consistent findings regarding timing of the overall sequence of chunks, the tentative conclusion that simple RT depends on the complexity of a single chunk needs further verification.

Any effect for simple RT due to within-chunk timing complexity (lower level of the hierarchy) is small compared with the effect due to complexity of the timing of chunk onsets (upper level of the hierarchy). This difference might be attributed to the durations of the timed intervals; intervals within a single chunk may be shorter than intervals between chunks. Perhaps the time required to compile timing is shorter if the interval to be timed is briefer; this possibility could also be examined in future research.

Coordination of compile processes within and across chunks

The possibility that the compiled code for timing of events within a single response chunk must be generated prior to initiation of each chunk of a multiple-chunk action raises issues concerning when this compile process can occur. The conclusion that a timing code must be compiled immediately prior to response initiation implies that, except for the first chunk in the sequence, the internal timing with respect to each chunk in the sequence must be compiled during the between-chunk interval just before initiation of the chunk. Compiling with respect to the first chunk would occur during RT.

Table 1 Summary of previous studies examining simple reaction-time (RT) effects due to within-chunk complexity differences

Reference	Within-chunk complexity manipulation	RT complexity effect
Klapp (1995)	Press-release vs. press-hold-release	+5.5 +7
Klapp (2003)	Syllables in one word, per syllable	+2.5
Klapp et al. (1973)	Syllables in one word, per syllable	+1.7
Klapp et al. (1974)	Press-release vs. press-hold-release	(−1)
Klapp et al. (1981)	Syllables in one word per syllable	(−1) +3 (−12)
Magnuson et al. (2008)	Press-release vs. press-hold-release	+3.25 +5 +15.5 +5
Sternberg et al. (1978)	Syllables in one word, per syllable	+4.5
Wright et al. (2004)	Press-release vs. press-hold-release	+3.3 +3 +16.5 +7

This model is consistent with the Sternberg et al. (1978) data for speech articulation in which simple RT increases as a function of the number of words with a slope that is independent of the number of syllables per word; the intercept of this function is larger if the first word includes more syllables. This finding fits the model described in the above paragraph as follows: The intercept depends on the number of syllables in the first word because of the need to compile the timing of the onsets of articulation of the syllables making up that word during RT. The overall increase in simple RT as a function of the number of words is due to compiling the time of onsets of articulation of the words in the sequence. The internal complexity of the words does not influence the slope of this relationship, because compiling the timing of articulation of gestures (syllables) within words, except for the first word, occurs during between-chunk intervals rather than during RT.

If, as assumed in this model, compiling of timing for within-chunk events occurs during between-chunk intervals, the duration each of these intervals would be expected to increase when the time required to compile timing of the following chunk is longer. For various reasons it is unlikely that this prediction can be tested adequately. First, as noted earlier, the time required to compile timing of within-chunk gestures is expected to be small relative to the larger and more consistent effect observed at the higher level of the hierarchy. Thus, any change in duration of between-chunk intervals due to compiling the internal timing of the next chunk in a sequence would be small and difficult to detect. Second, the between-chunk intervals must conform to the timing parameters that were established for onsets of the chunks. These parameters are likely to specify intervals that are longer than the time required to compile internal timing of the subsequent chunk so that the effect of compiling would be

obscured. Third, the sequence of chunks may become restructured to combine several gestures into a single chunk. The role of practice and training to produce restructuring was described in a previous section. After restructuring, the duration of the interval between actions that became combined into one chunk would not reflect compiling time. These three considerations can lead to a very complex pattern of between-chunk intervals, especially at high levels of practice, as shown by the interresponse intervals of multielement responses in Table 2 of Immink and Wright (2001). Because our primary analysis concerns compiling the code for timing of chunk onsets (upper level of the hierarchy), verification of this model for the lower level of the hierarchy is not needed to support our main point.

Why must the timing code be compiled just prior to the moment that it is needed?

The conclusion that compiling the timing code for onsets of chunks must be delayed until immediately before response initiation is strongly supported by the research results we have considered; a smaller and less consistent effect seems to occur with respect to timing for within-chunk gestures. Postponement of compiling the timing code suggests that a processing function is missing, which could enable compiling to occur earlier. One possibility, to be considered in this section, is that the missing function involves working memory.

Working memory defined

The term “working memory” refers to a short-duration memory system that retains a limited amount of information in an

easily accessible state so that it can be used in cognitive processing. This has been identified with the classical notion of primary memory, which is in current awareness, in contrast to secondary memory, which has left awareness and must be retrieved (James, 1890). Some claims about working memory have been shown to be questionable. For example, Klapp, Marshburn, and Lester (1983) argued that the once popular notion that working memory corresponds to the span of seven plus or minus two items in serial recall, is not viable. The proposal that individual differences in cognitive ability may be attributed, in part, to differences in capacity of working memory has also been challenged (Burgoyne, Hambrick, & Altmann, 2019; Engle, 2018). By contrast, the hypothesis that there are multiple components of working memory (e.g., Baddeley & Hitch, 1974) has received continuing support (D’Esposito, 2007). The present analysis is consistent with this multiple-component viewpoint in the sense that it deals with only one of many aspects of working memory.

The role of working memory in the simple RT paradigm

Data considered earlier indicate that in the simple RT paradigm, the abstract response goal can be selected during the foreperiod prior to RT so that the difficulty of selection does not influence RT. For this advance selection to be effective, the selected response must be retained in a form of working memory until the go signal appears. Similarly, if there was effective working memory for the compiled timing code, it should be possible to compile this code during the foreperiod and retain it until the go signal appears so that compiling would not be needed during RT. If this was possible, the complexity of timing would not influence simple RT. However, our fundamental point is that simple RT does depend on the complexity of response timing. This conclusion holds even if the number of chunks held constant (Maslovat et al., 2016). This indicates that compiling of the timing code occurs during simple RT rather than during the foreperiod. The reason that preprogramming for timing does not occur might be that, unlike memory for response selection, there is no working memory function to retain the compiled timing code. Instead, the code is volatile and must be generated during simple RT.

But is it reasonable to assume that working memory can store abstract response goals, thereby permitting selection of the goal during the foreperiod, but that it cannot function in this way with respect to the compiled timing code? Perhaps it is. Selection of abstract response goals is likely to be a cortical function. There is strong evidence that working memory is also supported by structures in the cortex; working memory may be based on “recruitment” of cortical sensory regions (D’Esposito, 2007; Scimeca, Kiyonaga, & D’Esposito, 2018). Because processing for selection and retention of this selection both occur in the cortex, these functions may be easy to link. By contrast

compiling the timing code can be attributed to the cerebellum (Ivry & Keele, 1989), which may have only an indirect role in working memory (Spencer & Ivry, 2009). This anatomic separation of compiling (cerebellum) and retention (cortical) functions may make it difficult to store a compiled timing code while waiting for the correct moment to use it.

Cost and benefit of delaying compiling of timing code until response initiation

Regardless of the underlying cause, whether due to limited access to working memory or to some other factor, it is surprising that no solution to the restriction regarding compiling the timing code has evolved. This could be because this limitation has its costs, but it also is beneficial, and the benefit may be of more importance.

First, consider the cost. Compiling of timing immediately prior to response initiation obviously introduces a delay in responding—a few ms increase in simple RT. Delayed compiling may also be a factor in producing the psychological refractory period effect—a bottleneck that prevents preparation of a second response until preparation for the first response has been completed. A recent analysis of this effect (Klapp, Maslovat, & Jagacinski, 2019) indicates how the requirement to delay compiling of response timing can contribute to production of this bottleneck. The increase in RT and the duration of the bottleneck are both rather brief, definitely less than 1 second, and thus are unlikely to be noticeable in many everyday tasks. However, these delays can be important in some circumstances, such as athletic performance and vehicular control.

The benefit of preventing compiling until response initiation may be of greater significance than the cost due to these delays because compiling can act as a gate to hold back overt action even though the action has been planned in other respects. For example, in the simple RT paradigm, most of the planning can occur without risking automatic release of the action because release requires that the timing code must be compiled. As explained earlier, the compile function is part of the initiation process (Maslovat et al., 2019).

The determination of how a prepared response can be held in readiness but not initiated has recently received considerable attention and debate (Calderon, Gevers, & Verguts, 2018; Churchland, Cunningham, Kaufman, Ryu, & Shenoy, 2010; Kaufman, Churchland, Ryu, & Shenoy, 2014; Kaufman et al., 2010; Sinclair & Hammond, 2008). Our analysis suggests a very simple mechanism that would prevent unintentional release of a response. A prepared response is not initiated automatically because it is not actually possible to complete all aspects of preparation in advance. Completion of the final stage, compiling the timing code, cannot be accomplished until just prior to response initiation. This provides a gate to prevent premature responding.

Movement preparation and initiation are independent processes

The present conclusion corresponds to aspects of the influential proposal of Haith, Pakpoor, and Krakauer (2016, p. 3007) concerning “independence of movement preparation and movement initiation.” They emphasized that preparation and initiation can occur in independent time frames. The data from simple RT studies, reviewed here, support and extend this conclusion. Aspects of preparation, including selection of abstract goals, can occur during the foreperiod even though compiling the timing code, a component of initiation (Maslovat et al., 2019), is delayed until after the go stimulus.

Haith et al. (2016) also suggest that initiation processes (as measured in traditional RT experiments) are more “sluggish” than absolutely necessary because of an optional delay that can be removed by instructions or by using a startling stimulus as the go stimulus. In this manner, the authors question whether RT reflects the summed duration of processing stages as is commonly assumed (Donders, 1868/1969). While it is certainly true that RT can be reduced especially by startle, timing complexity still influences simple RT when responses of varying complexity are triggered at short latency by a startling stimulus (Maslovat et al., 2014). This indicates that, although RT latencies may indeed overestimate the minimal time required for response preparation, they are still indicative of differences in processing time due to such factors as response complexity. Thus, the approach in which *differences* in RT are used to assess *differences* in processing time is valid even though the overall mean RT as typically measured can overestimate the total time required for processing.

But what does the optional delay then tell us about processing needed to control responses? This situation can be analyzed by expanding our model of processing in the simple RT paradigm (see Fig. 1b) to incorporate this delay. Response selection and element sequencing occur after the precue and thus during the foreperiod, with the selected response held in readiness until the imperative go stimulus is presented. The optional delay can be inserted after the go stimulus; this becomes a component of RT. After the delay, the timing code is compiled in conjunction with initiation and the time needed for compiling is also a component of RT. The optional delay may relate to participants trying to maximize movement accuracy. Haith et al. (2016) used a forced RT condition in which participants were required to respond at various times following the target presentation, whether they felt that the response was ready or not. Instruction to respond more quickly reduced average RT and it also produced more errors, leading the researchers to suggest that the optional RT delay may function to reduce the risk of movement initiation prior to the response being fully and appropriately prepared. Importantly, the finding that the effect of increased timing complexity to require longer simple RT is present with or without this delay

indicates that compiling the timing code is an enduring and inseparable component of response initiation.

Conclusion

The findings reviewed here demonstrate that programming of action timing cannot be completed until immediately prior to initiation of the response to be controlled. Within the simple RT paradigm, abstract response goals, including the timing goal, are selected and planned during the foreperiod, but compiling the code that implements the timing goal must be postponed so that this occurs during RT. This mandatory postponement of processing seems counterintuitive, and yet, in retrospect, we can see that it is consistent with research results first reported in 1960. These results indicated that at least some aspect of response programming must be postponed to occur during simple RT. Now we know that this restricted aspect relates to compiling the code that implements response timing; other aspects of response programming are not subject to this limitation.

This postponement of processing with respect to response timing represents a fundamental constraint that needs to be incorporated into any theory of motor response timing. A theory of action timing must also account for the marked tendency for actions involving distinct effectors to be temporally coordinated to the extent that independently timed actions are nearly impossible to produce (Jagacinski et al., 2016; Klapp, 1979; Klapp, Nelson, & Jagacinski, 1998).

These two constraints, delay of compiling the timing code and lack of temporal independence, would arise if timing is controlled by a temporary neural network that is established just prior to the start of the response. It would utilize neural transmission delays to implement corresponding delays in initiation of the elements of an action sequence. Because the network is assumed to be temporary and volatile, this interpretation fits the constraint regarding postponing of compiling the timing code. This interpretation also fits the second constraint, lack of temporal independence, if one assumes that all actions are triggered by nodes within the same network. Although this idea is vague and preliminary, it may point in a useful direction. Regardless of what theory of response timing emerges to fit these constraints, it seems clear that timing is an important limitation in the production of rapid and coordinated actions.

References

- Alouche, S. R., Sant’Anna, G. N., Biagioni, G., & Ribeiro-do-Valle, L. E. (2012). Influence of cueing on the preparation and execution of untrained and trained complex motor responses. *Brazilian Journal of Medical and Biological Research*, 45(5), 425–435. <https://doi.org/10.1590/S0100-879X2012007500053>

- Anson, J. G. (1982). Memory drum theory: Alternative tests and explanations for the complexity effects on simple reaction time. *Journal of Motor Behavior*, 14(3), 228–246. <https://doi.org/10.1080/00222895.1982.10735276>
- Aycock, J. (2003). A brief history of just-in-time. *ACM Computing Surveys*, 35(2), 97–113. <https://doi.org/10.1145/857076.857077>
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8). New York, NY: Academic Press.
- Burgoyne, A. P., Hambrick, D. Z., & Altmann, E. M. (2019). Is working memory capacity a causal factor in fluid intelligence? *Psychonomic Bulletin & Review*, 26(4), 1333–1339. <https://doi.org/10.3758/s13423-019-01606-9>
- Calderon, C. B., Gevers, W., & Verguts, T. (2018). The unfolding action model of initiation times, movement times, and movement paths. *Psychological Review*, 125(5), 785–805. <https://doi.org/10.1037/rev0000110>
- Canic, M. J., & Franks, I. M. (1989). Response preparation and latency in patterns of tapping movements. *Human Movement Science*, 8(2), 123–139. [https://doi.org/10.1016/0167-9457\(89\)90013-4](https://doi.org/10.1016/0167-9457(89)90013-4)
- Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Ryu, S. I., & Shenoy, K. V. (2010). Cortical preparatory activity: Representation of movement or first cog in a dynamical machine? *Neuron*, 68(3), 387–400. <https://doi.org/10.1016/j.neuron.2010.09.015>
- D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, 362(1481), 761–772. <https://doi.org/10.1098/rstb.2007.2086>
- Donders, F. C. (1868/1969). Over de snelheid van psychische processen [On the speed of mental processes]. In W. G. Koster (Ed.), *Attention and performance II. Acta Psychologica* (Vol. 30, pp. 412–431). Amsterdam, Netherlands: North Holland. (Original work published 1868) [https://doi.org/10.1016/0001-6918\(69\)90065-1](https://doi.org/10.1016/0001-6918(69)90065-1)
- Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, 13(2), 190–193. <https://doi.org/10.1177/1745691617720478>
- Fischman, M. G., Christina, R. W., & Anson, J. G. (2008). Memory drum theory's C movement: Revelations from Franklin Henry. *Research Quarterly for Exercise and Sport*, 79(3), 312–318. <https://doi.org/10.1080/02701367.2008.10599494>
- Fitts, P. M., & Seeger, C. M. (1953). S–R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 46(3), 199–210. <https://doi.org/10.1037/h0062827>
- Franks, I. M., & van Donkelaar, P. (1990). The effects of demanding temporal accuracy on the programming of simple tapping sequences. *Acta Psychologica*, 74(1), 1–14. [https://doi.org/10.1016/0001-6918\(90\)90031-A](https://doi.org/10.1016/0001-6918(90)90031-A)
- Haith, A. M., Pakpoor, J., & Krakauer, J. W. (2016). Independence of movement preparation and movement initiation. *The Journal of Neuroscience*, 36(10), 3007–3015. <https://doi.org/10.1523/JNEUROSCI.3245-15.2016>
- Henry, F., & Rogers, D. (1960). Increased response latency for complicated movements and a “memory drum” theory of neuromotor reaction. *Research Quarterly of the American Association for Health, Physical Education, and Recreation*, 31(3), 448–458. <https://doi.org/10.1080/10671188.1960.10762052>
- Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4(1), 11–26. <https://doi.org/10.1080/17470215208416600>
- Immink, M. A., & Wright, D. L. (2001). Motor programming during practice conditions high and low in contextual interference. *Journal of Experimental Psychology: Human Perception and Performance*, 27(2), 423–437. <https://doi.org/10.1037/0096-1523.27.2.423>
- Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, 1(2), 136–152. <https://doi.org/10.1162/jocn.1989.1.2.136>
- Jagacinski, R. J., Rizzi, E., Kim, T. H., Lavender, S. A., Speller, L. F., & Klapp, S. T. (2016). Parallel streams vs. integrated timing in multi-limb pattern generation: A test of Korte's third law. *Journal of Experimental Psychology: Human Perception and Performance*, 42(11), 1703–1715. <https://doi.org/10.1037/xhp0000263>
- James, W. (1890). *Principles of psychology* (Vol. 1). New York, NY: Holt. <https://doi.org/10.1037/10538-000>
- Kaufman, M. T., Churchland, M. M., Ryu, S. I., & Shenoy, K. V. (2014). Cortical activity in the null space: Permitting preparation without movement. *Nature Neuroscience*, 17(3), 440–448. <https://doi.org/10.1038/nn.3643>
- Kaufman, M. T., Churchland, M. M., Santhanam, G., Yu, B. M., Afshar, A., Ryu, S. I., & Shenoy, K. V. (2010). Roles of monkey premotor neuron classes in movement preparation and execution. *Journal of Neurophysiology*, 104(2), 799–810. <https://doi.org/10.1152/jn.00231.2009>
- Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin*, 70(6, Pt. 1), 387–403. <https://doi.org/10.1037/h0026739>
- Klapp, S. T. (1979). Doing two things at once: The role of temporal compatibility. *Memory & Cognition*, 7, 375–381. <https://doi.org/10.3758/BF03196942>
- Klapp, S. T. (1995). Motor response programming during simple and choice reaction time: The role of practice. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5), 1015–1027. <https://doi.org/10.1037/0096-1523.21.5.1015>
- Klapp, S. T. (1996). Reaction time analysis of central motor control. In H. N. Zelaznik (Ed.), *Advances in motor learning and control* (pp. 13–35). Champaign, IL: Human Kinetics Publishers.
- Klapp, S. T. (2003). Reaction time analysis of two types of motor preparation for speech articulation: Action as a sequence of chunks. *Journal of Motor Behavior*, 35(2), 135–150. <https://doi.org/10.1080/00222890309602129>
- Klapp, S. T. (2010). Comments on the classic Henry and Rogers (1960) paper on its 50th anniversary: Resolving the issue of simple versus choice reaction time. *Research Quarterly for Exercise and Sport*, 81(1), 108–112. <https://doi.org/10.1080/02701367.2010.10599634>
- Klapp, S. T., Anderson, W. G., & Berrian, R. W. (1973). Implicit speech in reading reconsidered. *Journal of Experimental Psychology*, 100(2), 368–374. <https://doi.org/10.1037/h0035471>
- Klapp, S. T., & Jagacinski, R. J. (2011). Gestalt principles in the control of motor action. *Psychological Bulletin*, 137(3), 443–462. <https://doi.org/10.1037/a0022361>
- Klapp, S. T., Marshburn, E. A., & Lester, P. T. (1983). Short term memory does not involve the “working memory” of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General*, 112(2), 240–264. <https://doi.org/10.1037/0096-3445.112.2.240>
- Klapp, S. T., Maslovat, D., & Jagacinski, R. J. (2019). The bottleneck of the psychological refractory period effect involves timing of response initiation rather than response selection. *Psychonomic Bulletin & Review*, 26(1), 29–47. <https://doi.org/10.3758/s13423-018-1498-6>
- Klapp, S. T., Nelson, J. M., & Jagacinski, R. J. (1998). Can people tap concurrent bimanual rhythms independently? *Journal of Motor Behavior*, 30(4), 301–322. <https://doi.org/10.1080/00222899809601346>
- Klapp, S. T., Greim, D. M., & Marshburn, E. A. (1981). Buffer storage of programmed articulation and articulatory loop. Two names for the same mechanism or two distinct components of short-term memory? In Long, J. B. & Baddeley, A. D. (Eds.) *Attention and Performance IX*. Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 459–472

- Klapp, S. T., Wyatt, E. P., & Lingo, W. M. (1974). Response programming in simple and choice reactions. *Journal of Motor Behavior*, 6(4), 263–271. <https://doi.org/10.1080/00222895.1974.10735002>
- Magnuson, C. E., Robin, D. A., & Wright, D. L. (2008). Motor programming when sequencing multiple elements of the same duration. *Journal of Motor Behavior*, 40(6), 532–544. <https://doi.org/10.3200/JMBR.40.6.532-544>
- Magnuson, C. E., Wright, D., & Verwey, W. (2004). Changes in the incidental context impacts search but not loading of the motor buffer. *The Quarterly Journal of Experimental Psychology: Section A*, 57(5), 935–951. <https://doi.org/10.1080/02724980343000675>
- Maslovat, D., Chua, R., Klapp, S. T., & Franks, I. M. (2016). Independent planning of timing and sequencing for complex motor movements. *Journal of Experimental Psychology: Human Perception and Performance*, 42(8), 1158–1172. <https://doi.org/10.1037/xhp0000220>
- Maslovat, D., Chua, R., Klapp, S. T., & Franks, I. M. (2018). Preparation of timing structure involves two independent sub-processes. *Psychological Research*, 82(5), 981–996. <https://doi.org/10.1007/s00426-017-0877-3>
- Maslovat, D., Hodges, N. J., Chua, R., & Franks, I. M. (2011). Motor preparation and the effects of practice: Evidence from Startle. *Behavioral Neuroscience*, 125(2), 226–240. <https://doi.org/10.1037/a0022567>
- Maslovat, D., Klapp, S. T., Forgaard, C. J., Chua, R., & Franks, I. M. (2019). The effect of response complexity on simple reaction time occurs even with a highly predictable imperative stimulus. *Neuroscience Letters*, 704, 62–66. <https://doi.org/10.1016/j.neulet.2019.03.056>
- Maslovat, D., Klapp, S. T., Jagacinski, R. J., & Franks, I. M. (2014). Control of response timing occurs during the simple reaction time interval but on-line for choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, 40(5), 2005–2021. <https://doi.org/10.1037/a0037522>
- Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, 109(4), 444–474. <https://doi.org/10.1037/0096-3445.109.4.444>
- Rosenbaum, D. A., Saltzman, E., & Kingman, A. (1984). Choosing between movement sequences. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 119–134). Hillsdale, NJ: Erlbaum.
- Scimeca, J. M., Kiyonaga, A., & D'Esposito, M. (2018). Reaffirming the sensory recruitment account of working memory. *Trends in Cognitive Sciences*, 22(3), 190–192. <https://doi.org/10.1016/j.tics.2017.12.007>
- Shin, Y. K., & Proctor, R. W. (2018). Evidence for distinct steps in response preparation from a delayed response paradigm. *Acta Psychologica*, 191, 42–51. <https://doi.org/10.1016/j.actpsy.2018.08.010>
- Sinclair, C., & Hammond, G. R. (2008). Reduced intracortical inhibition during the foreperiod of a warned reaction time task. *Experimental Brain Research*, 186(3), 385–392. <https://doi.org/10.1007/s00221-007-1241-4>
- Spencer, R. M., & Ivry, R. B. (2009). Sequence learning is preserved in individuals with cerebellar degeneration when the movements are directly cued. *Journal of Cognitive Neuroscience*, 21(7), 1302–1310. <https://doi.org/10.1162/jocn.2009.21102>
- Spencer, R. M., Zelaznik, H. N., Diedrichsen, J., & Ivry, R. B. (2003). Disrupted timing of discontinuous but not continuous movements by cerebellar lesions. *Science*, 300(5624), 1437–1439. <https://doi.org/10.1126/science.1083661>
- Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York, NY: Academic Press.
- van Donkelaar, P., & Franks, I. M. (1991). The preparation and initiation of simple rhythmical patterns. *Human Movement Science*, 10(6), 629–651. [https://doi.org/10.1016/0167-9457\(91\)90020-X](https://doi.org/10.1016/0167-9457(91)90020-X)
- Verwey, W. B., & Eikelboom, T. (2003). Evidence for lasting sequence segmentation in the discrete sequence-production task. *Journal of Motor Behavior*, 35(2), 171–181. <https://doi.org/10.1080/00222890309602131>
- Wright, D. L., Black, C. B., Immink, M. A., Brueckner, S., & Magnuson, C. (2004). Long-term programming improvements occur via concatenation of movement sequences during random but not during blocked practice. *Journal of Motor Behavior*, 36(1), 39–50. <https://doi.org/10.3200/JMBR.36.1.39-50>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.