

Analysis of timing variability in human movements by aligning parameter curves in time

Lisa K. Maurer¹ · Heiko Maurer¹ · Hermann Müller¹

Published online: 8 August 2017
© Psychonomic Society, Inc. 2017

Abstract The analysis of timing in human movements requires a reference with which timing can be quantified. In reactive movements this reference is given by the stimulus. However, many movements do not respond to such an external event. In throwing, for instance, the hand opening for release has to be timed to an acceleration of the throwing arm. A common approach to analyzing release-timing variability is to choose a landmark in the movement that is supposed to have a fixed temporal relation to the release. Such distinct landmarks, however, are not always well definable. Therefore, the present article describes an alternative approach analyzing timing variability on the basis of the alignment of different trials relative to their kinematic shape, by shifting the trials in the time domain. The basic assumption behind this approach is that single throwing movements are one instance of an acquired movement template, and thus show a considerable similarity. In contrast, the location of the temporal moment of release varies from trial to trial, generating imprecision regarding the release timing. In trials synchronized with respect to the release, this variability can be assessed by shifting the kinematic profiles of the throwing movements in time such that they superimpose as closely as possible. As a result, the corresponding time shifts for all trials represent a measure of the release time deviations across trials, and the standard deviation of these deviations represents the timing variability. Aside from timing analyses in such movements as throwing, the approach can be applied to very different tasks with timing demands—for example, to neurophysiological signals.

Keywords Timing · Throwing · Release · Precision · Variability · Kinematic profile alignment

The ability to generate well-timed movements is critical to a wide range of human behaviors such as driving a car or riding a bicycle, playing musical instruments or to perform high-performance skills like high diving. When studying human movement timing, often behavioral reaction times to a stimulus or tap to beat intervals in rhythmic tapping tasks are used to quantify timing performance. In real-world settings this would apply, for instance, to reacting to other road users in traffic or to keeping pace with a metronome when playing the piano. However, movement timing is not limited to reactions. We often have to combine two or more movements or submovements involved in an action in a timely fashion with no external event providing a timing reference, like clapping both hands. Furthermore, there are even movements that, in addition to having a missing external timing reference, also have no single detectable internal event that could represent a consistent position in time. This makes it difficult to examine the timing of different submovements performed concurrently—for instance, coupling the flexion of the wrist to arm elevation when pouring a cup of coffee, or measuring how high-divers combine a somersault with a twist. A task that has been widely studied with respect to such intrinsic timing is throwing, in which the hand opening for release has to be coupled to the accelerating movement of the throwing arm.

An evaluation of the quality of timing in throwing can be carried out using two different measures related to variability: precision and/or accuracy. Precision means the closeness of agreement between repeated measurements under specific conditions (Bureau International des Poids et Mesures [BIPM], 2012), which is usually expressed by measures of imprecision—for example, the standard deviation. Accuracy,

✉ Lisa K. Maurer
lisa.k.maurer@sport.uni-giessen.de

¹ Department of Psychology and Sport Science, Justus Liebig University, Giessen, Germany

on the other hand, is the “closeness of agreement between a measured quantity value and a true quantity value of a measurand” (BIPM, 2012). In goal-oriented throwing for such games as darts or basketball, both the timing precision and timing accuracy are determinants of the success rate, and both need to be controlled. For instance, if a thrower is able to precisely reproduce the release time but his throws are not accurate, he will systematically miss the target. On the other hand, when precision is poor, 100% accuracy can never be reached. Concretely, if the throws are on average accurately on target but prone to unsystematic fluctuations (noise), the thrower may never hit the target. Although random errors can at least partly be compensated using certain movement strategies (Müller & Loosch, 1999; Nasu, Matsuo, & Kadota, 2014). Thus, imprecision can be considered the lower limit for inaccuracy.

In our research on timing variability we focus on precision measures because we are primarily interested in these random timing errors. They can give insights into intrinsic variability of timing in the control of submovements and its development with learning, maturation, aging, and diseases (like Parkinson). In throwing, release timing is of special interest because the time window for the release is relatively short (between 12 and 45 ms; with reference to own pilot data) and humans can nevertheless throw with a precision of, depending on the task, 3–15 ms (standard deviation; Becker, Kunesch, & Freund, 1990; Hore, Watts, Martin, & Miller, 1995). The quantification of release timing precision is challenging because throws need to be collocated in time and the release can obviously not be used as a reference point. The most common method is to determine the variability of the temporal moment of release relative to other landmarks of the movement. Possible landmarks can be the movement onset/movement initiation (Chowdhary & Challis, 1999) or different arm motion kinematic parameters during the throw like the zenith of the hand or forearm trajectory (Hore, Timmann, & Watts, 2002; McNaughton, Timmann, Watts, & Hore, 2004; Nasu et al., 2014; Smeets, Frens, & Brenner, 2002), the peak speed of the movement (Smeets et al., 2002), the crossing of a certain threshold of the angular hand/forearm velocity or of the elbow or wrist extension/flexion as well as a threshold crossing of the elbow/wrist extension/flexion velocity (McNaughton et al., 2004). The underlying assumption of these approaches is that the coupling between release and the corresponding landmark is stable, or at least is supposed to be stable to accomplish the task (stability requirement), and variability between throws arises solely from variations in the duration of the landmark-to-release interval. However, there are problems with this assumption. First, in some cases the suitable landmarks are relatively far (in time) from release, creating a risk of violating the assumption of a stable relative time interval between the two events. Second, landmarks are more or less arbitrary chosen by the experimenter—that is, often without testing the stability

requirement. And third, some throwing movements do not even offer distinct landmarks that could be used as a reference. But even in those cases in which a landmark exists, the production of the landmark itself is prone to fluctuations as well. If we use a variable landmark as reference, its variability is then falsely attributed to release variability. Moreover, it is unclear whether throwers really try to stabilize exactly this landmark.

In cases in which any of these reservations apply, a coupling between release and movement might be better achieved by not relying on a single moment in time, but rather making use of the similarity of a larger part of the kinematic movement profiles (e.g., for a defined time window around release). We suggest shifting the movement profiles relative to each other in the time domain in order to superimpose the movement profiles in the specified time window. Hence, the magnitude of temporal shifts necessary for this alignment should be a good measure for timing fluctuations in the analyzed trials.

Reviewing the literature on adequate approaches to superimpose kinematic profiles one can find the Procrustes analysis and dynamical time warping (DTW) algorithms as well-established methods. The Procrustes analysis is a method for statistical shape analysis (Dryden & Mardia, 2016) that allows different linear transformations (translation, reflection, orthogonal rotation, and scaling) to superimpose different data sets. The method has extensively been used in analyzing kinematic data (for recent examples, see Jäger & Schöllhorn, 2012; Munzert, Maurer, & Reiser, 2014). However, the Procrustes analysis does not offer the possibility to superimpose parameter curves by shifting them in the time domain. At least two-dimensional data sets (e.g., horizontal and vertical position data) are required, in order to transform them to minimize the sum of squared distances between corresponding points. But it is not possible to use timing information as one component and another kinematic feature (e.g., angular position) as a second component, since there is no reasonable metric to determine a sum of squared distances in this non-Euclidian space. Furthermore, the Procrustes transformations always operate on all data components and do not allow only translating time information, as it would be necessary to superimpose the profiles by shifting them in the time domain. In contrast, algorithms based on the dynamical time warping approach (e.g., Ramsay & Li, 1998; Wang & Gasser, 1997) align parameter curves by transforming the temporal structure of time series measures and not the signal itself (e.g., position data). To do so, the time structures of different sequences are changed nonlinearly by a flexible warping function. Thus, the signal is not linearly shifted (i.e., translated) in the time domain and absolute timing information, a necessary raw value to calculate timing variance, gets lost. Hence, DTW also seems not to be an adequate method to implement the suggested approach to study timing variability.

We present a method that analyzes timing precision on the basis of the alignment of different trials relative to their shape

by shifting the profiles in the time domain as opposed to examining the temporal difference between release and a single reference landmark. The method has been used in previously published experiments in which differences in timing of throwing releases in patients with Parkinson's disease (Pendt, Reuter, & Müller, 2011) were assessed to investigate neural correlates of timing imprecision (Maurer, Sammer, Bischoff, Maurer, & Müller, 2014).

In the following, we refer to the term “landmark approach” to define timing analyses that use the temporal difference between release and a single landmark in the movement as timing measure as opposed to approaches that align larger parts of the kinematic movement profiles. In addition, when we speak of alignment, we mean that several kinematic profiles are superimposed as closely as possible in order to assess release timing differences. In contrast to this shape alignment, the profiles can also be temporally aligned, which we call “synchronization.” For instance, during data collection, kinematic throwing profiles of various trials are usually synchronized with respect to release. This means that possible timing deviations in release are eliminated inevitably in the latter case.

Assuming that each single throwing movement is one instance of an acquired movement template, a considerable similarity between the kinematic throwing profiles can be expected. Due to limitations of temporal control, the location of the temporal moment of release varies from trial to trial, generating imprecision of release timing. By aligning the kinematic profiles via shifting them in the time domain, release points shift apart, and the release time deviations between trials can be quantified. That is, the time shifts that are needed to superimpose the profiles as closely as possible represent a measure of timing deviations; and the standard deviation of these deviations represents timing variability. We present a numerical procedure, which we call *timeshift* analysis. It incrementally aligns the kinematic profiles of different throws by shifting them in the time domain. As a consequence, deviations in release timing between throws can be observed as the different release points temporally shift apart. These time shifts finally represent release timing in milliseconds and the timing of single trials can be expressed relative to all other trials. Concretely, the time shift will be defined as positive for trials with a delayed release as compared to all other trials, negative for trials with an earlier release, and zero when the release matches the average release time. Thus, the standard deviation of these time-shift values is representative of release timing variability. It stands for the random (uncontrollable) timing error—that is, temporal imprecision.

The application of the *timeshift* analysis is not limited to certain kinematic profiles. Other signals that evolve over time can be analyzed with respect to timing in the same way. To express this generality, the protocol steps are firstly phrased in more abstract terms and then illustrated by the example of timing analysis in a goal-oriented throwing task. The results

of the suggested way of quantifying temporal variability will be compared to the well-established landmark approaches. Furthermore, we will present some tests of the used optimization algorithm to superimpose the kinematic profiles as well as possible alternative algorithms.

Method

Participants

We present data of the *timeshift* analysis of 14 participants (nine females, five males) with a mean age of 22.4 years (± 2.5) and no specific experience in goal-oriented throwing. All were self-declared right-handers with normal or corrected-to-normal vision. All participants were informed about the purpose of the study and gave written informed consent. The experimental procedures have been approved by the ethics review board of the Justus Liebig University Giessen.

Task

The task we used is a semivirtual variant of a goal-oriented throwing task based on a British pub game in which the player throws a ball hanging from a post, with the goal of hitting a target skittle on the other side of the post (Fig. 1a). The task was displayed on a projection surface, and to perform it, participants manipulated a metal lever (the manipulandum) on which they rested their forearm. Participants sat with their front to the projection surface, on which the ball, the post, the target, and a virtual version of the lever were displayed in two dimensions from a top-down view (Fig. 1b). The post in the center of the projection surface was represented by a cone of 25-cm radius at position $x = 0.0$ cm, $y = 0.0$ cm. The target was a circle of 5-cm radius displayed 35 cm to the right and 100 cm above the center of the post. The virtual lever was represented as a solid bar of 40 cm length, fixed at one end and placed 150 cm below the post. The virtual ball was attached to the free end of the lever as soon as participants touched with their index finger a contact sensor at the tip of the manipulandum (Fig. 1b). To throw the ball, participants moved the manipulandum, which pivoted around an axle centered at the elbow joint, in an outward horizontal direction. An extension of the index finger opened the contact sensor and, in consequence, triggered the ball release. After release, the ball traversed a path initialized by the angle and velocity of the arm (i.e., manipulandum) at the instant of release. This trajectory was determined by the simulated physics of the task and described an elliptic path around the post (Müller & Sternad, 2004). The minimum distance (d in Fig. 1b) between the ball trajectory and the target center was representative of the participant's performance. This was converted to a score that decreased linearly from 100 points, for hits with zero distance ($d = 0.0$ cm), to 0 points, for distances larger than

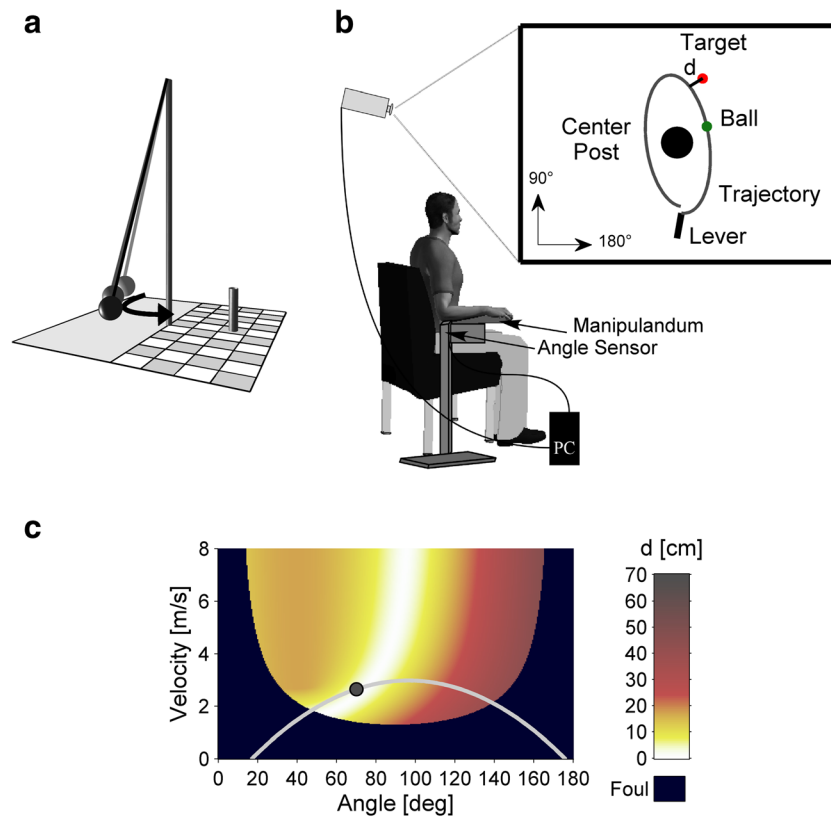


Fig. 1 Experimental task. (a) Sketch of the real skittles task. A ball is suspended on a string and swings around the center post, with the objective of knocking down the skittle on the opposite side. (b) Experimental setup. Participants operated a lever to throw the virtual ball on the screen in front of them with the goal of hitting the target located behind the center post. The angular displacement of the participant's forearm was measured by a magnetic angle sensor and recorded by the computer. (c) Execution and result space of the skittles task. For each combination of the execution variables release angle and

velocity, the color codes the result variable: the minimal distance (d) of the resulting ball trajectory from the target (error). White denotes the solution manifold with zero-error solutions. As an illustration, the angle-velocity trajectory of one throw is plotted, with its corresponding release point. From "Timing Accuracy in Self-Timed Movements Related to Neural Indicators of Movement Initiation," by L. K. Maurer, G. Sammer, M. Bischoff, H. Maurer, and H. Müller, 2014, *Human Movement Science*, 37, Fig. 3. Copyright 2014 by Elsevier B.V. Adapted with permission

50 cm (including post hits). The movements of the virtual lever and the ball trajectory were displayed in real time using custom-made software written in C++. The angle data were captured by a magnetic angle sensor with a sampling rate of 1000 Hz and a 12-bit resolution.

The instruction was to throw the ball in a counterclockwise direction around the post, aiming to hit the target. To do so, participants moved the manipulum in a clockwise direction similar to performing a Frisbee backhand. On the first day, participants tested the task for 10–15 trials. Afterward, 200 trials had to be performed per day for five consecutive days. A participant's performance, as represented by the score, was displayed on the projection surface every ten trials. Participants were motivated to keep their score as high as possible.

Requirements for the *timeshift* analysis

As we stated above, the alignment of parameter curves to assess timing variability is based on the assumption of

relatively uniform time-dependent profiles, meaning that timing variability and not shape variability is primarily responsible for release variability between trials. Beyond that, the kinematic profiles in our virtual throwing task had to be uniform even in an absolute sense, because the release parameters (release angle and velocity) were directly connected to the result. In other cases there might be a stronger focus on relative similarities, and amplitude normalization could then be applied prior to the *timeshift* analysis. In any case, the assumption of absolute or relative similarities should be tested. If the parameter curves vary too much between the trials that are to be compared, the analysis can lead to unrealistic results. To put this similarity assumption to a proof, the variability of the parameter curves within subjects can be compared with the variability between subjects by using, for example, the root-mean square error (RMSE). The RMSE within subjects should be smaller than the RMSE between subjects because subjects may vary interindividually in their profiles, but intra-individually they should be relatively stable.

Step-by-step protocol of the *timeshift* analysis

Within the data collection procedure, parameter curves can be synchronized differently. For instance, the kinematic profiles of the throwing task used here were synchronized at the point of release ($t = 0$) (Fig. 2a) because release was clearly defined. To assess timing differences between the release points of different trials, the *timeshift* algorithm shifts the parameter curves in a defined time window around the event of interest (e.g., release) in time, with decremental steps starting from shift value sh_{start} (e.g., 10 ms), to reduce the sum of all pairwise RMSE values. In the following, the quantification procedure is explained via the example of three sample trials ($n = 3$) of our throwing task. The algorithm was programmed in MATLAB (The MathWorks, Inc.).

Step 1: “Define parameter curves” Choose parameters that represent the curves and define a time window around the event (e.g., release) whose timing is to be analyzed. The time window should satisfy the described requirements for the *timeshift* analysis. It should be large enough to represent the parameter curve around the timing event appropriately, where the curve is the most stable over

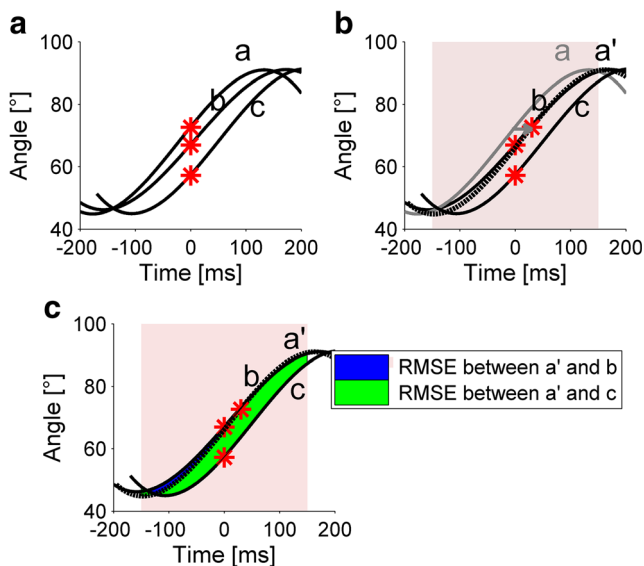


Fig. 2 Schematic illustration of the *timeshift* analysis. (a) The angle profiles of three trials (a , b , c) are plotted with their release times (red stars) synchronized at time $t = 0$. (b) The profile alignment procedure shifts profile a by 30 ms to profile a' . For reasons of better illustration, a large shift value is used; in the procedure, the first shift value is 10 ms. The shift and the following RMSE calculation are done for 300 ms (± 150 ms) around release. (c) Illustration of the RMSEs between trials. Only the areas between a' and b and a' and c for RMSE determination are displayed here. The RMSE between b and c is also calculated, and all three RMSEs determine the total $RMSE_{\text{tot}}$. From “Timing Accuracy in Self-Timed Movements Related to Neural Indicators of Movement Initiation,” by L. K. Maurer, G. Sammer, M. Bischoff, H. Maurer, and H. Müller, 2014, *Human Movement Science*, 37, Supplemental Fig. S1. Copyright 2014 by Elsevier B.V. Adapted with permission

several trials. It should be pointed out that (at least in throwing movements) the trajectory after the temporal event of interest (e.g., release) is of relevance to representing the movement path properly. Since there is variation in timing, a good movement strategy is to execute kinematic profiles that will lead to success even if the intended time of release is missed—that is, using an equifinal movement trajectory (e.g., Müller & Loosch, 1999).

In our case, kinematic profiles were represented by angle-time profiles of the throwing movement for a window of 300 ms (± 150 ms) around release. However, the exact size was not particularly relevant, because similar results can be achieved with different window sizes (see the section below on the robustness of the algorithm).

To minimize computational effort, we determined the coefficients of third-order polynomials of these angle-time profiles (average $R^2 = .99$; e.g., by using the *polyfit* function in MATLAB), and shifting (Step 2) was done by calculating new polynomial coefficients p_1' , p_2' , p_3' , and p_4' by algebraic conversion, such that

$$p_1 + p_2 * (t + sh) + p_3 * (t + sh)^2 + p_4 * (t + sh)^3 = p_1' + p_2' * t + p_3' * t^2 + p_4' * t^3, \quad (1)$$

with sh representing the shift value.

Step 2: “Shift signal and calculate pairwise distances” Shift one parameter curve by the current time-shift value sh_i and calculate the total root mean square error ($RMSE_{\text{tot}}$) between all parameter curves. Discard the shift if $RMSE_{\text{tot}}$ does not reduce. Otherwise, repeat shifting of the current signal followed by calculating $RMSE_{\text{tot}}$ until $RMSE_{\text{tot}}$ does not reduce further. This step is illustrated for the throwing example in Fig. 2b (shifting a to a' by a total time shift of 30 ms) and 2c (calculating the RMSE values).

For the three sample trials, $RMSE_{\text{tot}}$ is computed as follows:

$$RMSE_{\text{tot}} = RMSE_{a'-b} + RMSE_{a'-c} + RMSE_{b-c}, \quad (2)$$

where $RMSE_{a'-b}$ equals the RMSE between trials a' and b , $RMSE_{a'-c}$ is the RMSE between a' and c , and $RMSE_{b-c}$ represents the RMSE between b and c .

When using polynomial fits, time shifts are accomplished by replacing the old with the new coefficients. The RMSE values between two profiles $g(t)$ and $h(t)$ can be calculated by integration of the polynomial $f(t) = [g(t) - h(t)]^2$. In cases in which profiles cannot be fitted by an integrable function, the shifting as well as the RMSE determination has to be carried out index-wise.

Step 3: “Cycle trials” Repeat the previous step for all other trials. That is, in the three-trial example, trial b and after that trial c are shifted by shift value sh_i until $RMSE_{tot}$ does not reduce further.

Step 4: “Invert shift value” Change the direction (positive vs. negative, or vice versa) of the shift value

$$sh_i = -sh_i. \quad (3)$$

Step 5: Repeat Step 2 to Step 4 n times That is, repeat the shifting procedure for all trials with the current shift value and alternating sign. We used $n = 15$ in our throwing example.

Step 6: “Halve step” Halve the shift value size:

$$sh_{i+1} = 0.5 * sh_i. \quad (4)$$

Step 7: Repeat Step 2 to Step 6 Repeat these steps until the step size is less than a defined limit—for example, 0.1 ms. The chosen value for the step size limit is a trade-off between accuracy and computational effort. The limit should depend on the expected variability in the analyzed measure and should be clearly below that variability.

This procedure moves the parameter curves backward or forward in time, resulting in a time-shift measure in milliseconds for each trial. The standard deviation of time-shift values represents the timing variability.

Robustness of the algorithm against implementation details

Some additional analyses have been performed to test whether the algorithm tends to converge toward local minima. (1) We checked the assumption that the results should be independent of the order of trials transferred to the shifting algorithm. To do so, we ran the algorithm ten times per block (50 trials each) with permuted trial order. As a result, the absolute time-shift values and the timing variability (standard deviation of the time-shift values for blocks of 50 trials) hardly differed between the orders (the average standard deviations in the single time-shift values and in overall timing variability were 1.8 and 0.07 ms, respectively). (2) For each of the 280 blocks, we generated 100,000 vectors of randomly chosen shift values (uniformly distributed in the interval $[-75, 75]$) and checked whether the $RMSE_{tot}$ between the shifted profiles lay below the minimal RMSE value determined by the *timeshift* algorithm. However, this was never the case, indicating that no global minimum was detected with smaller RMSE values than those arising from our *timeshift* algorithm. (3) The kinematic profile of each trial within blocks was shifted by a random value (uniformly distributed in the interval $[-10, 10]$ ms)

before using the *timeshift* algorithm. Afterward, the random shifts were subtracted from the shift values determined by the algorithm, and these values were compared to those resulting from nonshifted profiles. High similarity between the resulting two sets of time-shift values was found, indicated by a mean difference of 0.09 ms when comparing the single shift values and a high correlation ($r > .99$) of the resulting measures of timing variability (standard deviation of shift values for each block).

Another influencing variable might be the size of the time window around the event of interest (e.g., release) for which the algorithm aligns the movement profiles by minimizing $RMSE_{tot}$. We used a window of ± 150 ms around release. Additionally, we run the *timeshift* analysis with windows of ± 100 and ± 50 ms around release and compared the results with those of window size ± 150 ms. To analyze the similarity of the results we used linear regression analysis for the resulting time-shift values for each block. We found slopes of the regression lines very close to 1 ($M_{\pm 150 \rightarrow \pm 100} = 1.02$, $SD_{\pm 150 \rightarrow \pm 100} = .07$, $M_{\pm 150 \rightarrow \pm 50} = 1.00$, $SD_{\pm 150 \rightarrow \pm 50} = .06$) but varying intercepts ($M_{\pm 150 \rightarrow \pm 100} = -10.05$, $SD_{\pm 150 \rightarrow \pm 100} = 10.24$, $M_{\pm 150 \rightarrow \pm 50} = -17.65$, $SD_{\pm 150 \rightarrow \pm 50} = 16.85$). Hence, the different implementations can cause constant offsets in the absolute time-shift values, but the time-shift values have a high relative similarity. Since we used the standard deviation of the time-shift values as a measure for timing variability, the absolute differences did not affect the final measure of interest. This was also demonstrated by linear regression analysis of the timing variability measures (tvm) determined for each analysis block on the basis of the different *timeshift* implementations ($tvm_{\pm 100} = 0.96 * tv_{\pm 150} + 1.59$, $R^2_{\pm 150 \rightarrow \pm 100} = .98$; $tvm_{\pm 50} = 0.97 * tv_{\pm 150} + 1.08$, $R^2_{\pm 150 \rightarrow \pm 50} = .98$). Hence, the quantification of timing variability using the *timeshift* analysis hardly relies on the exact window size, at least in our throwing example.

Finally, during data-shifting, data gaps can naturally occur between the profiles. These gaps need to be filled by adequate values, or the RMSE values have to be determined just for the overlapping parts. In the first case, gaps can be filled by original values of the profiles, estimated extrapolated values, or fitted interpolated values in the case of fitting the profiles beyond the analysis window. Whether the gaps should be filled or the width of the calculation window should be modified depends on the suitability of the generated values for the gaps. In our case, when a kinematic profile is shifted, the missing data points are added on the basis of an extrapolation via the polynomial fit. We also ran the *timeshift* analysis by fitting over a time window of ± 200 ms, but doing the alignment for a window of ± 50 ms. Linear regression analysis for the time-shift values within each analysis block revealed a high similarity between the two ways of performing the *timeshift* analysis ($M_{slope} = 1.03$, $SD_{slope} = 0.06$; $M_{intercept} = -0.78$, $SD_{intercept} = 5.76$). Doing the regression analysis for the

final measure of timing variability (tvm; standard deviation of time-shift values within each block) demonstrated highly consistent results ($tvm = 0.98 * tv_{\text{standard}} + 1.3121$, $R^2 = .99$).

Alternative algorithms

The presented *timeshift* algorithm to align movement profiles by shifting them in the time domain is an optimization algorithm that minimizes $RMSE_{\text{tot}}$. Hence, the optimization procedure might also be solved by other optimization algorithms. Two alternative algorithms were tested using the MATLAB functions *fminsearch* (“Nelder–Mead simplex direct search”) and *fminunc* (“Quasi-Newton Algorithm” Optimization Toolbox). In a first step, the resulting $RMSE_{\text{tot}}$ values of the three algorithms were compared. Although the *timeshift* analysis and the quasi-Newton algorithm resulted in similar values (the exact values depended on the stopping criteria used), the $RMSE_{\text{tot}}$ values of the Nelder–Mead algorithm were far higher. Visual inspection also demonstrated that the alignment of movement profiles was not successful using this algorithm, and it will not be considered any further. Inspection of the resulting shift values of the *timeshift* and quasi-Newton algorithms revealed highly similar results. Correlating the shift values of the single trials within the blocks resulted in mean correlation values of $r > .99$, and the same held for correlating the resulting measures of timing variability for blocks of 50 trials ($r > .99$). Finally, we analyzed the computational effort in terms of calculation time for aligning blocks of 50 trials, revealing slightly shorter times for the quasi-Newton algorithm (on average, 86% of the time required by the *timeshift* algorithm).

Validity check

To evaluate the validity of the *timeshift* analysis, we compared its timing results in the virtual throwing task with the results received when applying a timing analysis based on a landmark approach on the same data. Several landmarks may be considered with regard to the angle profiles in the throwing task. With reference to the options listed in the introduction, three landmarks seem most appropriate for our data: (1) movement initiation/onset (Onset), (2) maximal movement velocity (VeloMax), and (3) a specific fixed angle position (FixAngle). Movement onset was defined by the last point in time prior to the velocity maximum at which velocity was closest to zero. As the fixed angle we chose 70° , because this angle was typically passed by all participants and was on average relatively close to release (the effective range of manipulandum movement was between 0° and 180° ; cf. Fig. 1c). To assure that the results of the FixAngle analysis were not biased by the choice of angle, we also ran the analysis over a range of other reasonable reference angles (from 60° to 100°). The results did not differ significantly between different reference angles.

Release timing was calculated as the timing difference between the corresponding landmark and the release time.

As to the criteria of comparison, we first checked that the prerequisites of the approaches were not violated. In the case of the *timeshift* approach, the similarity requirement of the kinematic profiles had to be satisfied. This was tested using the RMSE between the angle profiles of the throws. For the landmark approaches, temporal stability was required between landmark and release, which we analyzed via the development of the absolute timing values over practice. If the temporal relation between the two events is stable, the absolute timing values should also remain stable in the course of practice. Second, we examined the development of timing variability over practice in all approaches, because an improvement in timing precision is expected with learning. Finally, we analyzed the amounts of timing variability (as quantified by the standard deviation of release timing) revealed by the three methods. We assumed that the measured value of release timing variability is a function of variance in release and of variance in the reference point (to which timing is determined), since the reference itself is also prone to fluctuations in time. Hence, the reference should be chosen to have minimal variance, because then the resulting timing variance would be ascribed primarily to the variance in release. With respect to our comparison between approaches, this means that the approach with the lowest measured release timing variability would have the lowest variance in reference, and thus would be the best representative of the true timing variability.

Results

This section, first, illustrates the results of the similarity requirement analysis of the kinematic throwing profiles as well as the stability requirement between landmark and release. Subsequently, we present the results of the *timeshift* analysis of two exemplary participants and 14 participants who practiced the virtual throwing task over five consecutive days with 200 trials each day. Finally, the timing development over practice as well as the timing variabilities of the *timeshift* analysis and the three landmark approaches are presented and compared.

Results of the prerequisite analyses

Similarity requirement of the *timeshift* approach To quantify the similarity between kinematic throwing profiles, we determined the variances between the angle-time profiles of trials within subjects over the five-day practice period. The measure for variance was the RMSE between angle profiles, which was calculated block-wise for blocks of 50 trials. RMSE reduced continuously over all practice blocks. The largest reduction, however, occurred on average over the first

150 trials. These first trials were interpreted as being representative of the exploration phase. {This was tested with exponential fits. $RMSE_{block}$ data were fitted with a function of the form $a - \exp[-b * (x + c)]$. The point at which the slope of the function fell below the regular dispersion of the RMSE (calculated as the RMSE between the actual RMSE values and the fit) values was defined as the end of the exploration phase.} After the exploration phase, the RMSE between trials within subjects was about 47% lower than a compilation of angle profiles randomly drawn from all participants. Hence, participants varied interindividually in their throwing manner, but intra-individually they were relatively stable.

Stability requirement of the landmark approaches With respect to the temporal stability between the landmarks and the release, Fig. 3 shows that the timing between all three landmarks and release varied over the 800 trials. Concretely, the temporal distance to release of the landmarks Onset and FixAngle decreased constantly with ongoing practice (Fig. 3, top and bottom panels). With respect to the stability requirement, one can argue that the stable relation only settles in the course of practice, and that the asymptotic profile in the FixAngle approach is in agreement with the stability requirement. In contrast, the temporal relation between VeloMax and

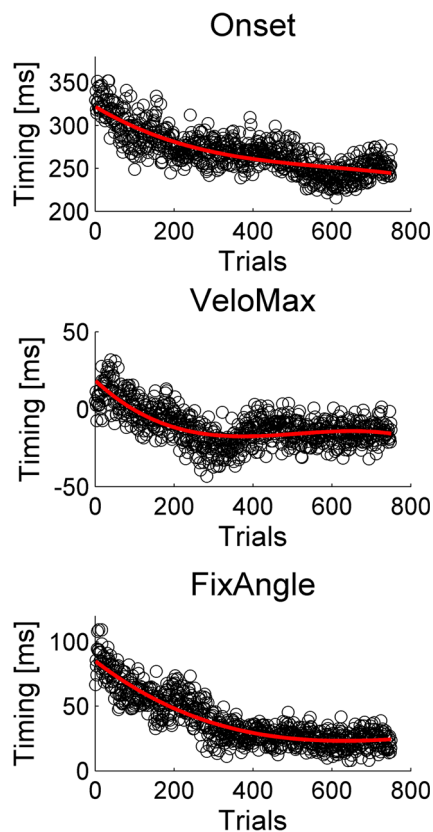


Fig. 3 Development of absolute timing over trials in the three landmark approaches. Depicted are the trial averages of timing in relation to the landmarks over all 14 participants (black circles) and third-order polynomial fits over the averages (red lines)

release almost features a U-shaped profile (Fig. 3, middle panel)—that is, participants release before the velocity maximum (>0) at the beginning of practice, then change to releasing after the velocity maximum (<0), and finally adapt to releasing relatively close to the velocity maximum. Here, we would not speak of a stable coupling between VeloMax and release. In conclusion, the stability requirement of the landmark approaches was only partially fulfilled.

Results of the *timeshift* analysis

The trials of each participant were blocked in groups of 50 trials, and the *timeshift* analysis was conducted block-wise. Figure 4a depicts 50 trials for Participant 1, synchronized with respect to release (as the data were recorded). Figure 4a' shows the result after the *timeshift* analysis for the same participant. The same trials are now aligned with respect to their angle profiles. Figures 4b and 4b' display the procedure on 50 trials for Participant 2. It can easily be seen that the two participants chose different throwing strategies with respect to the angle-time profiles. However, the profiles within the participants are relatively similar. Since the *timeshift* analysis produces relative values, timing variability can be compared between participants with different movement kinematics as long as the within-subjects similarity is given. We chose to block the trials to illustrate the development of timing variability with practice. If, on the other hand, the research interest lies in the total variability of timing, irrespective of temporal changes, it is reasonable to run the *timeshift* analysis over the full data set. When doing this, however, we strongly recommend also testing the similarity requirement over the full data set, since movement trajectories might change over time, and an analysis over the full range of different movements can lead to unreliable results (as Fig. 5 illustrates). Our participants' release timing variability as a function of practice blocks is displayed in Fig. 6a. Variability was determined as the standard deviation of the time-shift values per block. Timing variability gradually decreased from 39 ms (± 10.2) to 21 ms (± 7.3) on average.

Figure 5b and b' show an example in which the angle profiles (i.e., the throwing movements) differ extremely. In consequence, the alignment procedure of the *timeshift* analysis does not lead to satisfying results, and the time-shift values have no explanatory power. However, even in less severe cases (Fig. 5a and a'), the results of the *timeshift* analysis have to be considered with care, because differences between angle profiles might be more strongly connected to differences in throwing strategies than to differences in release timing.

Comparison to landmark approaches

Figure 6 comparatively depicts the results of the different landmark timing analyses in reference to the *timeshift*

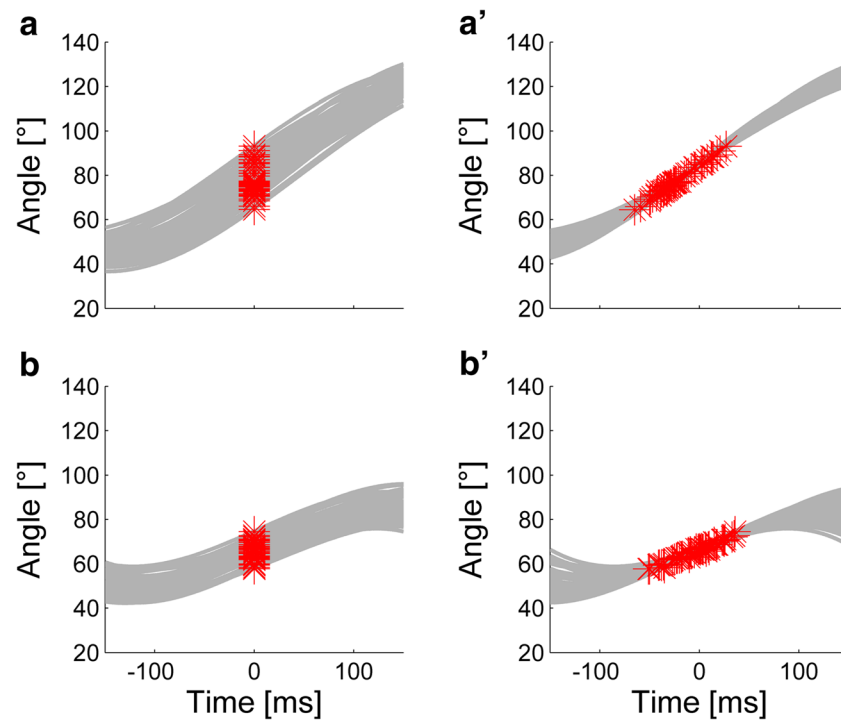


Fig. 4 *Timeshift* procedure with profile alignment for Participants 1 and 2. **(a)** Depicted in gray are the angle profiles of 50 trials for Participant 1, with their corresponding release points (red stars). The release is synchronized at time $t = 0$, and the profiles are plotted for a window of 300 ms (± 150 ms) around release. The *timeshift* analysis shifts the angle

profiles in time to reduce the total RMSE between trials, and thereby aligns them. **(a')** Time-shifting results for Participant 1. The release points of the different trials temporally shift apart, resulting in a time-shift (i.e., timing) value for each trial. **(b and b')** These panels show the same results for Participant 2

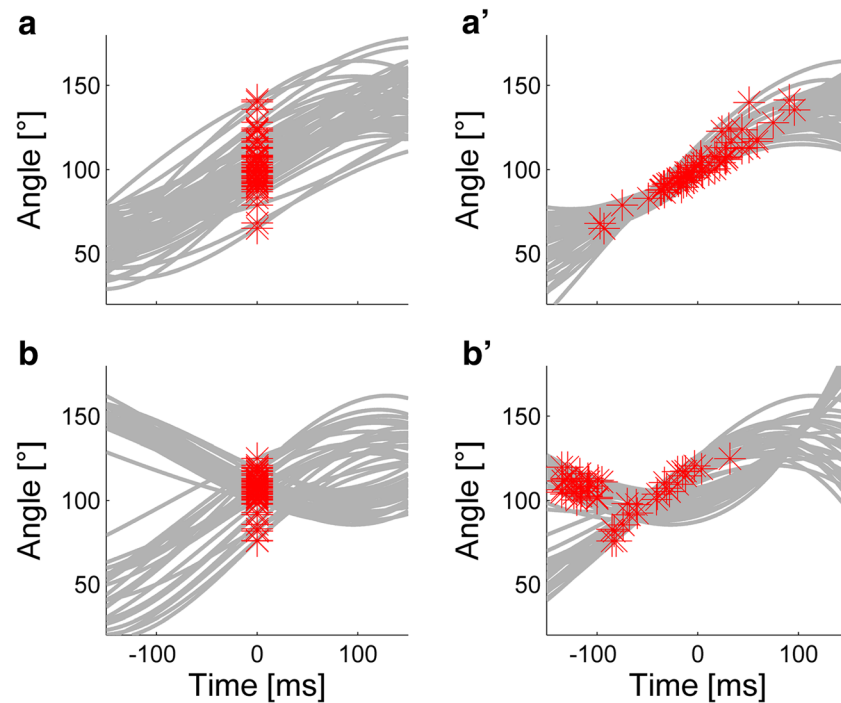


Fig. 5 Illustration of the limits of the *timeshift* analysis. **(a)** Angle profiles showing differences indicating variant throwing strategies. **(a')** Profile alignment cannot minimize the differences between trials satisfyingly. Thus, the time-shift values have to be considered with care. **(b)** Angle

profiles showing extreme characteristics that are not due to release timing differences but that arise from variant throwing strategies. **(b')** Profile alignment cannot minimize the differences between trials satisfyingly. Thus, time-shift values have no explanatory power in this case

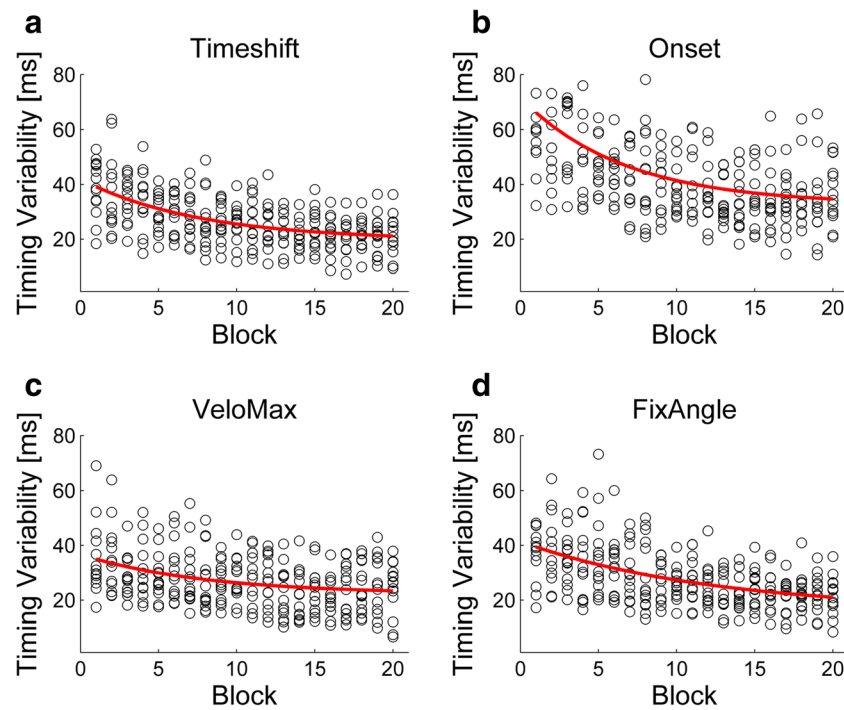


Fig. 6 Development of timing variability with practice for the four different methods. Timing variability is expressed as the standard deviation per block (50 trials each) of the time-shift values (**a**), the timing between release and Onset (**b**), the timing between release and

VeloMax (**c**), and the timing between release and FixAngle (**d**). Black circles represent the values for single participants. The red lines mark exponential fits of the averages over all 14 participants

analysis. Shown is the development of the timing variability (as determined by the corresponding approaches) averaged over blocks of 50 trials. The VeloMax approach (Fig. 6c) and the FixAngle approach (Fig. 6d) revealed results relatively similar to those from the *timeshift* analysis (Fig. 6a), with respect to starting level and development over practice, respectively. The timing variability yielded by the Onset approach (Fig. 6b) also reduced over time. With respect to the variability level, however, the approach produced constantly higher timing variability than the other three methods.

Discussion

The analysis of timing in nonreactive human movements like the release in throwing requires a timing reference to which timing can be quantified. The choice of the analysis approach depends on the task, the research question, and on the demands of the analysis. The most common method is to determine the variability of the temporal moment of release relative to a single event in the movement (a specific landmark). Prerequisite of this approach is the existence of a suitable landmark and the prevalence of the assumption that the coupling between release and the landmark is stable and variability between throws arises solely from variations between release and landmark. In cases in which these requirements are not valid, other approaches can be used that align larger parts of different

kinematic movement profiles in order to assess release timing relative to the whole movement instead of just a single landmark of it. Different approaches in this category exist (e.g., dynamical time warping, Procrustes analyses, *timeshift* analysis), which have been reviewed in the introduction. Which of these should be chosen depends on, for instance, whether it is necessary that relative timing remains intact or whether amplitudes of the signal can be changed. Regarding to our task and research question, the *timeshift* method presented here is appropriate. The *timeshift* method aligns different parameter curves using a range of the kinematic profile located around a single event (e.g., release in throwing movement) by shifting them in the time domain in order to access timing of that event. Importantly, single kinematic shapes are not distorted to be aligned such that absolute timing information remains intact.

We have applied the *timeshift* analysis to release timing in throwing; aligning was done with angle-time profiles of the throwing trajectories. As we have explained, it is critical for the method that the parameter curves to be aligned fulfill a similarity requirement. Only when single curves resemble each other, a realistic alignment can be accomplished. If this is the case, timing variability with respect to a specific event in the parameter curve (here, the release) can be determined in relation to the whole movement as opposed to a single landmark for example. In the present experiment, the alignment of the kinematic profiles allowed to quantify the release time deviations between trials. Concretely, the timing deviations

between trials were expressed by the time shifts that were needed to align the trials. These time shifts can be positive or negative depending on the position of the release on the kinematic profile. The time shift is positive when the release is delayed, and negative when the release is early. The standard deviation of the time shifts then represents the timing variability. The procedure can be applied to compare timing between individual trials of one set (Maurer et al., 2014) as well as between different sets of trials (Maurer et al., 2014; Pendt, Maurer, & Müller, 2012). When more than one set of trials of different conditions (e.g., the trials before and after a break) are compared, trials of all sets are passed to the algorithm at once and parameter curves are collectively aligned. After the *timeshift* analysis, the timing information (time shift) for the trials of the different sets needs to be averaged separately and the data compared with each other.

Our results show that the procedure produces plausible results for release timing in throwing. Timing variability of participants who had no specific experience in throwing was on average 39 ms at the beginning of practice. After about 1,000 throws the average variability decreased to 21 ms. The least timing variability in our participants was around 10 ms. These results are in accordance with other studies examining timing precision (i.e., random timing variability) using the landmark approach described in the introduction. They found a release timing variability of 14–58 ms in unskilled participants (Timmann, Citron, Watts, & Hore, 2001), relative to a variability range of 6–15 ms in recreational baseball/softball or cricket players (Hore et al., 1995; Hore, Watts, Tweed, & Miller, 1996). However, other studies measured significantly lower release timing variabilities (1–5 ms, Nasu et al., 2014; 1.8–7 ms, Smeets et al., 2002). These studies used other variability measures (i.e., accuracy; Nasu et al., 2014) and/or had other methodological differences with respect to the determination of timing (e.g., differences in the underlying modeling of the throwing movement; Smeets et al., 2002).

But there are other indicators that we can have confidence in the data presented here. One is the fact that participants decreased their timing variability by about almost 50% on average after 1,000 trials. Improving timing precision practically means noise reduction, and although noise is a random factor intrinsic to the human nervous system, it can be reduced with practice. Comparison of the unskilled and skilled participants in the above-mentioned studies demonstrates this: Participants with no experience in goal-oriented throwing show a release timing variability that is more than 50% higher than that of recreational baseball/softball or cricket players who have practiced several thousands of trials.

In addition to the *timeshift* analysis, we analyzed the timing variability of the same throwing data with different landmark approaches and then compared the approaches. The first landmark, the movement onset, produced much higher timing variability values than the other landmark approaches and than

the *timeshift* approach itself. This was possibly caused by the large temporal distance between movement onset and release (on average 260 ms), which increased the variability between the two events (movement onset and release) and thereby the timing variability. Thus, the movement onset does not seem to be a valid landmark for such movements. The timing variability results of the 70°-angle landmark were more similar to the results of the *timeshift* analysis as well as to the results of the other previously cited studies. However, the choice of the angle was arbitrary. Even with other possible reference angles, there is no real, substantive rationale to assume that participants try, of all things, to stabilize the distance between release and a certain angle during fast ballistic throwing. Indeed, the absolute timing results confirm that they do not; the temporal relation between release and reaching the reference angle constantly decreased over practice. Moreover, different participants used very different throwing strategies, as can be observed in Fig. 4. Hence, the distance between a certain angle position and release also varied. A timing analysis that is independent of the choice of a single landmark, like the *timeshift* approach, is clearly more advantageous in this regard.

The third landmark option, the velocity maximum, again yielded timing variability results comparable to those of the *timeshift* analysis. In addition, the distance between the velocity maximum and the release is relatively short. Participants typically released close to the velocity maximum. However, the timing between release and maximal velocity was not stable, as would be required. Concretely, participants released the ball at the beginning after the velocity maximum, then went over to releasing prior to the maximum, and ended at releasing the ball relatively close to the velocity maximum. Since this development was observed over tens of trials, this influence on timing variability can be compensated for by analyzing timing variability block-wise over several trials, as we have exemplarily done here. In the case that timing variability should be analyzed over larger trial numbers, the *timeshift* analysis is preferable, since it does not require the assumption of a stable timing between release and a single landmark.

Conclusion and further applications

We have presented the *timeshift* analysis as an adequate approach to quantifying timing variability in human movements by aligning parameter curves and shifting them in the time domain, which, as compared to the single-landmark approach, does not require a stable temporal relation of the event of interest to another single event. Importantly, it is a prerequisite of the *timeshift* method that the different parameter curves themselves are relatively stable. That is, they should not vary too much in shape, such that alignment is not possible. If this requirement is fulfilled, the *timeshift* analysis can be applied to different kinds of signals evolving over time—for example, electroencephalographic (EEG) signals. When analyzing

event-related potentials, for instance, the coupling of a time-varying event to a distinct landmark in the signal is even more challenging than in kinematic signals, and hence, a coupling to the whole signal might be preferable. In contrast, the similarity requirement might be less well satisfied because of greater amounts of noise in the EEG signals. But, in cases in which the absolute amplitude of the single EEG signals is not significant to the research question, amplitude normalization can be applied, mitigating irrelevant differences between the signals and thus improving the applicability of the *timeshift* analysis. We have previously used the *timeshift* analysis to compare the EEG signals that underlie different throwing movements, with plausible results (Maurer et al., 2014), and other applications are conceivable.

Author note This research was supported by the Deutsche Forschungsgemeinschaft (Grants MU 1374/3-1 and SFB/TRR 135 to H. Müller). The data are accessible via doi:10.5281/zenodo.583422.

References

- Becker, W. J., Kunesch, E., & Freund, H.-J. (1990). Coordination of a multi-joint movement in normal humans and in patients with cerebellar dysfunction. *Canadian Journal of Neurological Sciences*, *17*(03), 264–274. doi:10.1017/S0317167100030560
- Bureau International des Poids et Mesures (BIPM). (2012). *Vocabulaire international de métrologie: Concepts fondamentaux et généraux et termes associés (VIM) [International vocabulary of metrology: Basic and general concepts and associated terms (VIM)]* (3rd ed., 2008 version with minor corrections). Paris, France: BIPM.
- Chowdhary, A., & Challis, J. (1999). Timing accuracy in human throwing. *Journal of Theoretical Biology*, *201*, 219–229. doi:10.1006/jtbi.1999.1024
- Dryden, I. L., & Mardia, K. V. (2016). *Statistical shape analysis with applications in R* (2nd ed., Wiley Series in Probability and Statistics). Chichester, UK: John Wiley & Sons.
- Hore, J., Timmann, D., & Watts, S. (2002). Disorders in timing and force of finger opening in overarm throws made by cerebellar subjects. *Annals of the New York Academy of Sciences*, *978*, 1–15. doi:10.1111/j.1749-6632.2002.tb07551.x
- Hore, J., Watts, S., Martin, J., & Miller, B. (1995). Timing of finger opening and ball release in fast and accurate overarm throws. *Experimental Brain Research*, *103*, 277–286. doi:10.1007/BF00231714
- Hore, J., Watts, S., Tweed, D., & Miller, B. (1996). Overarm throws with the nondominant arm: Kinematics of accuracy. *Journal of Neurophysiology*, *76*, 3693–3704.
- Jäger, J. M., & Schöllhorn, W. I. (2012). Identifying individuality and variability in team tactics by means of statistical shape analysis and multilayer perceptrons. *Human Movement Science*, *31*, 303–317. doi:10.1016/j.humov.2010.09.005
- Maurer, L. K., Sammer, G., Bischoff, M., Maurer, H., & Müller, H. (2014). Timing accuracy in self-timed movements related to neural indicators of movement initiation. *Human Movement Science*, *37*, 42–57. doi:10.1016/j.humov.2014.06.005
- McNaughton, S., Timmann, D., Watts, S., & Hore, J. (2004). Overarm throwing speed in cerebellar subjects: Effect of timing of ball release. *Experimental Brain Research*, *154*, 470–478. doi:10.1007/s00221-003-1677-0
- Müller, H., & Loosch, E. (1999). Functional variability and an equifinal path of movement during targeted throwing. *Journal of Human Movement Studies*, *36*, 103–126.
- Müller, H., & Sternad, D. (2004). Decomposition of variability in the execution of goal-oriented tasks: Three components of skill improvement. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 212–233. doi:10.1037/0096-1523.30.1.212
- Munzert, J., Maurer, H., & Reiser, M. (2014). Verbal–motor attention-focusing instructions influence kinematics and performance on a golf-putting task. *Journal of Motor Behavior*, *46*, 309–318. doi:10.1080/00222895.2014.912197
- Nasu, D., Matsuo, T., & Kadota, K. (2014). Two types of motor strategy for accurate dart throwing. *PLoS ONE*, *9*, e88536. doi:10.1371/journal.pone.0088536
- Pendt, L. K., Maurer, H., & Müller, H. (2012). The influence of movement initiation deficits on the quantification of retention in Parkinson's disease. *Frontiers in Human Neuroscience*, *6*, 226. doi:10.3389/fnhum.2012.00226
- Pendt, L. K., Reuter, I., & Müller, H. (2011). Motor skill learning, retention, and control deficits in Parkinson's disease. *PLoS ONE*, *6*, e21669. doi:10.1371/journal.pone.0021669
- Ramsay, J. O., & Li, X. (1998). Curve registration. *Journal of the Royal Statistical Society: Series B*, *60*, 351–363. doi:10.1111/1467-9868.00129
- Smeets, J., Frens, M., & Brenner, E. (2002). Throwing darts: Timing is not the limiting factor. *Experimental Brain Research*, *144*, 268–274. doi:10.1007/s00221-002-1072-2
- Timmann, D., Citron, R., Watts, S., & Hore, J. (2001). Increased variability in finger position occurs throughout overarm throws made by cerebellar and unskilled subjects. *Journal of Neurophysiology*, *86*, 2690–2702.
- Wang, K., & Gasser, T. (1997). Alignment of curves by dynamic time warping. *Annals of Statistics*, *25*, 1251–1276. doi:10.1214/aos/1069362747