

# Speed, Accuracy, and Serial Order in Sequence Production

Peter Q. Pfordresher<sup>a</sup>, Caroline Palmer<sup>b</sup>, Melissa K. Jungers<sup>c</sup>

<sup>a</sup>*Department of Psychology and Institute for Music Research, The University of Texas at San Antonio*

<sup>b</sup>*Department of Psychology, McGill University*

<sup>c</sup>*Department of Psychology, Ohio State University, Newark*

Received 13 September 2005; received in revised form 14 June 2006; accepted 23 June 2006

---

## Abstract

The production of complex sequences like music or speech requires the rapid and temporally precise production of events (e.g., notes and chords), often at fast rates. Memory retrieval in these circumstances may rely on the simultaneous activation of both the current event and the surrounding context (Lashley, 1951). We describe an extension to a model of incremental retrieval in sequence production (Palmer & Pfordresher, 2003) that incorporates this logic to predict overall error rates and speed–accuracy trade-offs, as well as types of serial ordering errors. The model assumes that retrieval of the current event is influenced by activations of surrounding events. Activations of surrounding events increase over time, such that both the accessibility of distant events and overall accuracy increases at slower production rates. The model's predictions were tested in an experiment in which pianists performed unfamiliar music at 8 different tempi. Model fits to speed–accuracy data and to serial ordering errors support model predictions. Parameter fits to individual data further suggest that working memory contributes to the retrieval of serial order and overall accuracy is influenced in addition by motor dexterity and domain-specific skill.

*Keywords:* Sequence production; Memory; Motor control; Mathematical modeling; Speed–accuracy trade-offs

---

## 1. Introduction

Two topics that have dominated research in sequence production are the relation between accuracy and speed (the speed–accuracy trade-off; Woodworth, 1899) and the way in which events are retrieved from memory in the correct order (the serial order problem; Lashley, 1951).<sup>1</sup> Whereas accuracy and the speed–accuracy trade-off concern the overall likelihood of

---

Correspondence should be addressed to Peter Q. Pfordresher, University at Buffalo, State University of New York, Department of Psychology, Park Hall, Buffalo, NY 14260; e-mail: pqp@buffalo.edu or to Caroline Palmer, McGill University, Department of Psychology, 1205 Dr. Penfield Avenue, Montreal, Quebec H3A 1B1 Canada; e-mail: caroline.palmer@mcgill.ca

production errors,  $p(\text{err})$ , the problem of serial order concerns the likelihood of a given type of error ( $x$ ) when an error occurs,  $p(\text{err}_x | \text{err})$ . Most research has focused exclusively on one of the two issues, despite the fact that both involve a common data source and probably reflect a common processing constraint (cf. MacKay, 1982, 1987). We describe a model in which timing functions as that common constraint (the range model; Palmer & Pfordresher, 2003).

The core assumption of the range model is that performers use information about the surrounding context when retrieving events (notes and chords in music) during sequence production. In other words, events are represented contextually via the pattern of activation strengths across the entire sequence. Activation strengths vary as a function of the serial proximity and similarity of surrounding events to the current event, with the most active surrounding events being those that are proximal and similar. Time constrains the accessibility of events across distances in the original model, and thereby influences predicted patterns of serial ordering errors. Specifically, activations for distant events increase as production rate slows. Because event retrieval is contextual in nature, the extended model predicts that the activation strength of the current event results from the strengths of surrounding events. As a result of this contextual dependence, described in detail later, the strength of the current event increases relative to surrounding events as production rate slows. The extended model thus predicts a speed–accuracy trade-off that is tested in a new experiment reported here.

The range model, like other contextually based approaches, characterizes sequence production as an incremental process in which the planning and execution of multiple events overlap in time (Kempen & Hoenkamp, 1987; Smith & Wheeldon, 1999; Wheeldon, Meyer, & Smith, 2002; cf. McClelland, 1979). We consider planning (the preparation of to-be-produced events prior to their production) to arise from a process of *response preparation*, during which a set of possible events are activated. Response preparation may be (but is not necessarily) followed by *response selection*, which is the selection and execution of an overt response.<sup>2</sup> Whereas the original range model was primarily interested in response preparation (i.e., which events are activated at any given point in time), the extended model proposed here focuses on response selection (the likelihood that the correct event is selected). The extended range model proposes that response preparation and selection are mutually constrained by the rate at which sequences are produced.

Next we review models of serial order and of timing in accuracy (the speed–accuracy trade-off), and compare these models to the range model of planning. We then review the original range model (as in Palmer & Pfordresher, 2003) and introduce the extended model that accounts for overall accuracy. Next we describe an experiment on pianists' performances of melodies at different production rates and compare their accuracy and serial ordering errors to the model's predictions.

## 2. Constraints on serial order and accuracy

The retrieval of events in the correct serial order poses a particularly difficult challenge for sequence production tasks, such as performing music or speaking, as pointed out some time ago by Lashley (1951). First, actions must be generated rapidly in a series during sequence production. Under such circumstances, preparation cannot be limited to the current event. Sec-

ond, the meaning of an individual sequence event is often ambiguous when one fails to consider the surrounding context. Music is perhaps the most extreme example of this observation, in that individual events (notes or chords), which often repeat in multiple contexts, have meaning that is defined relative to their context (L. B. Meyer, 1956). Thus, it is important that performers associate individual events with the surrounding context during retrieval. In the range model, this constraint is reflected in the activation strengths of events surrounding the current event, which are used to predict patterns of serial ordering errors.

Serial ordering errors occur when people produce events intended for elsewhere in a sequence (e.g., slips of the tongue in speech [Bock, 1995; Dell, 1986; Garrett, 1980], slips of the finger in music [Palmer & van de Sande, 1995], or ordering errors in serial recall [Healy, 1974]). Serial ordering errors often resemble the intended event and are thought to occur because multiple events are similarly accessible at the same time. For instance, the word *bad* is likely to be replaced by *sad* rather than *was* when producing a statement like “the bad man was sad,” due to phonological or syntactic similarity. Such patterns suggest that serial ordering errors result from similarity-based interference among planned events. Serial ordering errors are referred to as *movement errors* when an event produced in error matches an intended event at another position.

The accessibility of surrounding context during response preparation can be measured by analyses of movement errors as a function of *distance*, the separation, in number of events, between the current position and the position for which the error was intended. For instance, if a performer begins the intended musical sequence G4–A4–B4 with the note B4, the error is said to have a distance of 2. The relative frequency with which movement errors originate from various distances is displayed in terms of a *movement gradient* (Brown, Preece, & Hulme, 2000; Vousden, Brown, & Harley, 2000), which the original range model was designed to predict. Movement gradients plot the conditional probability of an error from a certain distance, given that an error has occurred.

Error distances reflect a performer’s *range of planning*, the degree to which distant versus proximal events are mentally accessible during retrieval (Palmer & van de Sande, 1995). Performances characterized by errors from far distances represent a broad range of planning and greater access to the surrounding context. Although research in the past has explored the role of error distance in the retrieval of serial order (Drake & Palmer, 2000; Palmer & Drake, 1997; Palmer & Pfordresher, 2003; Palmer & van de Sande, 1993, 1995; for similar results in speech see García-Albea, del Viso, & Igoa, 1989), the range model is unique in considering the way in which production rate may influence accessibility of the surrounding context.

Retrieval of serial order is clearly symbiotic with accuracy, in that correct retrieval denotes perfect accuracy, although distributions of serial ordering errors (as in movement gradients) are statistically independent of overall accuracy (provided that errors occur). The extended model was developed to account for both accuracy and patterns of serial ordering errors. Because time is central to the range model framework, we incorporated temporal constraints to predict overall accuracy in the context of the speed–accuracy trade-off.

The speed–accuracy trade-off is one of the benchmark findings of experimental psychology: The possible accuracy with which a task can be performed diminishes as individuals perform the task more rapidly, and vice versa (first documented by Woodworth, 1899). Changes in accuracy with speed, or vice versa, form a *speed–accuracy function* that can describe the per-

formance of an individual or a sample. Speed–accuracy trade-offs are typically analyzed with respect to the performance of individual participants under conditions in which requirements for speed or accuracy are manipulated experimentally. Certain tasks focus on how fast participants can complete a task while maintaining a prescribed level of accuracy. For instance, in target acquisition tasks, participants are required to generate movements (i.e., of a stylus) to a target location as quickly as possible and the dependent measure is the time required to maintain this level of accuracy (Fitts, 1954). Other paradigms, like that used in this study, focus on how accuracy is influenced by time constraints. Although speed–accuracy trade-offs have been examined in many single-response tasks such as target acquisition (e.g., Fitts, 1954), choice reaction time (e.g., Ratcliff, 1978), and serial recall (e.g., McElree, 2001), fewer studies have addressed trade-offs in sequential tasks such as speech or music.

Interestingly, past research on music performance has not revealed the same speed–accuracy trade-off. Most studies have documented nonsignificant relations between error rates and mean production rate (Palmer & Drake, 1997; Palmer & van de Sande, 1993, 1995; Sloboda, Clarke, Parncutt, & Raekallio, 1998). One could infer from such null results that the temporal flexibility acquired by expert musicians frees them from the speed–accuracy trade-off. Drake and Palmer (2000) speculated that the absence of speed–accuracy trade-offs may reflect a dominance of relative timing over global production rate in music production. However, the range of prescribed rate conditions may have not been broad enough to elicit speed–accuracy trade-offs in those studies. Palmer and Pfordresher (2003), by contrast, found higher error rates in a fast tempo condition than in a more moderate tempo condition, but more than two tempo conditions are necessary to clarify the shape of the speed–accuracy function. As it stands, we know of no existing research that has systematically varied tempo in music performance across a wide enough range of tempi to test whether speed–accuracy trade-offs occur in music performance as well as the shape of the speed–accuracy function.

### 3. The range model of planning

#### 3.1. *The original range model*

We focus here on aspects of the range model (Palmer & Pfordresher, 2003) that are central to the extended model (other aspects of the original model are presented in the Appendix). Readers familiar with that model may continue to the section 3.2, in which we describe new predictions for overall accuracy and the speed–accuracy trade-off. The original range model predicts patterns of movement errors and theoretically relates to the process of response preparation. The extended model predicts overall accuracy and the speed–accuracy trade-off, and theoretically relates to response selection.

In the range model, event activations form a graded distribution across serial positions such that activations surrounding the current position diminish with decreasing serial proximity and similarity to the current event. Two examples are shown in Fig. 1. This framework differs from one in which accessibility is all-or-none, such as a model that would propose a moving time window around the current event or a chunking mechanism. An important implication of the range model for this research is that the activation of distant events is greater at slower tempi

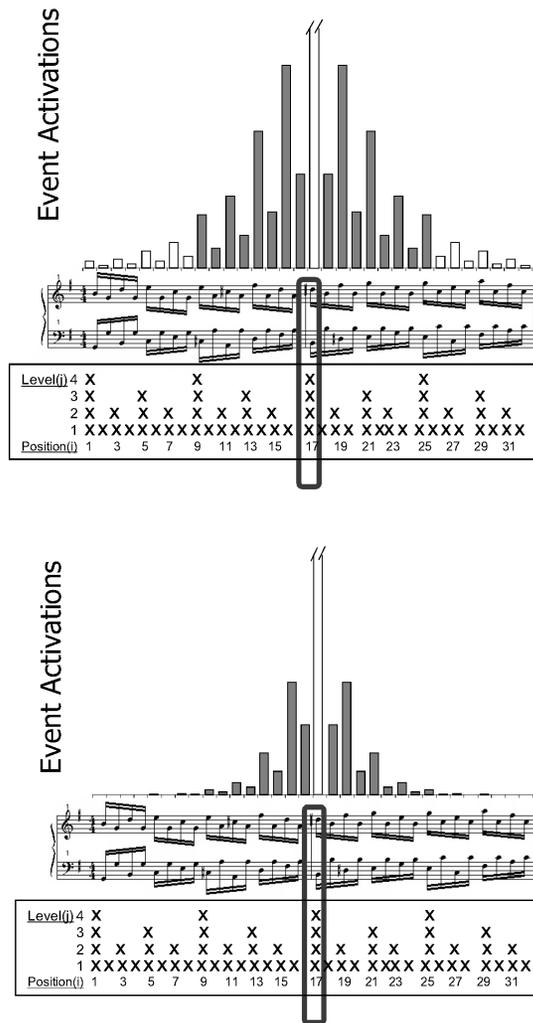


Fig. 1. Gradient of activations predicted by the range model of planning (Palmer & Pfordresher, 2003), for a slow tempo (top,  $t = .32$ ), and a fast tempo (bottom,  $t = .15$ ), shown above music notation. Rectangles in notation highlight the current intended event. Both gradients use  $a = .92$  and  $w_2 = .5$ .

(Fig. 1, top), than at faster tempi (Fig. 1, bottom). Bars showing activations for current events are open at the top because the extended model (described later) proposes that the activation of surrounding events alters the activation of the current event, whereas the original model predictions do not incorporate the current event's activation (which was provisionally fixed at one by Palmer & Pfordresher, 2003).

The original range model predictions emerge as the product of two components: serial proximity and metrical similarity. The serial component of the range model is most central to the predicted speed-accuracy trade-off because only the serial component explicitly involves timing (Palmer & Pfordresher, 2000). The serial proximity component predicts a decrease in event

activations ( $S_x$ ) as a function of the absolute distance of events from the current event ( $|x|$ , where  $x < 0$  for past events and  $x > 0$  for future events), with the most active events being those closest to the current event:

$$S_x = a^{(|x|/t)} \quad (1)$$

The  $t$  parameter, a fixed parameter, measures event durations (interonset intervals [IOIs], in seconds) as specified by the performance tempo. Because of possible differences in the way in which duration is processed as a function of time scale (Drake & Botte, 1993; Friberg & Sundberg, 1995; Hibi, 1983; Peters, 1989), we limit the  $t$  parameter to those durations that are typically associated with rhythmic sequences,  $0.1 < t \leq 2.0$  sec. The  $a$  parameter, a free parameter that is allowed to range from  $0.8 < a \leq 1.0$  (the lower limit was imposed because activations drop off precipitously for lower values of  $a$ ), is fit to proportions of serial ordering errors as a function of distance and may reflect the working memory capacity of the performer (Palmer & Schendel, 2002). The value of  $t$  does not vary for events with different durations (e.g., in the context of rhythms).

The serial proximity component always predicts decreases in event activation with distance from the current event (unless  $a = 1$ , for which  $S_x = 1$  for all  $x$ ); lower values for either  $t$  or  $a$  enhance the rate of decrease. Thus, events surrounding the current event are less active overall when an individual performs at a faster, as opposed to slower, tempo (lower  $t$ ) or for performers with smaller, as opposed to larger, short-term memory capacity (lower  $a$ ). Fig. 1 demonstrates the influence of tempo on the distribution of activations. Lower activations of surrounding events result in a smaller range of planning: Fewer sequence events are accessible and movement errors originate from nearby events. Activations of distant events are presumed to increase as tempo slows because more time is available for performers to retrieve information about the surrounding context during response preparation. The distribution of activations across distance is the same for each sequence position, and is symmetric across positive and negative distances. One necessary exception is that no activation is predicted for nonexistent sequence events (e.g., when the current position is 1, no activation of past events is predicted).

Event activations at each distance and sequence position are estimated by the product of the serial component and a metrical component ( $S_x * M_x$ ) across all distances and at every sequence position. Meter refers to the pattern of regularly alternating strong and weak accents that underlie sequences like speech and music (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983; Liberman & Prince, 1977; Palmer & Krumhansl, 1990). The metrical component ( $M$ ), which measures similarity of surrounding events to the current event based on metrical accent strengths, yields a second free parameter for fits of the original range model ( $w_2$ ) that represents the relative importance of different metrical levels. This parameter contributes substantially to fits of the original model to movement gradients, but has negligible influence on extended model fits to speed and accuracy data. Thus the metrical component and  $w_2$  are described in the Appendix. With respect to the predicted activations of the original model in Fig. 1, the metrical component accounts for the periodic alternations in activation strengths, whereas the serial component accounts for the degree to which activations spread to distant events.

The primary purpose of the original range model was to account for movement gradients, the conditional probability of an error at distance  $x$  given that an error occurs,  $p(\text{err}_x | \text{err})$ . Predicted movement gradients in the range model are generated by averaging activations across sequence positions, as well as positive and negative error distances, and then dividing each mean activation strength by the sum of mean activations across absolute distances. Model fits were limited to errors with a source within a distance of 8 events (the dark bars in Fig. 1, which includes most errors), because pitches typically repeat every 8 events in the stimuli used. Errors with a distance of greater than 8 events could reflect the random probability of pitch repetition (more so than events within a window of 8 events). Palmer and Pfordresher (2003) reported fits of predicted movement gradients to movement error data from adult and child pianists, data from memorized performances and performances from notation, performances from binary and ternary meters, and performances with differing metrical interpretations. In this article, we extend this perspective to model the probability that an incorrect pitch event will be selected in the first place,  $p(\text{err})$ .

### 3.2. The extended range model

The extended range model uses the same basic framework as the original model to predict the probability of selecting an incorrect pitch event,  $p(\text{err})$ , by introducing a new assumption regarding the activation strength of the current event. In the original range model, the activation of the current event is not used to generate predictions. Palmer and Pfordresher (2003) kept the current event's activation fixed at 1 by default. This assumption leads to counter-intuitive predictions (described later) if response selection is based on event activations, as is typically the case (e.g., Dell, 1986; MacKay, 1987). In addition, we introduce two new parameters to the extended model to account for factors that may constrain the accuracy of response selection, but may not constrain the preparation of multiple responses prior to selection. One new parameter,  $t'$ , was designed to account for individual differences in motor dexterity whereas the other,  $B$ , was designed to account for individual differences in domain-specific skill.

Errors in the extended model occur when response selection results in the production of some event other than the current event. Because activation strengths reflect accessibility, and thus probability of retrieval, error probability becomes the ratio of summed activations for all noncurrent events (i.e., all potential errors) relative to the summed activations for all events. This relation can be expressed as follows:

$$p(\text{err}) = \frac{\sum_{\sim \text{current}}}{(\sum_{\sim \text{current}}) + \text{current}} = \frac{\sum_{x \neq 0} S_x * M_x}{\sum_x S_x * M_x} \quad (2)$$

where  $\sim \text{current}$  refers to activations of events other than the current event, and  $\text{current}$  refers to the activation of the current event. The numerator is thus the sum of all event activations surrounding the current event, whereas the denominator includes the current event's activation in addition to the sum of surrounding event activations. Mean  $p(\text{err})$  across all positions generates

the estimated error rate for a performance. If the activation of the current event always equals one, as was provisionally suggested by Palmer and Pfordresher (2003), then  $p(\text{err})$  would increase as production rate slowed, because the activation of surrounding events would increase relative to the current event. The extended model alters this default assumption.

We introduce the assumption that the activation of the current event is altered by the activations of surrounding events. Specifically, we propose that the activation of the current event grows in proportion to the squared sum of the surrounding events' activations:

$$\text{current} = S_0 * M_0 = \left( \sum_{x \neq 0} S_x * M_x \right)^2 \quad (3a)$$

As a result, the extended model makes the following basic assumption:

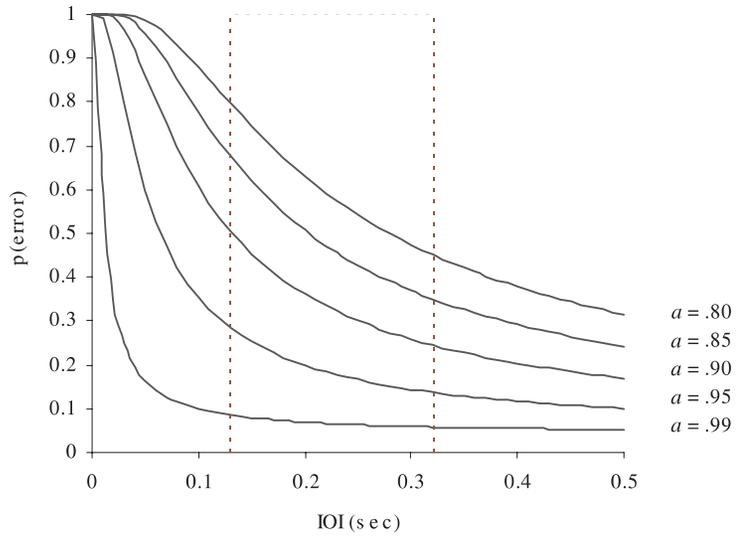
$$p(\text{err}) = \frac{\sum \sim \text{current}}{(\sum \sim \text{current}) + \text{current}} = \frac{\sum_{x \neq 0} S_x * M_x}{\left( \sum_{x \neq 0} S_x * M_x \right) + \left( \sum_{x \neq 0} S_x * M_x \right)^2} \quad (3b)$$

Note that this new assumption regarding the current event's activation does not alter the predictions of the original model, which excludes the current event entirely.

Fig. 2a plots predicted error rates from the extended model for different values of  $a$ . Error probabilities were generated for each position in a sequence (comprising 32 events, as in the stimuli used for the experiment) and then averaged. We include values of  $t$  lower than .10, which are below the limit that is applicable to the range model, to illustrate the fact that all functions share maxima at  $p(\text{err}) = 1.00$ . The  $a$  parameter influences the predicted shape of speed-accuracy functions, particularly in the range of tempi used for this experiment (see area within dashed lines), which reflect tempi used often in music performance. For low values of  $a$  (low working memory), predicted error rates decrease at a consistent rate as IOI increases, and error rates are always high relative to functions resulting from other values of  $a$ . For higher values of  $a$ , speed-accuracy functions decrease with IOI rapidly at first and then approach an asymptote. The distinction between the initial rapid descent in errors and the second, flatter phase is enhanced as  $a$  increases. Note also that within the range of tempi used for this experiment (see region within dashed lines), particularly high values of  $a$  can lead to a negligible speed-accuracy trade-off with consistently high accuracy.

Fig. 2b illustrates how activations of current and surrounding events lead to predicted error rates for a single representative value of  $a$ . Two important characteristics of the model can be seen here. The first characteristic concerns the two phases of the speed-accuracy function; the distinction between the initial rapid descent in error rates and the subsequent plateau. The portion of the function in which errors decrease rapidly with IOI occurs when the activations of both current and summed surrounding events increase with IOI. Recall that increases in activations across all events (current and surrounding) increase the denominator of Equation 3b. The plateau then occurs when activations of both current and surrounding events reach their respective plateaus. The second characteristic concerns the way in which activations determine the value of  $p(\text{err})$  at any point along the speed-accuracy function. As can be seen,  $p(\text{err}) > .5$  when the activation of the current event is less than the sum of surrounding events. Asymptotic  $p(\text{err})$

A



B

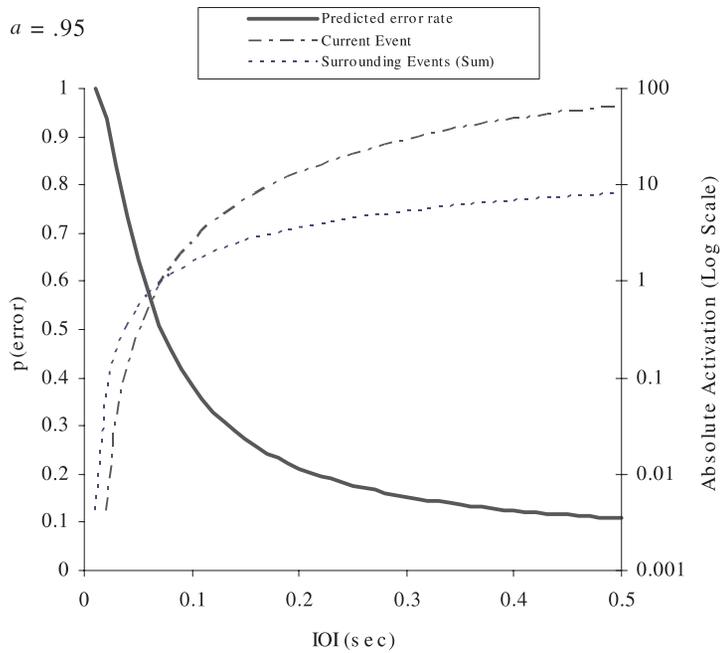


Fig. 2. (a) Predicted speed-accuracy functions from Equation 2, for six different values of  $a$ . (b) Predicted speed accuracy function (left ordinate) plotted with component activations for the current event and summed surrounding events (right ordinate) for  $a = .90$ . Dashed lines highlight model predictions for the prescribed tempi used in the reported experiment.

is related to the difference in activations between the current and surrounding events when they reach plateau.

We incorporated two additional parameters in the extended range model to account for factors that may contribute to response selection but not response preparation. The first parameter,  $t'$ , was designed to address differences in motor skills. The introduction of  $t'$  was motivated by research suggesting that individual performers differ in the maximum possible speed of performance (MacKenzie & Van Eerd, 1990; R. K. Meyer & Palmer, 2003; Palmer & Meyer, 2000), and that noise in the motor system may contribute to speed–accuracy trade-offs in general (Schmidt, Zelaznik, Hawkins, Frank, & Quinn, 1979). We reasoned that differences in motor abilities may be borne out in different error rates at the fastest tempi. Moreover, examination of obtained speed–accuracy functions from performers suggested that for some performers there is a critical IOI at which decreases in IOI lead to steep increases in error rates. Thus, we hypothesized that there may be an IOI that functions like a “motor threshold” for performers, such that IOIs above the threshold (slower tempi) can be played relatively accurately but IOIs below the threshold (faster tempi) lead to large error rates.

The  $t'$  parameter influences the IOI associated with a motor timing threshold by shifting speed–accuracy functions along the  $x$  axis. It is entered into the serial component of the range model (cf. Equation 1):

$$S'_x = a \left\lfloor \frac{|x|}{t-t'} \right\rfloor \quad (4)$$

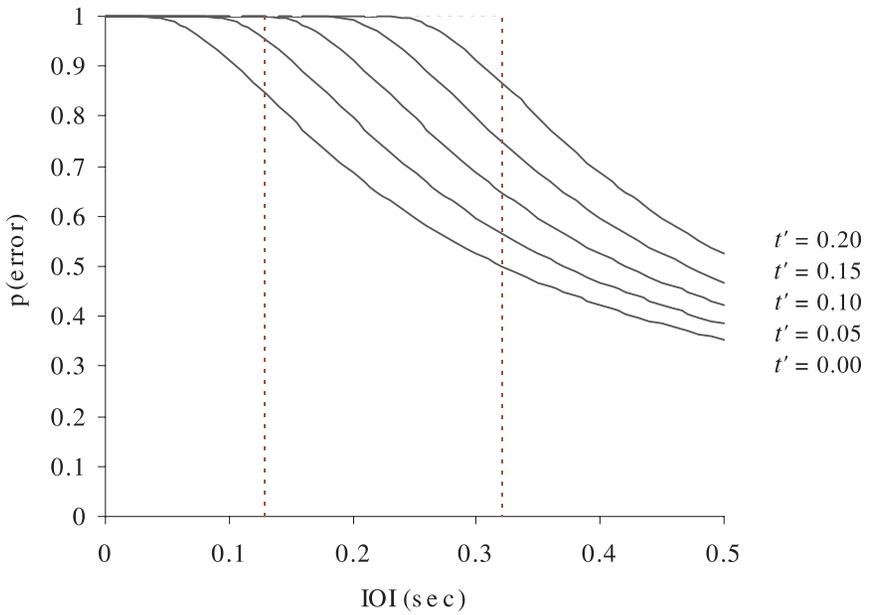
for which  $0 \leq t' < t$ . The  $t'$  parameter decreases event durations in the model from their actual produced values ( $t$ ). We varied  $t'$  in this way based on the logic that the difficulty of producing a sequence of IOIs reflects the limitations of an individual’s motor system. A performer with good motor dexterity should yield a  $t'$  of zero, with higher values of  $t'$  reflecting less well-developed motor systems that reach their upper tempo limit at a somewhat slower rate. In other words, a performer with a high  $t'$  (e.g.,  $t' = .10$ ) may produce IOIs of .30 sec with the same ease as would a more adept motor system performing .20-sec IOIs ( $t = .30$ ,  $t' = .10$ ,  $t - t' = .20$ ).

The effect of varying  $t'$ , shown in Fig. 3a for  $a = .80$ ,<sup>3</sup> implements the assumption given earlier:  $t'$  shifts the speed–accuracy function along the  $x$  axis, and as a result a particular IOI is predicted to elicit higher error rates for a participant with a high  $t'$  than for a participant with a low  $t'$ . This can be seen by comparing  $p(\text{err})$  across all functions for a particular IOI. Because Fig. 3a shows functions associated with a low value of  $a$ , speed–accuracy functions do not reach asymptote within the range of IOIs shown.

Variations of  $t'$  have no influence on the overall shape of the speed–accuracy function, as can be seen by comparing variations of  $t'$  in Fig. 3a with those in Fig. 3b ( $a = .99$ ). In addition, the high value of  $a$  in Fig. 3b demonstrates the effect of  $t'$  on asymptotic behavior is illustrated. As can be seen,  $t'$  influences the IOI at which decreased IOIs cause steep increases in error rates and increased IOIs lead to stable and highly accurate production.

The second new parameter was designed to address individual differences in domain-specific skill (cf. Ericsson, Krampe, & Tesch-Römer, 1993; Krampe & Ericsson, 1996). Performers differ in the rapidity with which they can learn to perform musical sequences to some criterion (Drake & Palmer, 2000). Such differences in skill may influence the way in

A



B

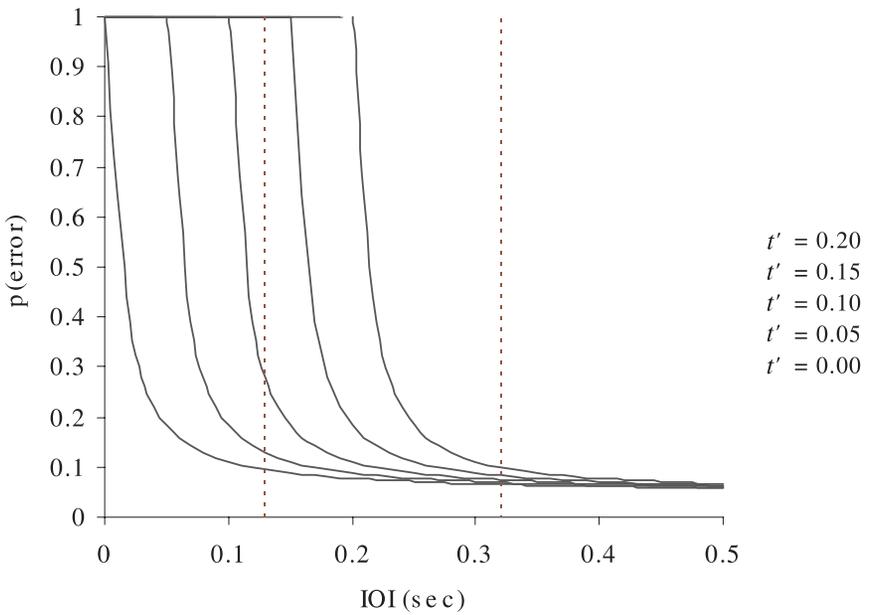


Fig. 3. Predicted speed-accuracy functions that incorporate  $S'$  from Equation 4, for five different values of  $t'$  when (a)  $a = .80$  and (b)  $a = 0.99$ . Dashed lines highlight model predictions for the prescribed tempi used in the reported experiment.

which surrounding events influence the activation level of the current event. The second new parameter,  $B$ , models the way in which skill facilitates retrieval, by weighing the current event's activation. Combining this parameter with  $t'$  yields a modified equation for the current event's activation strength (cf. Equation 3a):

$$current = B * \left( \sum_{x \neq 0} S'_x * M_x \right)^2 \quad (5a)$$

Thus,

$$p(err) = \frac{\sum_{x \neq 0} S'_x * M_x}{\left( \sum_{x \neq 0} S'_x * M_x \right) + [B * \left( \sum_{x \neq 0} S'_x * M_x \right)^2]} \quad (5b)$$

for which  $B \geq 1$ , with no upper limit (as  $B \rightarrow \infty$ ,  $p(err) \rightarrow 0$ ).

Fig. 4a displays the influence of  $B$  on predicted speed–accuracy functions when  $a = .80$  and  $t' = 0$  for a range of production rates.  $B$  has some influence on the shape of the function (as does  $a$ ), primarily along the y axis, but has no influence on predictions for patterns of movement errors. Furthermore, different values of  $a$  have a significant impact on predicted speed–accuracy functions regardless of the value of  $B$ , which can be seen by comparing Figs. 4a ( $a = .80$ ) and 4b ( $a = .99$ ).

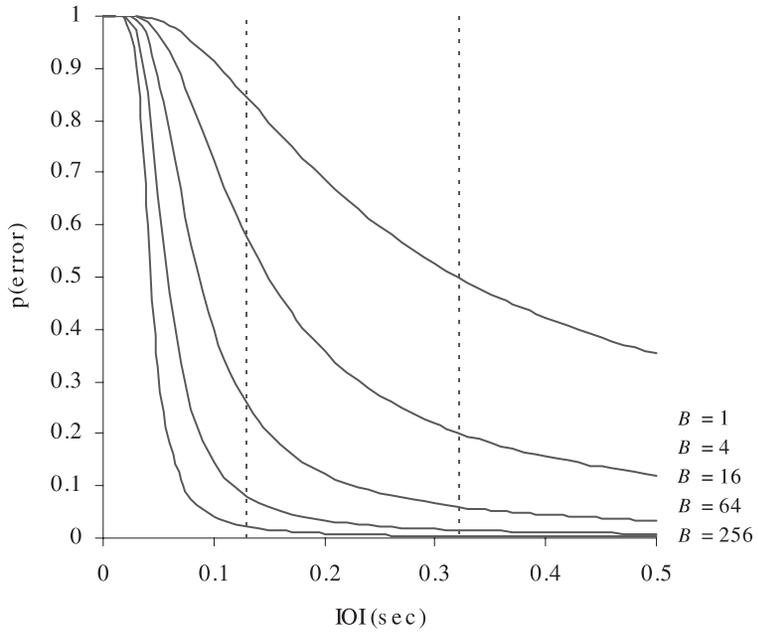
We conducted an experiment that was designed to elicit errors in music performance across a broad range of performance tempi. Pianists performed unfamiliar sequences that were practiced thoroughly at the beginning of the session until they could be performed without errors at a slow tempo; thus the task did not involve sight-reading. Sequences were always performed in view of music notation at a tempo set by a metronome. The sequences were musically complex finger exercises (also used in Palmer & Pfordresher, 2003), designed to be structurally similar to tongue-twisters used to elicit errors in speech (cf. Dell, 1986; Dell, Burger, & Svec, 1997; Rosenbaum, Weber, Hazelett, & Hindorff, 1986). We fit the original range model to distributions of movement error frequencies across error distance (movement gradients), and fit the extended range model to speed–accuracy data.

## 4. Method

### 4.1. Participants

Twelve adult pianists from the Ohio State University community ( $M$  age = 24.1 years) participated in exchange for course credit in introductory psychology or payment. Participants had between 7 and 32 years of private piano instruction ( $M = 14.1$  years) and between 10 and 34 years of piano playing experience ( $M = 16.3$ ). All participants reported playing piano regularly and reported no hearing problems. Eleven of the 12 pianists were right-handed.

A



B

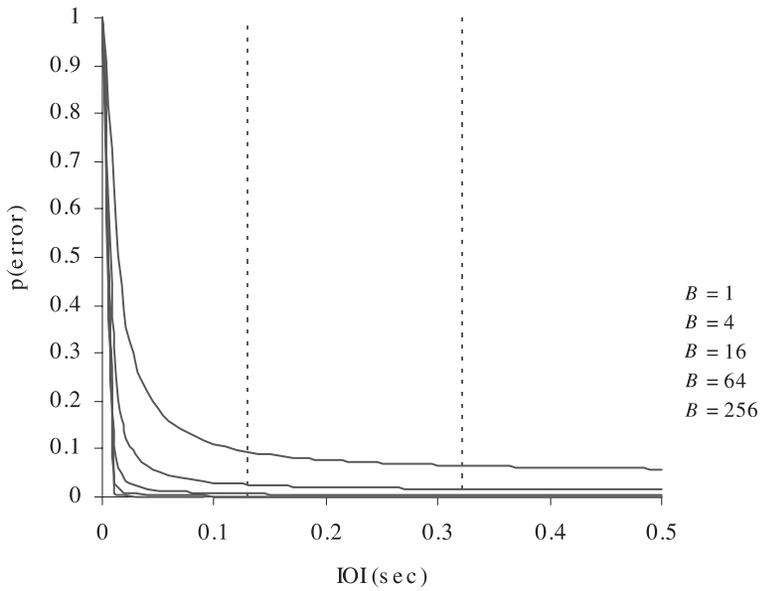


Fig. 4. Predicted speed–accuracy functions from Equation 5b, for six different values of  $B$  when (a)  $a = .80$  or (b)  $a = .99$ . For both plots,  $t' = 0$ . Dashed lines highlight model predictions for the prescribed tempi used in the reported experiment.

#### 4.2. Apparatus

Pianists performed on a computer-monitored Yamaha Disklavier MX100 acoustic upright piano. Optical sensors detected key press velocities without affecting the touch or sound of the acoustic piano. The pitch, timing, and hammer velocity values (correlated with intensity) for each note event were recorded on an IBM PC-compatible computer in MIDI format. The timing resolution of MIDI data acquisition was 2 msec; this precision was equivalent to 0.8% of the measured variability in produced timing (indexed by *SD* of IOIs in performances). Pitch errors were detected by comparing performances with the contents of music notation (Large, 1993; Palmer & van de Sande, 1993, 1995).

#### 4.3. Materials

The stimuli were four sequences taken from Palmer and Pfordresher (2003), adapted from finger exercises by Czerny (Opus 299). Each sequence contained 32 sixteenth-note events (event = a single note or a set of notes intended to sound simultaneously), notated with the same time signature (4/4). Similar to phonemic repetitions in tongue-twisters, sequences initially established repeating pitch and duration patterns that were later violated. Two of the stimuli had sixteenth-notes in both left and right hands and two had sixteenth-notes in the right hand and quarter-notes in the left hand. There were no successive repeating pitches and pitches repeated every 8 events on average.

#### 4.4. Design and procedure

Eight tempi (.12, .15, .18, .21, .24, .27, .30, and .33 sec per sixteenth-note IOI) were crossed with the four musical sequences for a total of 32 trials in a within-subjects design. Each trial comprised two performances of a tempo/sequence combination in succession. The session comprised 8 blocks of 4 trials; each trial within a block comprised a unique tempo–sequence combination. Within these constraints, two random orders of trials were generated according to a Latin square design. Each participant was randomly assigned to one of the two orders.

At the beginning of the session, each participant completed a musical background questionnaire. Then the participant was presented with a notated musical sequence and asked to perform each melody at a slow tempo of sixteenth-notes = .536-sec IOIs, indicated on a metronome, until an error-free performance was obtained. These slow performances were designed to ensure that errors at faster tempi did not result from incorrect learning of the sequences. The 8 experimental blocks followed the initial slow performance block.

In each trial, participants performed one of the musical sequences twice, pausing between repetitions, with the notation in view. The metronome established tempo at the quarter-note level (metronome IOI = 4× produced IOIs), and participants synchronized the first event in each group of 4 with the metronome. The metronome was set at this slower rate because setting the metronome at the sixteenth-note rate (the fastest of which would be 462 beats per minute) would yield a distractingly fast pulse sequence and is not common musical practice. Participants were instructed not to stop if they made errors. Thus, instructions emphasized speed, in terms of tempo, rather than accuracy (cf. Pew, 1969). At the conclusion of the experimental

session, the participants performed each of the four musical sequences again at the slow rate twice (536-msec IOIs) to ensure there were no learning errors that may have developed across the session.

#### 4.5. Error coding

Pitch error rates (number of errors relative to number of error opportunities) were averaged across the two repetitions of the same sequence–tempo pair for all trials. Pitch error rates were first calculated separately for single notes and chords (which occurred with different frequencies in the musical pieces) and then summed for the overall error rate per performance. Movement gradients for error distances were then computed from serial ordering (i.e., movement) errors. Serial ordering errors are defined here as errors that match an event intended for elsewhere in the sequence (i.e., the insertion of event  $i + x$  at position  $i$ ). The assumption that notated events were intended is based on pianists' error-free performances at the initial slow tempo. The nearest sequence event with the same pitch as the error is termed the error's source, and the absolute distance between the error and its source was computed. Corrections (interrupted errors, in which an event was performed incorrectly and then corrected), were excluded due to ambiguities in their coding (as in Dell, 1986; Garrett, 1980).

## 5. Results

Results are organized into four main sections. First we describe the empirical results of this study. Then we describe model fits for the original and extended range models, respectively.

### 5.1. Obtained error data

#### 5.1.1. Speed–accuracy trade-offs

There were a total of 4,586 pitch errors in all; the mean error rate per trial was 0.112. Fig. 5 shows mean error rates as a function of the mean IOIs (inverse of speed) produced by participants in each tempo condition. Error rates varied reliably as a function of tempo condition. A one-way within-participants analysis of variance (ANOVA) on mean error rates revealed a main effect of tempo condition,  $F(7, 77) = 24.08$ ,  $MSE = .011$ ,  $p < .01$ . No differences as a function of stimulus melody or practice within the session emerged.

#### 5.1.2. Movement gradients.

Analyses of movement errors were based on a subset of the error data used to analyze speed–accuracy trade-offs. Deletions (23% of the data) are ambiguous with respect to the error source and were excluded. Of the remaining errors, 81% had an identifiable source elsewhere in the sequence and were serial-ordering errors. Of those errors, 71% had an identifiable source within a distance of 8 events surrounding the current event (one metrical cycle, equivalent to the chance estimate of how often pitches repeated in the stimuli), and were incorporated in analyses of movement errors. Of movement errors, 44% were anticipatory (i.e., the source followed the error; 1,057 errors) and 56% were perseveratory (1,342 errors).

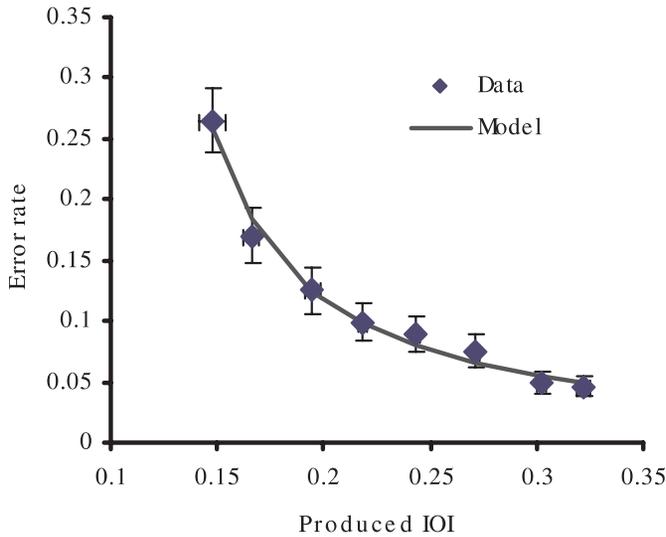


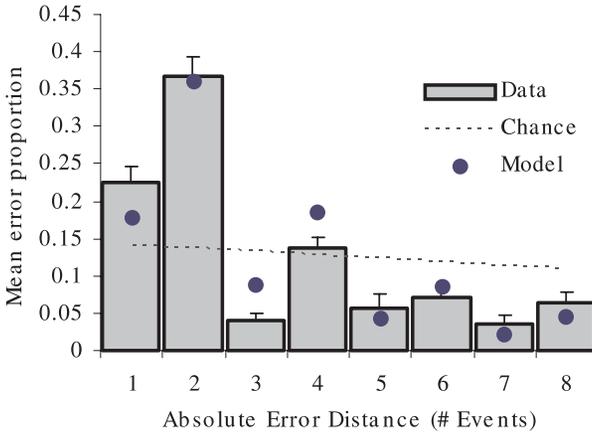
Fig. 5. Obtained and predicted speed–accuracy functions for mean data. For this function,  $a$  was fixed at .92 and  $w_2$  was fixed at .54. Other parameters were fit,  $B = 7.85$ ,  $t' = 0.08$ .

Movement gradients for each participant were calculated using the proportion of movement errors with sources from distances of  $\pm 1$  to 8 (error frequencies at each distance divided by total number of movement errors in this range). The mean movement gradient across participants and tempo conditions is shown in Fig. 6a. An 8 (distance)  $\times$  8 (tempo) repeated measures ANOVA on error proportions yielded a significant main effect of distance,  $F(7, 77) = 30.349$ ,  $MSE = .042$ ,  $p < .01$ , but no Distance  $\times$  Tempo interaction ( $p > .10$ ). As predicted by the original range model (Palmer & Pfordresher, 2003), errors associated with proximal events (shorter distances) were more prevalent than errors associated with more distant events, and errors associated with metrically similar events (distances that are multiples of 2, see Appendix) were more prevalent than errors at other distances. Fits of the original range model to the mean data, shown as dots in Fig. 6, are discussed in the modeling section later.

We also examined patterns of movement errors at the extreme tempo conditions. An ANOVA on the fastest and slowest tempo conditions yielded a significant Tempo  $\times$  Distance interaction,  $F(7, 77) = 2.25$ ,  $MSE = 0.056$ ,  $p < .05$ , in addition to the significant effect of distance,  $F(7, 77) = 25.54$ ,  $MSE = 0.015$ ,  $p < .01$ . Fig. 6b shows mean data (across participants) from these tempo conditions. As can be seen, the frequency of errors from farther distances was markedly increased in the slowest compared with the fastest condition.

Finally, we tested the range model's prediction that the overall scope of planning increases with tempo. Mean absolute range was calculated for each tempo condition from the mean of all absolute error distances within 8 events. Fig. 7 displays mean range averaged across participants as a function of produced IOI (also averaged across participants) for each tempo condition. The influence of tempo on range approached significance,  $F(7, 77) = 1.90$ ,  $MSE = 1.081$ ,  $p = .08$ . As with the analysis of error frequency by distance, we also examined the two extreme tempo conditions. The difference between these two conditions was significant,  $F(1, 11) = 4.67$ ,  $MSE = 1.96$ ,  $p = .05$ . The correlation between produced IOI and mean absolute range was

A



B

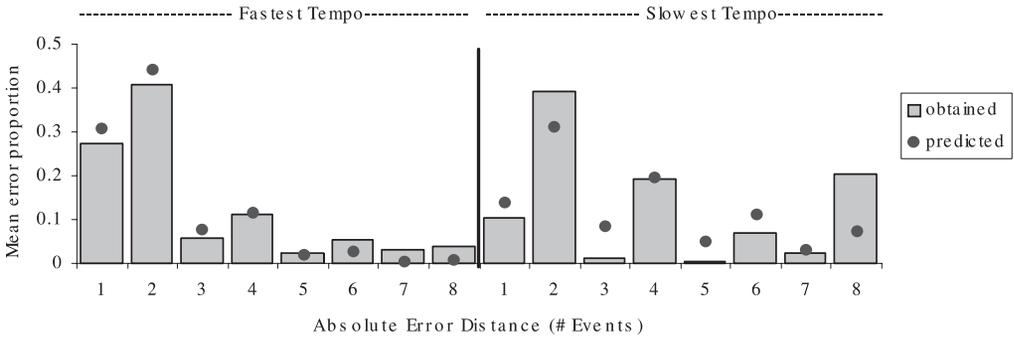


Fig. 6. (a) Obtained movement gradient, averaged across participants and tempo conditions, with predictions from the original range model and chance estimates. Mean  $a$  across participants was used for this fit,  $a = .92$ ,  $w_2 = .54$ . (b) Corresponding fits for the fastest and slowest tempo conditions using same parameters.

significant when the data were broken down by both tempo and individual ( $df = 94$ ),  $r = .19$ ,  $p < .05$  ( $r = .44$  for the two extreme tempi).

5.2. Model fits to movement gradients: Original range model

The original range model was fit to movement gradients separately for each participant and tempo condition. Fits were carried out in two steps, following the same procedure used by Palmer and Pfordresher (2003). First, the range model was fit to each participant’s movement gradient using one free parameter,  $a$ , which was allowed to vary between .8 and .99 (see Equation 1). The value of  $a$  was then fixed to the best fitting value from the first iteration and the

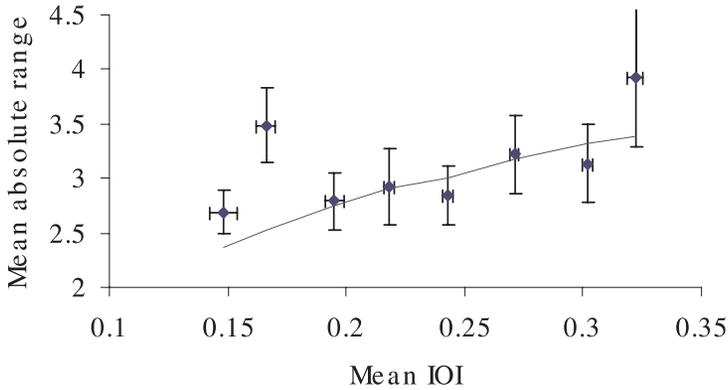


Fig. 7. Obtained mean absolute range by mean interonset intervals (dots), and range model prediction (line). Error bars represent  $\pm 1$  SE.

model was fit again, allowing metrical weight on level 2 of the metrical grid ( $w_2$ ) to vary between .01 and .99 (see Appendix).

Fits of the original range model were generated in this way for each participant and tempo condition. Goodness of fit was established through the metric variance accounted for (VAF). VAF measures goodness of fit based on the ratio of squared error between model and data to the squared error within the data.<sup>4</sup> On average, predictions accounted for 73% of the variance in movement gradients for each participant and tempo condition. In fact, 78 of the 89 fits (88%) for each participant and tempo condition were significant (some cells were missing due to an absence of errors). Table 1 shows VAF for model fits as well as model parameters for each individual data, averaged across tempi. The model provided a good fit to the data based on VAF for all but 2 participants. The participant with lowest VAF made very few errors (< 1%) resulting in an anomalous movement gradient. Fig. 6a shows predicted values averaged across participants and tempi, which provided a good fit of mean obtained data (VAF = 93%) and for the extreme tempo conditions (Fig. 6b; *M* VAF across the extreme conditions = 84%).

### 5.3. Model fits to speed–accuracy trade-offs: Extended range model

We then fit the extended range model, described earlier, to obtained speed–accuracy data. All fits used values of  $a$  and  $w_2$  from previous fits to movement gradients. The two remaining free parameters ( $B$ ,  $t'$ ) were fit simultaneously through the least squares optimization procedure in MATLAB (The MathWorks, Natick, MA). Performed, rather than prescribed, mean IOIs were used as input to all fits. Parameters were not allowed to vary with tempo condition; values from fits of the original range model reflected average parameter values across tempi.

Fig. 5 shows the range model's fit to the mean data across participants. This fit accounted for almost all the variance in the error data (VAF = 99%,  $p < .01$ ). Best fitting parameter values and VAFs for individual fits are shown in Table 1. We do not include the  $w_2$  parameter (from the metrical component of the original model), which yields negligible effects on predicted speed–accuracy functions. Model fits accounted for 95% of the variance, on average, in individual participants' error data and all individual fits were statistically significant.

Table 1  
Performance measures, model parameters, and VAF for individual performers

	Performance Measures				Model Parameters			VAF	
	Err-Rate	Range	Practice	T-Err	<i>a</i>	<i>B</i>	<i>t'</i>	Original (%)	Extended (%)
Performer									
1	0.11	3.60	2.50	0.000	0.935	9.79	0.087	88	95
2	0.07	2.86	2.31	0.004	0.917	12.45	0.064	76	74
3	0.02	1.93	2.50	-0.001	0.898	178.33	0.082	79	99
4	0.26	3.11	4.31	0.050	0.917	2.48	0.121	95	84
5	0.26	2.57	3.63	0.010	0.904	3.90	0.083	95	96
6	0.01	4.78	2.00	-0.001	0.945	245.30	0.092	<0	97
7	0.04	2.63	2.75	-0.001	0.916	45.38	0.073	81	96
8	0.03	2.65	2.75	0.000	0.893	97.48	0.080	73	>99
9	0.06	3.04	2.25	0.005	0.908	104.04	0.120	28	>99
10	0.24	3.44	5.38	0.013	0.935	1.00	0.000	95	91
11	0.14	2.86	6.00	0.016	0.912	4.96	0.073	92	60
12	0.11	4.16	3.63	0.001	0.954	4.28	0.068	82	96
Median	0.09	2.95	2.75	0.003	0.917	11.12	0.081	81	96

Note. VAF = variance accounted for; Err-rate = mean error rate across all trials for a participant; Range = mean absolute error distance; Practice = mean number of repetitions in block 1; t-err = mean signed difference between produced and prescribed IOI.

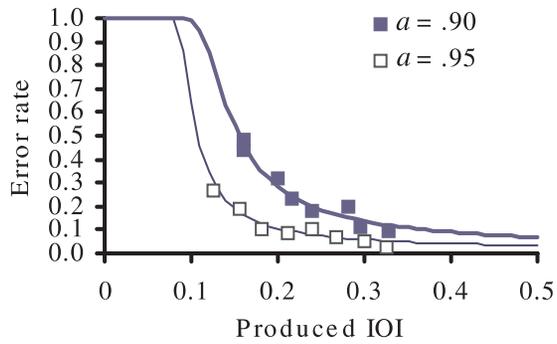
### 5.3.1. Individual differences

The ability of the extended range model to account for individual differences is demonstrated by examining pairs of participants who differ primarily along one of the three primary model parameters (*a*, *B*, *t'*). For each parameter, two individuals were selected who primarily differed in one parameter, while maintaining the highest possible similarity across the other two parameters. Fig. 8 shows data from the participants who were selected. These plots verify qualitatively that the different influences of parameters on speed-accuracy functions, shown in Figs. 2 through 4, can be observed in individual participants. Each participant's extended model fit was extrapolated to range from  $0 \leq \text{IOI} \leq .5$ , to allow comparisons with Figs. 2 through 4.

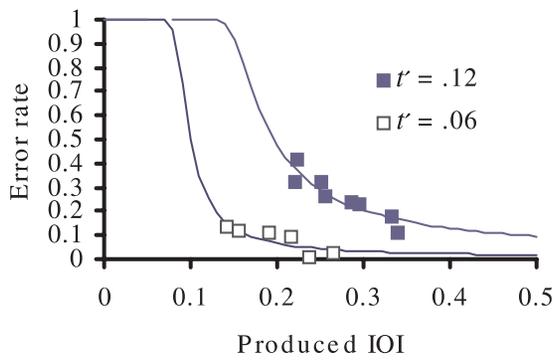
Fig. 8a demonstrates the influence of *a* on speed-accuracy functions using the data of Participants 5 (low *a*) and 12 (high *a*) from Table 1 (cf. Fig. 2a). The function for the participant with a high *a* (better performance) is more bowed than that of the participant with a low *a* (worse performance). As a result, the high-*a* participant reaches asymptote more quickly. However, both participants' errors converge at the slowest tempi (largest IOIs). Otherwise, both these performers exhibit relatively low values of *B* ( $M = 4.09$ , reflecting generally high error rates), as well as moderately high values of *t'* ( $M = .065$ ). Fig. 9 demonstrates how different values of *a* relate to obtained movement gradients and fits of the original range model for each of these participants. Note that the high-*a* participant committed more long-distance serial ordering errors than did the low-*a* participant.

Fig. 8b demonstrates the influence of *t'* on speed-accuracy functions using the data of Participant 2 (low *t'*) and 4 (high *t'*) from Table 1 (cf. Fig. 3). These participants are similar with re-

A



B



C

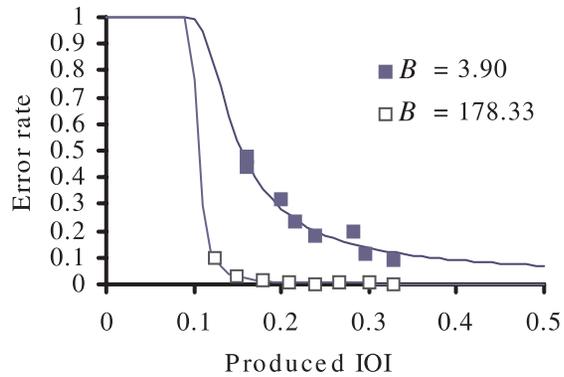


Fig. 8. Speed–accuracy data from participants representative of differences in parameters for (a)  $a$ , (b)  $t'$ , and (c)  $B$ .

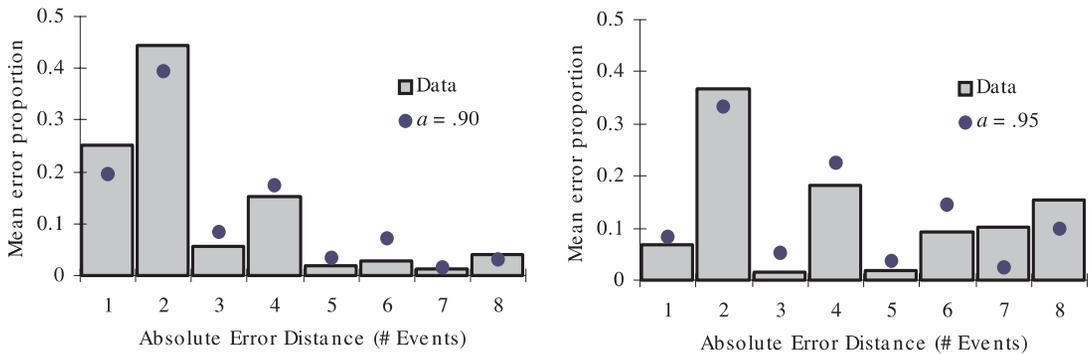


Fig. 9. Movement gradients for individual participants represented in Fig. 7a, who represent error proportions reflecting low (left) and high (right) values of  $a$ .

spect to the best fitting  $a$  parameter ( $a = .92$  for each), but differed somewhat with respect to  $B$  ( $B$  for Participant 2 = 12.48, for Participant 4 = 2.48). As a result the two functions differ with respect to asymptotic performance. The  $t'$  parameter influences the difference in offset along the  $x$  axis between the two functions, which causes the function for Participant 4 to reach  $p(\text{err}) = 1$  at a larger (slower) IOI than the function for Participant 2. We suggest that this difference happens because the high- $t'$  performer (Participant 2) reaches the upper limit of his or her possible speed at around .20 sec per IOI.

Finally, Fig. 8c demonstrates the influence of  $B$  using the data of Participants 3 (high  $B$ ) and 5 (low  $B$ ), also shown in Fig. 8a. As with manipulations of  $a$ , differences in  $B$  influences the shape of the speed–accuracy function, but have a stronger influence than  $a$  does on the asymptotic level of accuracy, observed here in the separation between the lines. Both participants exhibited similar values of the  $a$  and  $t'$  parameters ( $M = 0.90$  and  $M = 0.08$ , respectively).

### 5.3.2. Model complexity

Because we fit the two extended parameters,  $B$  and  $t'$ , simultaneously, our initial fits do not establish the necessity of both parameters (Palmer & Pfordresher, 2003, established the importance of both original model parameters). Furthermore, the additional parameters add complexity to the model that may undermine its generalizability (Pitt, Myung, & Zhang, 2002). Are both parameters necessary? We addressed this issue by comparing the model fits with both parameters to model fits with one or both of the extended model parameters ( $B$ ,  $t'$ ) fixed at its default value. Default values are those that are consistent with the basic extended model framework illustrated by Equation 3b, in which  $B = 1$  and  $t' = 0$ . We used the Akaike Information Criterion (AIC; Akaike, 1973) to determine the best model.<sup>5</sup> AIC is preferable to VAF as a goodness-of-fit measure when one compares models of different complexity, because AIC takes into account number of parameters in the model in addition to its goodness of fit. Lower AIC indicates a better fit. AIC is an ordinal measure that simply determines which model “wins,” thus we did not calculate significance tests for different mean AIC values.

Table 2 shows mean AIC, VAF, and Pearson's  $r$  for model fits to the mean data (see Fig. 5) for models in which different numbers of parameters were allowed to vary. The full model,

Table 2

Mean goodness-of-fit for extended model fits with differing numbers of parameters

Parameters	AIC	VAF (%)	Pearson's $r$
All ( $a$ , $t'$ , $B$ )	-44.36	99	.99
$t'$ fixed ( $a$ , $B$ )	-25.92	72	.97
$B$ fixed ( $a$ , $t'$ )	-1.88	< 0	.96
$a$ only	0.12	< 0	.96

Note. AIC = Akaike Information Criterion; VAF = variance accounted for.

with two free parameters ( $B$ ,  $t'$ , with values of  $a$  and  $w_2$  fixed) provided the best fit (lowest AIC). Fits to models with reduced numbers of parameters according to AIC and VAF indicated a hierarchy of importance across parameters. Although  $t'$  improved fits, even when taking into account complexity,  $B$  appeared to have a stronger influence on fits than did  $t'$ . We also fit the models to individual data. Every participant's AIC increased (indicating a poorer fit) relative to the full model when  $B$  was fixed at 1 and only  $t'$  was fit. Similarly, AICs for 10 of 12 participants increased (indicating a poorer fit) when  $t'$  was fixed at 0, a significant margin (binomial sign test,  $p = .016$ ).

Pearson's  $r$ , in contrast to other goodness-of-fit measures, indicated good fits for all models. In particular, a good correlation resulted even when both  $B$  and  $t'$  were set to default values and the model incorporated only those parameter values that were derived from fits of the original range model to movement gradients. This result was observed for individual performances as well; no correlations between the parameter-free model and error rates fell below  $r = .95$  ( $p < .01$  for each). Why? Both VAF and AIC take into account how close in scale predicted values are to obtained values. Pearson correlations, on the other hand, only consider whether predicted and obtained values covary. The fact that  $r$  shows good fits, even when no parameters are allowed to vary, suggests that the basic shape of the speed-accuracy function (irrespective of scale) can be predicted simply by using parameters that are derived from movement errors (i.e., the original range model). Thus, the additional parameters primarily scale the speed-accuracy function. In the next section, we consider whether variability in best fitting parameters across participants has any significance beyond scaling.

### 5.3.3. Construct validity of new parameters

We next examined whether the model parameters reflect underlying psychological mechanisms proposed in the introduction by regressing measures of performance on the three model parameters that most strongly influence predicted speed-accuracy functions ( $B$ ,  $t'$ , and  $a$ ). We attempted to derive measures that relate to proposed mechanisms that were independent of both speed and accuracy (Table 1 shows mean error rates for individual participants for purposes of comparison). None of the best fitting model parameters correlated significantly with each other ( $M r = -.01$ ). Our analyses thus focus on simple correlations.

We measured domain-specific skill by using the number of repetitions participants performed to reach the zero-error criterion for each piece during the learning phase of the experiment. The number of repetitions that each participant performed is shown in Table 1, taking

into account partial repetitions (e.g., 2.5 would indicate that a performer performed an entire melody twice plus one half). We reasoned that participants who require fewer repetitions to learn pieces have acquired higher skill levels, leading to a higher optimal  $B$  parameter, than those who required more repetitions. The simple correlation between repetitions during practice and  $B$  was significantly negative,  $r = -.58, p < .05$ . Simple correlations of other parameters with number of practice repetitions were not significant.

We measured individual differences in timing ability by examining tempo error: the mean signed deviation of produced IOIs from prescribed (metronomic) IOIs across all conditions for each performer. We reasoned that performers who produced tempi that were reliably slower than the target tempo would have a higher intrinsic motor threshold than those who were accurate or performed quicker than the produced tempo (suggesting motor dexterity). The correlation between tempo error and participants' optimal  $t'$  parameters was significant,  $r = .57, p < .05$ , whereas other parameters did not correlate significantly with tempo error.

Finally, we tested one of the central assumptions of the original range model: the relation between  $a$  and overall planning scope. We measured planning scope using mean absolute range, described earlier. Mean range significantly correlated with  $a$ ,  $r = .89, p < .01$ , but was not correlated with other parameters.

## 6. Discussion

This research introduces a number of new findings. Empirically, we demonstrate speed–accuracy trade-offs in music performances by highly skilled pianists. Such trade-offs have not been demonstrated previously in this population to our knowledge. Fits of the extended range model to speed–accuracy data, tested here for the first time, suggest that changes in range of planning with tempo (found here, as in Palmer & Pfordresher, 2003) facilitate retrieval of the correct event. At the same time, the extended model suggests that accuracy is influenced by other factors that vary across individuals (although not necessarily with tempo): motor dexterity and domain-specific skill in music performance.

Timing was expected to influence accuracy, in keeping with the classic speed–accuracy trade-off. Empirical data verified this claim. Among populations with expertise in sequence production, trade-offs may only appear when individuals perform at the limits of their abilities under strict temporal demands, as in the experiment reported here. In a more general sense, the current data suggest that trained musicians performing music are subject to the same constraints as are individuals from the broader population. This observation converges with other research that suggests expertise in music performance results from the refinement of skills through deliberate practice that are present in the general population, and do not necessarily result from inherited “talents” that differ from the qualities present in the rest of the population (Ericsson et al., 1993; Howe, Davidson, & Sloboda 1998; Krampe & Ericsson, 1996).

In this section we discuss how the findings reported here support the basic assumptions of our model. Then we compare our model to other models of sequence production and of the speed–accuracy trade-off.

### 6.1. The extended range model

The research reported here supports the hypothesis that the retrieval of serial order and overall performance accuracy share a common cause: timing. Timing in the original range model allows the performer to access distant events during response preparation. The activation of the surrounding context in the extended model facilitates activation of the current event, leading to improved accuracy. How well do the data support these joint effects of timing?

The original range model proposed that slowed timing increases range of planning (Palmer & Pfordresher, 2003). This prediction extends from the assumption that incorporating the surrounding context helps performance, because musical events are defined by their context. When more time is allocated to response preparation, a performer will be able to incorporate more contextual information during planning (i.e., a broader range of planning). As a result, errors are less frequent on the whole, but those errors that do occur tend to originate from relatively distant sources. Palmer and Pfordresher (2003) supported this prediction in analyses of errors during performances at two tempi. The experiment reported here incorporated a broader range of tempi and again demonstrated increased range of planning with tempo. Change in range of planning with tempo was gradual, as predicted by the model. Whereas differences across all tempo conditions fell short of significance, differences at extreme tempi were robust. Importantly, obtained mean range values were proximal to predicted values from the range model.

The extended model introduced a further assumption about timing in both range of planning and accuracy. Specifically, increases in range of planning brought about by slowed timing were predicted to enhance the activation of the current (correct) sequence event. Fits of the extended model supported this claim. Fits of the extended model, based entirely on parameters from the original model (Equation 3b) predicted the shape of obtained speed–accuracy trade-offs as measured by Pearson's  $r$ . However, these predictions differed in scale from obtained error rates. This difference in scale suggests that variables beyond range of planning influence the accuracy of response selection. We incorporated two new parameters in the extended model to model the influence of variables beyond response preparation on overall accuracy.

Parameters related to motor dexterity ( $t'$ ) and domain-specific skill ( $B$ ) were included in the extended model of accuracy but not in the original range model fits. Neither parameter was allowed to vary with production rate (IOI). Additional parameters therefore modulate the speed–accuracy relation but are not integral parts of the predicted speed–accuracy trade-off. Objectively, these parameters serve the purpose of scaling the speed–accuracy trade-off to fit the data. However, correlations between parameters and measures of performance suggested that they do relate to underlying mechanisms involved in accuracy. Whereas  $t'$  correlated with accuracy in produced tempo, but not the number of repetitions required to learn melodies,  $B$  correlated with repetitions during learning but not tempo accuracy. Thus we consider the new parameters to serve a purpose beyond curve fitting. Despite the contribution of extended range model parameters to the model's fit, the fact that extended model predictions correlated well with obtained error rates even when additional parameters were fixed at their default values suggests that the most fundamental source of speed–accuracy trade-offs in music performance is the way in which planning constrains range of planning.

## 6.2. Comparisons with other approaches

In this section, we compare the predictions of the extended range model to similar models that have addressed speed–accuracy trade-offs or serial order. The range model accounts for retrieval during the production of musical sequences. We know of no similar model in the domain of music performance. However, models of serial order in speech production address task components that we consider fundamental to music production. Both behaviors concern retrieval during the production of auditory sequences in which both event timing and event contents (e.g., pitches or syllables) are used to communicate an intended message. Importantly, speech and music production differ from behaviors like typing, in which timing does not contribute to the communicated message. Other models of the speed–accuracy trade-off focus on single, rather than serially ordered, responses in contexts that focus specifically on motor constraints while minimizing memory load (e.g., Beilock, Bertenthal, McCoy, & Carr, 2004; Fitts, 1954; D. E. Meyer, Smith, & Wright, 1982; Plamondon & Alimi, 1997; Schmidt et al., 1979), or focus on memory constraints while minimizing motoric factors (e.g., Boldini, Russo, & Avons, 2004; McElree, 2001; Ratcliff, 1978; Sternberg, 1969).

## 6.3. Models of sequence production and errors

Table 3 compares the predictions made by the range model to other models of serial production in speech. To review, the primary phenomena predicted by the range model are frequencies of movement errors as a function of error distance (i.e., movement gradients), overall accuracy, and the speed–accuracy trade-off. We thus focus on these properties of the models summarized in Table 3. Note that the speech production models that we compare to the range model were designed to make additional predictions that are not summarized in the table.

Dell (1986; see also Dell, 1985; Dell et al., 1997) proposed a model of speech production in which activation spreads through a hierarchical network that represents the linguistic structure. When an utterance is planned, activation begins at nodes representing the highest level of the hierarchy (semantics). Frames establish the serial order of linguistic categories within each hierarchical level across the utterance. Over time, activation spreads to different nodes in the frame that are associated with specific categories, and speech errors can occur when random fluctuations in activation lead to the production of an incorrect node (which represents a lin-

Table 3  
Comparisons of the range model with other sequence production models

Predictions	Dell (1986)	MacKay (1982)	Vousden et al. (2000)	Extended Range
Overall accuracy	✓ <sup>a</sup>	✗	✗	✓
Speed–accuracy trade-off	✓ <sup>a</sup>	✓ <sup>a</sup>	✗	✓
Error distance	✗	✓	✓	✓

*Note.* Check marks (✓) indicate that the model makes testable predictions for the behavioral phenomenon in question, and an ✗ indicates the absence of such a prediction.

<sup>a</sup>Not tested empirically.

guistic unit) from a given category. Thus, the activation of many different nodes, some but not all of which are associated with correct events, increases over time. At slow speaking rates, the likelihood that a correct node's activation surpasses that of an incorrect node is greater than at fast speaking rates, leading to a predicted speed–accuracy trade-off. This prediction has not been tested quantitatively to our knowledge.

As can be seen in Table 3, Dell's (1986) model predicts overall accuracy and speed–accuracy trade-offs, as does the extended range model, although not all of these effects are tested. We, like others (Vousden et al., 2000) characterize Dell's model as not accounting for movement errors with respect to error distance. This claim warrants some discussion. Dell's model does predict different distances for phoneme versus word errors, based on the fact that these errors originate at different levels of the linguistic hierarchy (cf. Garrett, 1980). However, it is not clear whether the model would predict the kind of variations of error distance with speaking rate as does the extended range model.

Another model that was designed to account for both serial ordering errors and error rates in sequence production was node structure theory (MacKay, 1982, 1987). An account of speed–accuracy trade-offs from node structure theory was detailed in MacKay (1982), whereas the account for serial ordering errors was detailed in MacKay (1987). Like Dell's (1986) model, node structure theory has primarily been applied to speech, and incorporates spreading of activation (termed *priming*) through a hierarchical representation, using “sequence nodes” in place of Dell's frame-based representation. Oddly, the model makes slightly different claims with respect to the basis of the speed–accuracy trade-off and serial ordering errors. With respect to serial ordering errors, node structure theory presents a similar account to that offered by Dell (1986): Activation of related nodes (some correct, some not) increases over time, and with more time the activation of the correct node will exceed that of incorrect nodes by a critical amount (MacKay, 1987). However, the proposed account of speed–accuracy trade-offs (MacKay, 1982) suggests that the activity of nodes associated with errors remains at a constant level, whereas the activation of the correct node increases with time (a similar proposal was outlined by McClelland, 1979). Activity of incorrect nodes thus functions as a background “noise” level that the correct node must exceed by a critical amount for the correct event to be produced, which takes time. By contrast, the range model uses the same assumptions about the activations of events other than the current event in modeling serial ordering errors and in modeling overall error rates. As with Dell (1986), we know of no formal test of node structure theory's predicted speed–accuracy trade-off (although we propose one later). Moreover, the model makes no quantitative predictions regarding overall accuracy and does not predict error distance beyond distinctions between phoneme and word errors (see Table 3).

The extended range model is similar in a number of ways to a model that does account for distance relations but was not designed to account for speed–accuracy trade-offs, the oscillator model of speech production proposed by Vousden and colleagues (2000; Brown et al., 2000; for a similar approach see Church & Broadbent, 1990). This model includes two primary components—a phoneme feature vector and a phonological context vector—with phoneme sequences arising from associations between the two over time. The phonological context vector encodes the passage of time as a function of the products of many sinusoidal oscillators. Patterns resulting from summed oscillator activity result in greater similarities for close points in time than for distant points in time, like the range model's serial component, and oscillations

can be manipulated to create periodicities in temporal context that resemble the product of serial and metrical components in the range model. Indeed, Palmer and Pfordresher (2003) suggested that such an oscillatory mechanism may underlie metrical similarity (cf. Gasser, Eck, & Port, 1999; Large & Jones, 1999; Large & Kolen, 1994; Large & Palmer, 2002; McAuley, 1995).

The oscillator model of Vousden et al. (2000), however, does not address production rate and thus is not designed at present to predict a speed–accuracy trade-off, nor does it predict overall accuracy (see Table 3). Nevertheless, increases in production rate may result in associations of events with more similar states of the context vector, assuming that the component oscillator periods do not vary with production rate. Thus, it is possible (although not verified) for such an oscillator approach to predict speed–accuracy trade-offs. Unlike the models discussed previously, this model was designed explicitly to account for error distances. Its predictions differ in orientation from the range model, however. Rather than focus on the role of production rate in error distance, Vousden and colleagues accounted for distances as a function of whether errors are anticipatory and perseveratory, a distinction that appears not to influence errors in music performance in the same way as in speech (Palmer & Pfordresher, 2003).

#### 6.4. *Test of an alternative model based on node structure theory*

The account of speed–accuracy trade-offs developed in node structure theory (MacKay, 1982), although not quantitatively specified, is based on a related model that yields testable predictions: the cascade model of McClelland (1979; cf. Wickelgren, 1977). The cascade model, unlike node structure theory, was originally designed to account for single responses rather than the production of sequences. Theoretically, the link between these models suggests that constraints on the retrieval of single responses that lead to speed–accuracy trade-offs may also lead to speed–accuracy trade-offs in the production of complex sequences. This implication is intuitively compelling and parsimonious, and response preparation and selection in music performance may be related in a cascaded fashion (Palmer, 2005). However, we question whether such a link is plausible given the need to plan multiple responses at once during sequence production. In this section we compare the predictions of the range model with an alternative based on the single-response approach as proposed by MacKay.

The cascade model is based on assumptions about multiple processes that overlap in time. The full model is highly complex and involves a series of equations. However, the speed–accuracy trade-off that it predicts has been shown to reduce to a simple exponential activation function under certain conditions (cf. Equation 12, McClelland, 1979):

$$Y = 1 - e^{-k(t-T)} \quad (6)$$

In which  $k$  establishes the rate of change (a free parameter),  $t$  is the duration of time that is allowed to make a decision (a variable identical to  $t$  in the range model), and  $T$  is a temporal shift caused by overlapping processes (a free parameter similar to our  $t'$  parameter).

We adopted this framework in testing a plausible alternative to the range model of planning. Because this equation relates to the activation of the correct event, it predicts accuracy, but the inverse could obviously be used to predict error rates, given that  $0 \leq Y' \leq 1$ . We also included an

additional asymptote parameter ( $\alpha$ ), for two reasons. First, we wanted to give the alternative model optimum flexibility. Second, the form of the model we use is identical to a descriptive model proposed by Wickelgren (1977) that has frequently been applied to research in memory retrieval that uses the “response signal” method to generate speed–accuracy functions (e.g., Boldini et al., 2004; McElree, 2001). Predicted accuracy from the cascade model (which results from a modification of the preceding equation that is less easily applicable to our data), were shown to be very similar to the predictions of Wickelgren’s (1977) model. Thus the resulting equation for error rates for the alternative model tested here is:

$$p(\text{err}) = 1 - p(\text{acc}) = 1 - [\alpha * (1 - e^{-k(t-T)})] \quad (7)$$

which is the inverse of the exponential approach to the limit,  $\alpha$ . The form of this equation assumes that the data set comprises proportions ( $0 < Y < 1$ ). This alternative model constitutes a viable contender to the extended range model, given its historical use in the speed–accuracy literature.

We fit all three parameters of the alternative model simultaneously through the least squares optimization procedure in MATLAB. The fit of the alternative model to mean data is shown in Fig. 10. As can be seen, the fit is good (VAF = 92%), although not as good as the fit of the extended range model (VAF = 99%). A  $t$  test on the associated Pearson correlations ( $r$  with data for Wickelgren model = .96, for extended range model = .99) yielded a significant difference,  $t(5) = -2.59, p < .05$  (test for dependent  $r$ s; Cohen & Cohen, 1983). In contrast to the extended range model, the alternative model tended to underestimate error rates at extreme tempi because it predicts a function that is less bowed than the function predicted by the extended range model. Although the alternative model has one fewer free parameters than the extended range model (given the conservative assumption that the extended range model uses four free parameters), the alternative model’s poorer fit was not an artifact of complexity according to AIC, which was higher for the alternative model (AIC =  $-34.86$ ) than for the extended range model

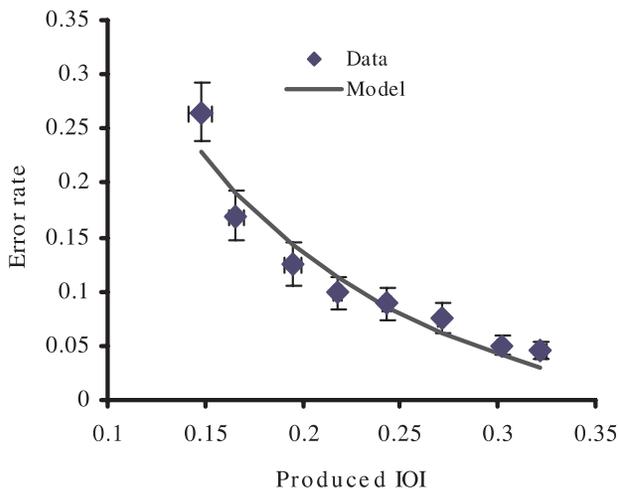


Fig. 10. Best fitting speed–accuracy function based on an alternative model (Equation 7), with mean data. Best fitting model parameters are  $\alpha = 1.0261, k = 8.63, T = -0.01$ . Error bars represent  $\pm 1 SE$ .

(AIC =  $-45.34$ ). Likewise, fits of the extended range model to individual data were superior to fits of the alternative model in a majority of cases (7 of 12), which suggests that the superior fit of the range model to mean data was not an artifact of averaging (cf. Anderson & Tweeney, 1997; Myung, Kim, & Pitt, 2000). Finally, we note that two of the best fitting parameter values are counterintuitive: The asymptote is greater than 1 (suggesting accuracies of greater than 100%), and the offset is negative. We allowed such values (which do not deviate greatly from 1 and 0, respectively) to accommodate for the fact that the model was originally designed to fit  $d'$  rather than  $p(\text{err})$ .<sup>6</sup>

### 6.5. *Limitations and alternative approaches*

The extended range model represents a new line of theorizing in a simple, easily testable framework. The model makes minimal assumptions about the cognitive representation involved and about item-specific information. The new model adds only two new parameters to the original model. Given the complexity of the behavior being modeled, certain aspects of planning and execution are not incorporated in the range model, some of which are discussed here.

It is possible that the limitation of skill and motor dexterity to response selection may be too restrictive; range of planning may be influenced by motor dexterity or the ability of more skilled performers to organize stimulus information into higher order units (e.g., Chase & Simon, 1973; Ericsson & Kintsch, 1995). These limitations may account for the fact that the relation between tempo and range of planning was only marginally reliable when data were averaged across participants. Even so, the range model does not appear to predict large differences and the data conform to this prediction (Fig. 7).

The range model also does not account for possible contributions of motor movements to serial ordering errors, instead linking errors to abstract sources of similarity and proximity among sequence events in memory. By contrast a class of models not summarized in Table 3 relates sequencing errors directly to finger movements. For example, Rumelhart and Normal (1981) produced an effector-based model of skilled typing that could, in principle, be adapted to account for speed-accuracy trade-offs in music performance. Their model focuses on the way in which finger movements constrain both speed and accuracy. Because typing is typically produced as quickly as possible, they do not consider the role of rate and their model thus does not predict speed-accuracy trade-offs, although it could be adapted to do so.

An important question for the range model therefore concerns whether confusions among finger movements can account for errors in production. We allowed pianists to choose their own fingering and therefore are not able to incorporate fingering explicitly in current model fits. However, another recent study did address the possibility that fingering contributes to serial ordering errors in music performance (Pfordresher & Palmer, 2006). That study incorporated simpler stimuli with prescribed fingering. Pfordresher and Palmer (2006) found patterns of serial ordering errors across distances 1 through 3 that fit the assumptions of the range model. However, an alternative model that replaced the metrical similarity metric with one based on the proximity of fingers did not match obtained patterns of serial ordering errors. It is likely that similar results would obtain for the current data.

The effector independence assumed by the range model converges with research demonstrating that transfer of sequence learning occurs across effectors (e.g., Grafton, Hazeltine, &

Ivry, 1998; Keele, Cohen, & Ivry, 1990; R. K. Meyer & Palmer, 2003; Palmer & R. K. Meyer, 2000). In such cases, transfer is governed by task similarity based on sequence structure rather than the kinds of actions used to generate sequences (cf. MacKay & Bowman, 1969). Effector independence is also supported by similarity in error patterns across domains that involve very different actions: speech and music (viz. piano performance). Similar patterns of serial ordering errors have been found in different domains, such as speech production (e.g., Garrett, 1980), typing (e.g. Rumelhart & Norman, 1982), and piano performance (e.g., Palmer & van de Sande, 1993, 1995), and similar patterns of serial ordering errors have been found when the same participants produce speech or sequences of key presses (MacKay, 1971; Rosenbaum et al., 1986).

## 7. Conclusions

The performance of music exemplifies a serial ordering task of the highest difficulty; sequences comprise events that recur in different contexts and must be produced with great rapidity and temporal precision. Even so, musicians are capable of performing at very fast rates with very low error rates, often circumventing the traditional speed–accuracy trade-off relation (although not always, as our data demonstrate). We have described a model that accounts for this behavior in terms of a contextual representation that predicts both error rates (accuracy) and error types (serial order) with parameters that model working memory, domain-specific skill, and motor dexterity among performers. The fundamental assumption of the extended range model is that the errors in production originate in the incremental retrieval of events over time. Thus serial order and accuracy may both share a common source: the time over which the contextual representation of a sequence is updated during production.

The extended range model makes assumptions that are consistent with past work but do not fully resemble any single model. An important issue for other models of speed–accuracy trade-offs to consider is the role of the surrounding context, which may facilitate rather than interfere with event retrieval. In addition, the approach presented here suggests that individual differences in speed–accuracy trade-offs during complex tasks may emerge from multiple sources, including both motoric and cognitive factors.

## Notes

1. By *event*, we refer to a stimulus, or stimuli that cooccur in time; events in music are single notes or chords. Although events are commonly considered to be perceptual, recent research suggests that actions may be conceptualized with respect to the resulting perceived stimulus (Hommel, Müsseler, Aschersleben, & Prinz, 2001). Based on this research, we use the term *event* to refer interchangeably to actions and consequent sounds in performance.
2. We note that this treatment differs slightly from the use of these terms in other research (e.g., Osman, Moore, & Ulrich, 2003) in which preparation is exclusively motoric in origin and follows response selection in time. Such research often incorporates a go–no

go paradigm in which selection involves a decision regarding which effector to use (e.g., left vs. right hand), and response preparation involves the initiation of that movement. By contrast, the need to select multiple responses in a sequence may reverse the temporal order of these processes.

3. Values for extended range model parameters ( $t'$ ,  $B$ ) used to generate model predictions in Figs. 3 and 4 reflect the range of fitted values in the reported experiment and were based on preliminary fits.
4. VAF was computed as  $VAF = 1 - (SSE / SST) = 1 - [\Sigma(\text{data} - \text{model})^2 / \Sigma(\text{data} - M)^2]$ , where  $SSE$  = sum of squares–error,  $SST$  = sum of squares–total, and  $M$  = mean for obtained data (Ruml & Caramazza, 2000). Note that it is mathematically possible to obtain a negative VAF if  $SSE > SST$ , even though the interpretation of VAF is similar to  $r^2$ , which cannot result in a negative value.
5. We estimated AIC using sum of squared errors:  $AIC = n * \ln SSE + 2 * k$  (Myung & Pitt, 1998), where  $n$  = number of observations (which is 8 for all fits reported here) and  $k$  = number of parameters.
6. Measures of percentage correct and  $d'$  often do not converge (MacMillan & Creelman, 1991). Nevertheless, we considered the extension of this model to error rates to be reasonable. Wickelgren's equation predicting  $d'$  is highly similar to the "limiting case" activation equation proposed by McClelland (1979; see Equation 6). This equation predicts activation of the correct event, a pattern that should converge with  $p(\text{correct})$  and by extension  $p(\text{err})$ . Conversely, limiting predicted activation functions to  $d'$  seems overly restrictive.

## Acknowledgments

This research was sponsored in part by National Science Foundation Grant BCS–0344892 to Peter Q. Pfordresher and by the Canada Research Chairs program, National Sciences and Engineering Research Council of Canada Grant 298173 and National Institute of Mental Health Grant R01–45764 to Caroline Palmer. Portions of these results were presented at the 2003 Meeting of the Psychonomic Society and the 2003 meeting of the Society for Music Perception and Cognition.

We thank Bruno Repp, Renee Timmers, and an anonymous reviewer for helpful comments on an earlier version of this article.

## References

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrox & F. Caski (Ed.), *Second international symposium on information theory* (pp. 267–281). Budapest, Hungary: Akademiai Kiado.
- Anderson, R. B., & Tweeney, R. D. (1997). Artfactual power curves in forgetting. *Memory & Cognition*, 25, 724–730.

- Beilock, S. L., Bertenthal, B. I., McCoy, A. M., & Carr, T. H. (2004). Haste does not always make waste: Expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychonomic Bulletin & Review*, *11*, 373–379.
- Bock, K. (1995). Sentence production: From mind to mouth. In J. L. Milles & P. D. Eimas (Eds.), *Handbook of perception and cognition: Vol. 11. Speech, language, and communication* (pp. 181–216). Orlando, FL: Academic.
- Boldini, A., Russo, R., & Avons, S. E. (2004). One process is not enough: A speed–accuracy tradeoff study of recognition memory. *Psychonomic Bulletin & Review*, *11*, 353–361.
- Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*, 127–181.
- Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York: Academic.
- Church, R. M., & Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition*, *37*, 55–81.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*. Chicago: University of Chicago Press.
- Dell, G. S. (1985). Positive feedback in hierarchical connectionist models: Applications to language production. *Cognitive Science*, *9*, 3–23.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.
- Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, *104*, 123–147.
- Drake, C., & Botte, M. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics*, *54*, 277–286.
- Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition*, *74*, 1–32.
- Duke, R. A. (1989). Musicians' perception of beat in monotonic stimuli. *Journal of Research in Music Education*, *37*, 61–71.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.
- Fitts, P. M. (1954). The information capacity of the motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *67*, 381–391.
- Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America*, *98*, 2524–2530.
- García-Albea, J. E., del Viso, S., & Igoa, J. M. (1989). Movement errors and levels of processing in sentence production. *Journal of Psycholinguistic Research*, *18*, 145–161.
- Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production: Speech and talk* (pp. 177–220). London: Academic.
- Gasser, M., Eck, D., & Port, R. (1999). Meter as mechanism: A neural network model that learns metrical patterns. *Connection Science*, *11*, 187–216.
- Grafton, S. T., Hazeltine, E., & Ivry, R. B. (1998). Abstract and effector-specific representations of motor sequences identified with PET. *The Journal of Neuroscience*, *18*, 9420–9428.
- Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 644–655.
- Hibi, S. (1983). Rhythm perception in repetitive sound sequence. *Journal of the Acoustical Society of Japan*, *4*, 83–95.
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–937.
- Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences*, *21*, 399–442.

- Keele, S. W., Cohen, A., & Ivry, R. (1990). Motor programs: Concepts and issues. In M. Jeannerod (Ed.), *Attention & performance XIII* (pp. 77–110). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201–258.
- Krampe, R. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, 125, 331–359.
- Large, E. W. (1993). Dynamic programming for the analysis of serial behaviors. *Behavior Research Methods, Instruments, & Computers*, 25, 238–241.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119–159.
- Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, 6, 177–208.
- Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, 26, 1–37.
- Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8, 249–336.
- MacKay, D. G. (1971). Stress pre-entry in motor systems. *American Journal of Psychology*, 84, 35–51.
- MacKay, D. G. (1982). The problems of flexibility, fluency, and speed–accuracy trade-off in skilled behavior. *Psychological Review*, 89, 483–506.
- MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag.
- MacKay, D. G., & Bowman, R. B. (1969). On producing the meaning in sentences. *American Journal of Psychology*, 82, 23–39.
- MacKenzie, C. L., & Van Eerd, D. L. (1990). Rhythmic precision in the performance of piano scales: Motor psychophysics and motor programming. In M. Jeannerod (Ed.), *Attention & performance XIII* (pp. 357–408). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge, UK: Cambridge University Press.
- McAuley, J. D. (1995). *On the perception of time as phase: Toward an adaptive-oscillator model of rhythm*. Unpublished doctoral dissertation, Indiana University, Bloomington, IN.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287–330.
- McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 817–835.
- Meyer, D. E., Smith, J. E., & Wright, C. E. (1982). Models for the speed and accuracy of aimed movements. *Psychological Review*, 89, 449–482.
- Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
- Meyer, R. K., & Palmer, C. (2003). Temporal and motor transfer in music performance. *Music Perception*, 21, 81–84.
- Myung, I. J., Kim, C., & Pitt, M. A. (2000). Toward an explanation of the power-law artifact: Insights from response surface analysis. *Memory & Cognition*, 28, 832–840.
- Myung, I. J., & Pitt, M. A. (1998). Issues in selecting mathematical models of cognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 327–355). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Osman, A., Moore, C. M., & Ulrich, R. (2003). Temporal organization of covert motor processes during response selection and preparation. *Biological Psychology*, 64, 47–75.
- Palmer, C. (2005). Time course of retrieval and movement preparation in music performance. *Annals of the New York Academy of Sciences*, 1060, 360–367.
- Palmer, C., & Drake, C. (1997). Monitoring and planning capacities in the acquisition of music performance skills. *Canadian Journal of Experimental Psychology*, 51, 369–384.

- Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 728–741.
- Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. *Psychological Science*, 11, 63–68.
- Palmer, C., & Pfordresher, P. Q. (2000). The faces of meter in performance and perception. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the 6th international conference on music perception and cognition* [CD-ROM]. Keele, UK: Keele University Department of Psychology.
- Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review*, 110, 683–712.
- Palmer, C., & Schendel, Z. A. (2002). Working memory constraints in sequence production. *Abstracts of the Psychonomic Society*, 7, 30.
- Palmer, C., & van de Sande, C. (1993). Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 19, 457–470.
- Palmer, C., & van de Sande, C. (1995). Range of planning in skilled music performance. *Journal of Experimental Psychology: Human Perception & Performance*, 21, 947–962.
- Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11, 409–464.
- Peters, M. (1989). The relationship between variability of intertap intervals and interval duration. *Psychological Research*, 51, 38–42.
- Pew, R. W. (1969). The speed–accuracy operating characteristics. *Acta Psychologica*, 30, 16–26.
- Pfordresher, P. Q., & Palmer, C. (2006). Effects of hearing the past, present, or future in music performance. *Perception & Psychophysics*, 68, 362–376.
- Pitt, M. A., Mung, I., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, 109, 472–491.
- Plamondon, R., & Alimi, A. M. (1997). Speed/accuracy trade-offs in target-directed movements. *Behavioral and Brain Sciences*, 20, 279–349.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Rosenbaum, D. A., Weber, R. J., Hazelett, W. M., & Hindorff, V. (1986). The parameter remapping effect in human performance: Evidence from tongue twisters and finger fumlbers. *Journal of Memory and Language*, 25, 710–725.
- Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science*, 6, 1–36.
- Ruml, W., & Caramazza, A. (2000). An evaluation of a computational model of lexical access: Comment on Dell et al. (1997). *Psychological Review*, 107, 609–634.
- Schmidt, R. A., Zelaznik, H. N., Hawkins, B., Frank, J., & Quinn, J. T. (1979). Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review*, 86, 415–451.
- Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 185–203.
- Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, 73, 205–246.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153, 652–654.
- Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 57, 421–457.
- Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology*, 41, 101–175.
- Wheeldon, L. R., Meyer, A. S., & Smith, M. (2002). Incrementality. In R. Goldstone (Ed.), *Encyclopedia of cognitive science*. London: Macmillan.
- Wickelgren, W. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 41, 67–85.
- Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychological Monographs*, 3.

## Appendix

The range model combines the serial component with a metrical component that reflects similarity between the metrical accent strength of the current event and surrounding events. Metrical accents typically alternate between strong and weak (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983; Liberman & Prince, 1977). The similarity between the current event ( $i$ ) and an event at some distance ( $x$ ) takes this form:

$$M_{x,i} = 1 - \frac{|m_i - m_{i+x}|}{m_i + m_{i+x}} \quad (\text{A1})$$

Where  $M$  is an array of similarity relations by position and distance (where distance  $x$  can be positive for future events or negative for past events) and  $m$  is a variable that reflects metrical accent strength in metrical grid notation (Lerdahl & Jackendoff, 1983; Liberman & Prince, 1977; Palmer & Pfordresher, 2003). According to Equation A1, metrical similarity is the inverse of the absolute difference in metrical accent strength between the two positions, divided by the sum of accent strengths. This relative similarity metric is a generalized form of Weber's law, and was found by Palmer and Pfordresher (2003) to provide better fits to movement error data than an absolute similarity metric based on the numerator alone. The result of applying this equation to binary meters is that events separated from the current event by multiples of 2 are more similar to the current event, leading to higher activations for events at these distances than for events at other distances. This general rule holds for each sequence position, although the specific pattern of activation strengths across distance in the metrical component varies depending on the metrical accent associated with the current position (see Palmer & Pfordresher, 2003, for details).

Finally, metrical weights ( $w_j$ ) increase the salience of one level in the metrical grid, and correspond to the level at which people would clap or tap to the beat of a musical sequence (Duke, 1989; Parncutt, 1994). Such behaviors indicate that a certain level ( $j$ ) in the metrical hierarchy receives prominence; such levels are referred to as the *tactus* in music (e.g., R. K. Meyer & Palmer, 2001). In the range model the *tactus* level may be weighted more highly than other metrical levels through an additional free parameter  $w_j$ , such that:

$$m_i = \sum_j^k w_j \cdot g_{j,i} \quad \text{where} \quad \sum_j^k w_j = 1, \quad \text{and} \quad 0 < w_j < 1 \quad (\text{A2})$$

where  $g$  indicates the presence ( $g = 1$ ) or absence ( $g = 0$ ) of a metrical accent at each metrical level ( $j$ ) and  $k$  is the total number of levels ( $k = 4$  in Fig. 1). In the metrical grid notation of Fig. 1 (shown below the music notation),  $g = 1$  for positions and levels marked with an X. The weight ( $w$ ) for  $j = \textit{tactus}$  is assumed to exceed  $1/k$  (equal weights across levels), and the weights of remaining levels are equally divided from the quantity  $1 - w_{j = \textit{tactus}}$  to ensure that weights across all levels sum to 1. We chose to vary weights exclusively on Level 2 ( $w_{\textit{tactus}} = w_2$ ) in the fits reported here, based on previous fits to performances with the same stimuli, which indicated that Level 2 rather than other metrical levels is chosen as the *tactus* (Palmer & Pfordresher, 2003).

Note that Fig. 1 shows event activations at a single point in time and at a single sequence position. However, the range model predicts that the distribution of activations changes across sequence positions, due to change in metrical similarity relations across position (*i*). Furthermore, although the range model predictions shown in Fig. 1 are symmetrical with respect to past and future events, this symmetry does not hold at each sequence position for the same reason (although it does hold for average activations across all positions). Fits to different sequence positions are shown in Palmer and Pfordresher (2003).