Journal of Experimental Psychology:
Human Perception and Performance
2004, Vol. 30, No. 1, 212–233

0096-1523/04/$12.00 DOI: 10.1037/0096-1523.30.1.212

Decomposition of Variability in the Execution of Goal-Oriented Tasks: Three Components of Skill Improvement

Hermann Müller
University of the Saarland

Dagmar Sternad
Pennsylvania State University

A central ability of the motor system is to achieve goals with great reliability, although never with zero variability. It is argued that variability is reduced with practice by 3 separate means: reduction of stochastic noise (*N*), exploitation of task tolerance (*T*), and covariation (*C*) between central variables. A method is presented that decomposes variability into these components in relation to task space that is defined by the execution variables. Successful variable combinations form the solution manifold. In a virtual skittles task, it is demonstrated that participants' improvement over repetitions, indicated by increasing accuracy, is accounted for by *N*, *T*, and, to a lesser degree, *C*. The relative contribution of these components changes over the course of practice and task variations.

In many everyday activities and sports skills, actors engage in movements in which they try over successive trials to achieve a desired outcome. Examples in sports are throwing a dart to hit the bull's-eye, bowling a ball to hit the skittles, or running up to hit the take-off board for a long jump. Similar goal-oriented skills are found in everyday life when one aims to hit a key on the keyboard or guide a fork to one's mouth. Historically, hammering on an anvil was the first of such skills to be investigated, and it has become almost legendary through Bernstein's (1935, 1967) studies. It can also be found in even earlier writings by Drill (1933) and Woodworth (1899; see also Newell & Vaillancourt, 2001). As Bernstein demonstrated, the achieved accuracy in the result—hitting the same spot on the anvil—contrasts with the observation that the trajectories of the multijoint arm are virtually always different. Similarly, throwing a dart to the same target position can be achieved with many different release positions and release angles. The performance is successful when the dispersion in the result is within some tolerance band that is given by the size of the target (i.e., the size of the bull's-eye). If the target is relatively small, the success is limited by the variability in the result. One focus of this study is on how relative constancy in the result is achieved even when there is variability in the process of movement execution. A second focus is on how the structure of this variability in execution changes across practice.

A fundamental assumption in our approach is that biological systems are always subject to perturbations, originating from both internal physiological processes and from the environment. Hence, a certain irreducible level of "noise" is a basic expression of biological systems and can even play a functional role. Abnormal levels of variability are a characteristic of dysfunctional behavior, as evidenced in tremor and excessive movements, and also in the lack of variability, known as stereotypy (Newell, van Emmerik, & Sprague, 1993). However, the magnitude of this variability is also a function of the task and expertise. We assume that the acquisition and improvement of a skill is characterized by the exploration and exploitation of this natural variability within a given task.

The assumption that biological systems have a significant stochastic component in their movement system is at the core of several current approaches to motor control. For instance, Harris and Wolpert (1998) proposed a model in which the neural control signals have inherent noise that increases with the magnitude of the control signal. Given a task, the controller selects a trajectory shape that minimizes the variance of the final position. Similarly, Todorov and Jordan (2002) developed a stochastic optimal feedback controller that operates with multiplicative noise in its control signals and with noisy sensors with delay. The controller only corrects deviations that interfere with task goals. Both models demonstrate their feasibility by reproducing selected behavioral results, such as Fitts's law and the ⅔ power law.

Whereas these modeling studies suggest specific control algorithms that explicitly include system noise to replicate variability results, typically presented in traditional statistical measures, such as standard deviations around a given target mean, the goal of the present study was to develop a new analysis of variability that quantifies features observed in empirical data. We suggest a novel method that decomposes variability in the performance of a redundant system into three components. This method differentiates purely stochastic components from goal-dependent components of variability. Further, beyond assuming stochastic processes to be only a nuisance, our method extracts aspects of variability that can be interpreted as having a functionally positive role, such as the exploration of task space. This method can provide new cornerstones for testing control mechanisms that deal with this inherent variability or noise.

Hermann Müller, Institute for Sport Science, University of the Saarland, Saarbrücken, Germany; Dagmar Sternad, Department of Kinesiology, Pennsylvania State University.

This research was supported by Deutsche Forschungsgemeinschaft Habilitation Grants MU 1374/1-1–1-3 to Hermann Müller and by National Science Foundation Behavioral Neuroscience Grant 00-96543 to Dagmar Sternad. We thank our anonymous reviewers for their extremely constructive comments.

Correspondence concerning this article should be addressed to Dagmar Sternad, 266 REC Hall, Department of Kinesiology, Pennsylvania State University, University Park, PA 16803. E-mail: dxs48@psu.edu

A fundamental step for the analysis of variability is the distinction of features that characterize the execution process and features that capture the result of a movement. Let's take dart throwing as an example to illustrate this distinction (see Figure 1). The critical moment that determines the entire dart throw is the moment of dart release. This execution of the throw can be fully described by a 6-D vector **e**, which contains the position coordinates at release (x, y, and z), as well as the release velocities in the three dimensions: $\mathbf{e} = (x, y, z, v_x, v_y, v_z)$. The result can be described in terms of the position on the planar target disk in a 2-D vector, $\mathbf{r} = (r_x, r_y)$, which quantifies the location of the hit either in Cartesian or in polar coordinates. If there is a series of throws with i trials ($i = 1, 2, \ldots, n$), there is a set of execution vectors $\mathbf{E} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ and a set of result vectors $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$.

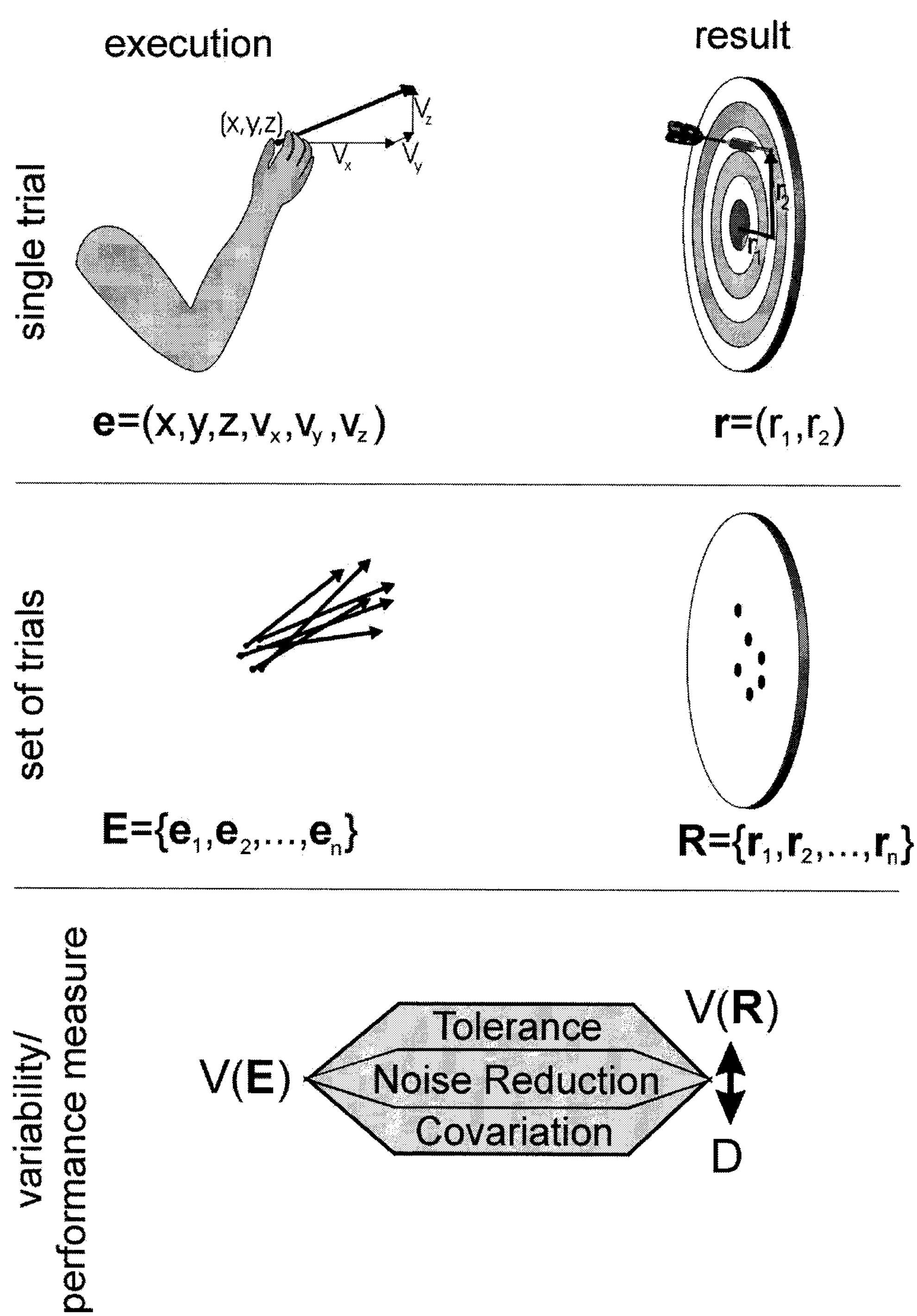

Figure 1. Illustration of the concepts and notation introduced in the text.

Repeated throws are unlikely to be produced in an identical fashion. Therefore, successive trials will have a variability V in both the execution variables and the result variables. The central issue introduced above can now be formulated more succinctly: Which relation do execution variability V(E) and result variability V(R) have? This question has particular relevance in situations in which V(R) is the critical limiting factor for successful performance measured in D (e.g., distance from the target). The problem in answering this question is that the variability of two sets of vectors are compared that can have different numbers of entries and different physical dimensions. In the example of the dart throw, the vector **e** has six entries, whereas the result vector **r** only has two position dimensions.

With these concepts, a few paradoxical observations can now be formulated as follows: The variability in the movement result V(R) can decrease, or even approach zero, although the variability in the execution V(E) remains significantly above zero and constant. Conversely, it is possible for V(E) to decrease without producing a concomitant decrease in V(R). To understand these paradoxical relations, it is necessary to understand that V(E) is not the single determinant of V(R). The present study shows that changes in V(R) can be described as the sum of three components: noise reduction, exploitation of task tolerance, and covariation between the relevant execution variables (Figure 1). We present a method for calculating the quantitative contributions of these components and exemplify this method in an experiment in which participants practiced a goal-oriented task. On the basis of the data, we show that the observed improvement in the accuracy of performance can be broken down into these three components.

First Component: Noise Reduction

The first systematic empirical studies on variability in movement execution were conducted at the end of the 19th century by Fullerton and Cattell (1892) and Woodworth (1899). In his now famous studies on line drawing, Woodworth highlighted the fact that accurate movement amplitudes depended on movement speed and other target criteria. This work was the beginning of a series of studies that led to the formulation of Fitts's law (see, e.g., Crossman & Goodeve, 1983; Fitts, 1954; Fitts & Peterson, 1964; Meyer, Smith, Kornblum, Abrams, & Wright, 1989; Plamondon & Alimi, 1997). Besides the well-known systematic relation between movement time, amplitude, and accuracy, a central insight from these early studies was that a certain degree of variability or noise in execution is inevitable, and this noise can never be entirely eliminated.

Studies that investigated movements over a course of practice sessions, however, observed repeatedly that this variability decreased while the accuracy for the desired task increased. It was not astonishing for Woodworth (1899) that it was the worker with the least practice who performed the only mishit among 4,000 observed hammer movements. Reduction of motor noise as a consequence of practice is a widely acknowledged component in skill improvement. Yet these findings were possibly too obvious to motivate further examinations of the functional processes that underlie this decrease in variability.

When variability was investigated, it was typically only the variability in the result that was studied. In contrast to the ubiquitous presence of this decrease in result variability, relatively few studies have explicitly related variability in the result to the variability in the execution across practice. A notable exception is the concomitant recording of joint trajectories and hitting position in hammering by Bernstein (1935, 1967). Other studies are by Arutyunyan, Gurfinkel, and Mirskii (1968); Darling and Cooke (1987b); Higgins and Spaeth (1972); Vorro (1973); and Worringham (1991, 1993). However, what has not been addressed in these studies is why and how it happens that movements become less "noisy" and why only to a certain nonzero level. How can the variability in the result V(R) be reduced, and how does the variability in the execution V(E) change in this context?

Noise reduction measured in the outcome alone, typically measured in variables that assume a normal distribution of a stochastic process, cannot address these questions. Although variability in the result is most often accompanied by decreasing variability in execution, the latter is not a necessary condition for the former. Admittedly, the analytic method developed below is also not able to give answers about the "why" and the "how," but it does detail when and with what time course noise is reduced. However, in order to theorize about the underlying mechanisms that deal with biological noise, one must first obtain a sound understanding of the phenomenon. Aside from this apparent component, there are two other, less obvious possible components of how variability in the result can decrease during practice, and these partially deterministic sources are illustrated and explained in more detail below. We then quantify the amount of noise that is reduced, given that other functional components are also present.

Second Component: Task-Specific Covariation

How can variability in the result V(R) decrease without a proportional decrease in the variability of the execution V(E)? It was again Woodworth (1899) who highlighted in his studies on drawing and pointing movements the fact that if two variables x and y have an additive relation and have respective standard deviations s_x and s_y, then the variance of the sum $x + y$, s^2_{x+y} is equal to the sum of the two individual variances and twice their covariance:

$$s^2_{x+y} = s^2_x + s^2_y + 2cov(x, y). \tag{1}$$

If x and y do not covary, such that $cov = 0$, the variance in the end result is the sum of the variances of x and y. Hence, $s^2_x + s^2_y$ is referred to as covariation-free variability. Because the covariation can also assume negative values, the total variance can decrease in comparison with the sum of the individual variances. This simple equation illustrates the fact that variability, or more precisely deviations in the individual variables from the mean, can compensate for each other to achieve a result that has less variability than the sum of individual variabilities. The contribution of covariation to the reduction of variability in the result is therefore the difference between the variability of the sum of the two individual variables, $s^2_x + s^2_x$, and the empirically measured variability s^2_{x+y}. This simple notion of subtracting the empirically measured variability from the covariation-free variability to extract the contribution of covariation holds not only for the linear bivariate case but also for the nonlinear multivariate case, which is more common in problems of motor control. However, for this more complex situation, other methods for calculating the covariation-free variability have to be developed. This issue is expanded on below (Müller & Sternad, 2003).

In an early example, Stimpel (1933) studied repeated forearm throwing actions to a target from a gestalt-theoretical perspective. He showed that the dispersion in the result was smaller than what would have been expected from the dispersion measured in the execution variables. He recorded angle and velocity of ball release at fixed release positions. His experimental task was so constrained that these two variables were the only variables that determined the ball's trajectory and, therefore, determined hitting success. The experiments showed that the variability in the movement result (i.e., deviations from target hits) was smaller than the variability measured in the two execution variables, release angle and velocity. In a similar dart-throwing experiment, the same observations were reported by Müller and Loosch (1999) and Loosch (1990). Arutyunyan and colleagues showed for pistol shooting that variations in body and pistol angles compensated for each other (Arutyunyan et al., 1968; Arutyunyan, Gurfinkel, & Mirskii, 1969). More recently, the concepts of these studies on pistol shooting were picked up again by Scholz and colleagues (Latash, Scholz, Danion, & Schöner, 2001; Scholz, Danion, Latash, & Schöner, 2002; Scholz & Schöner, 1999; Scholz, Schöner, & Latash, 2000) and further supported with a more modern methodology and more computationally advanced data analysis. Central to their approach is that control is indicated if a specific constancy is observed—one that is brought about by nonrandom covariation among component processes. Furthermore, there is a series of other investigations in which the existence of covariation has been discussed as indication of control (Abbs, 1986; Abbs, Gracco, & Cole, 1984; Hughes & Abbs, 1976; Bernstein, 1975; Bootsma & van Wieringen, 1990; Cordo, 1988, 1990; Darling & Cooke, 1987a, 1987b; Darling & Stephenson, 1993; Darling, Cole, & Abbs, 1988; Gutman, Latash, Almeida, & Gottlieb, 1993; Kudo, Tsutsui, Ishikura, Tomoki, & Yamamoto, 2000; McDonald, van Emmerik, & Newell, 1989; Newell, van Emmerik, & Sprague, 1993; van Emmerik, 1992; Vereijken, van Emmerik, Whiting, & Newell, 1992; Voigt, 1933). For the present context, covariation between execution variables is a second essential component that can reduce variability in the result.

Third Component: Task Tolerance

In the tasks reviewed above, the desired result was brought about by more than a single combination of values of execution variables. This is the case if the result can vary within boundaries or if the movement system is redundant. For example, in dart throwing, different combinations of release angle and velocity can lead to the same result. For instance, the bull's-eye has a radius such that all results within this tolerance radius count as hits. The space of all variable combinations is referred to as *task space*. In the dart-throwing example, this is a 6-D space that is spanned by the components of the execution vector **e**. A combination of execution variables that lead to "success" (i.e., hitting the bull's-eye) is referred to as a *solution* of the task. The subset in task space that contains solutions is referred to as *solution space* or *solution manifold*, which is in robotics more commonly referred to as *null space*.

In task space, one can find areas where many successful solutions are adjacent (i.e., the solution manifold is large). If an actor intends to achieve a specific result, he or she should aim for those solutions that have other solutions adjacent to them in solution space. Even if an actor is unable to realize this particular solution, due to the inevitable noise, he or she will still achieve values that are close to, or even identical with, the desired solution. Areas of the task space in which there is a high probability that also neighboring combinations of variable values lead to a successful solution are, hence, more stable or "tolerant" with respect to motor noise. Therefore, the third component for improving the probability of success is referred to as *task tolerance*. Graphing the task space with its solution manifold(s) illustrates these areas of high tolerance as subspaces.

In the following, we develop a method by which the relative contributions of the three components discussed above to successful achievement of a task goal can be separated and quantified across sets of trials. We show that the operationalization of these three components can completely describe the improvement in performance. The core idea of this method is demonstrated with the simplified task of playing skittles. The same task is subsequently studied in an experiment, and the data are analyzed with the method. This task has only two execution variables, which simplifies the exposition. However, the method is generalizable to tasks with any number of variables.

The Task: Virtual Skittles

The task in focus is a simplified virtual version of the British pub game "skittles." In skittles, a ball is suspended from a string attached to the tip of a tall vertical post. The player takes the ball, displaces it from vertical while keeping the string taut, and swings it around the post such that it describes an elliptical trajectory. At the opposite side of the post is one skittle or a set of skittles. The goal of the swing is to knock down the skittle or skittles with the ball (Figure 2).

In order to make the task experimentally tractable and to capture it in a formal model, we modified it in the following fashion (Figure 3): The participant held a "paddle" that was fixed to a hinge joint and could be moved horizontally. A ball was shown on a monitor and was "held" by the paddle. When the participant

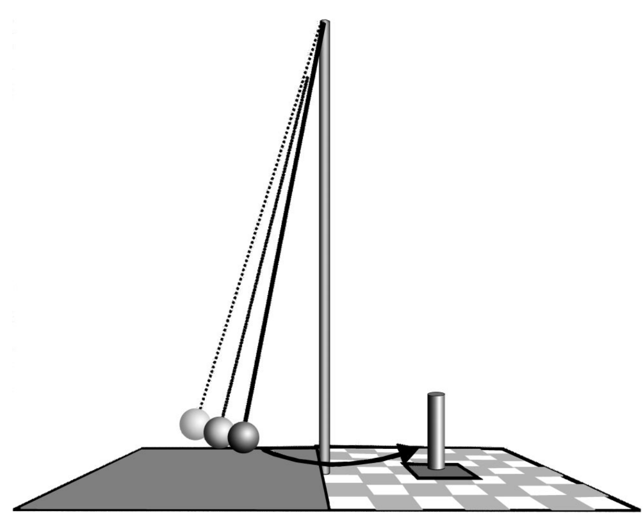

Figure 2. Sketch of the skittles task. A ball is suspended on a string and swings around the center post, with the objective of knocking down the skittle at the opposite side.

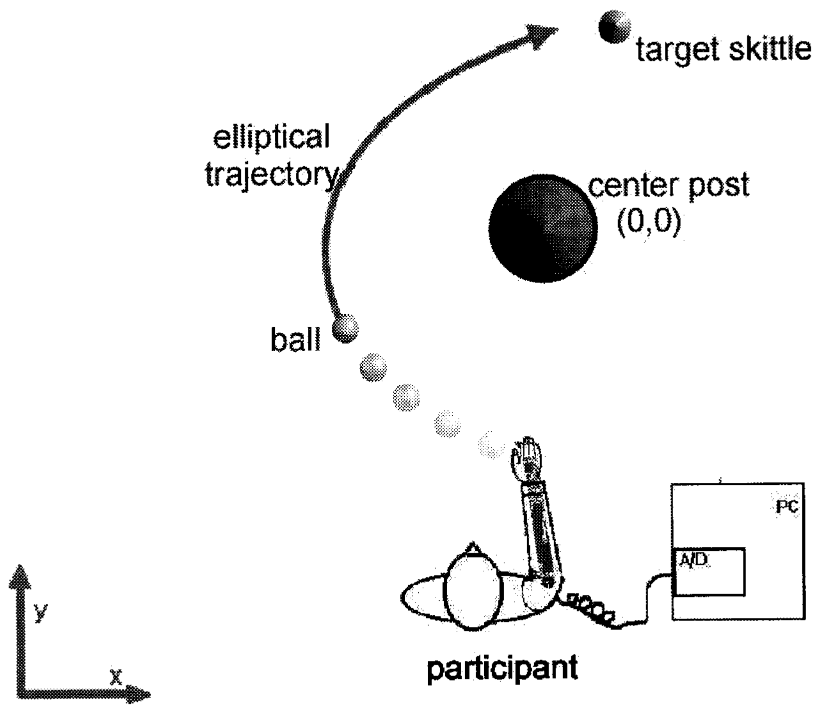

Figure 3. Top-down view of the skittles task with the *x–y* coordinate system. The ball swings in this horizontal plane. The center post is broader than in the real task. The position of the participant's forearm is recorded by the computer.

threw the ball, it accelerated onto a circular trajectory around the hinge joint. The moment of ball release was controlled by the participant by the release of a contact switch. At release, the ball took off with a tangential velocity determined by the release and traversed around the central post in a circular trajectory toward a skittle at the opposite side. This ball trajectory was virtual, and the participant could see it on the computer screen. The participant's view of his or her workspace is shown Figure 4. Because of restoring forces, which were proportional to the distance between the ball and the rest position at the center post, the ball was accelerated back to the center post. This rest position was the origin of the coordinate system ($x = 0$; $y = 0$).

The participant's goal was to hit the circular target skittle exactly at its center of mass. The measure for goodness of performance was the minimal distance *d* between the trajectory and the center of the skittle (see inset of Figure 4). A perfect hit with zero distance to the skittle center was given the value of zero (Trajectory B in Figure 4). The participant could throw the ball with a range of release angles (see the polar coordinates around the paddle in Figure 4). As is shown below, two angle ranges were particularly successful in hitting the skittle and were chosen most frequently by the participants: A release angle of approximately 90° is illustrated by Trajectory A (Figure 4). The angle adopted for Trajectory B is close to −90° where the paddle points toward the participant. To perform with this −90° release angle, participants had to change their orientation to the throwing device and adopt a position different from the one shown in Figure 3, which they were free to do. Trajectories A and B traverse the center post in a clockwise direction, but participants were also allowed to throw the ball onto a counterclockwise path (Trajectory C). The latter direction was chosen less often. Release velocity was defined to be positive when the paddle rotated in a clockwise direction and negative when it rotated in a counterclockwise direction. Note that it was possible to throw the ball onto a counterclockwise path with a positive release velocity (Trajectory C).

The Model

The elliptic trajectories of the ball were generated by a 2-D model system in which the ball was attached by a massless spring to the origin of the coordinate system at the center post (Figure 4). The swinging motions were lightly damped to approximate realistic behavior. At time *t* the equations for the position of the ball in the *x*- and *y*-direction were written as follows:

$$x(t) = A_x sin(\omega t + \varphi_x)e^{-\frac{t}{\tau}}$$

$$y(t) = A_y sin(\omega t + \varphi_y)e^{-\frac{t}{\tau}}. \quad (2)$$

The amplitudes A_x and A_y, as well as the phase differences φ_x and φ_y resulted from the ball's release angle and release velocity including energy conservation. The physical parameters chosen for the experimental task (ω, τ) and the precise calculations are given in Appendix A.

Task Space

If all model parameters are fixed, the performance depends on only two variables: the angle of the paddle at release α and the

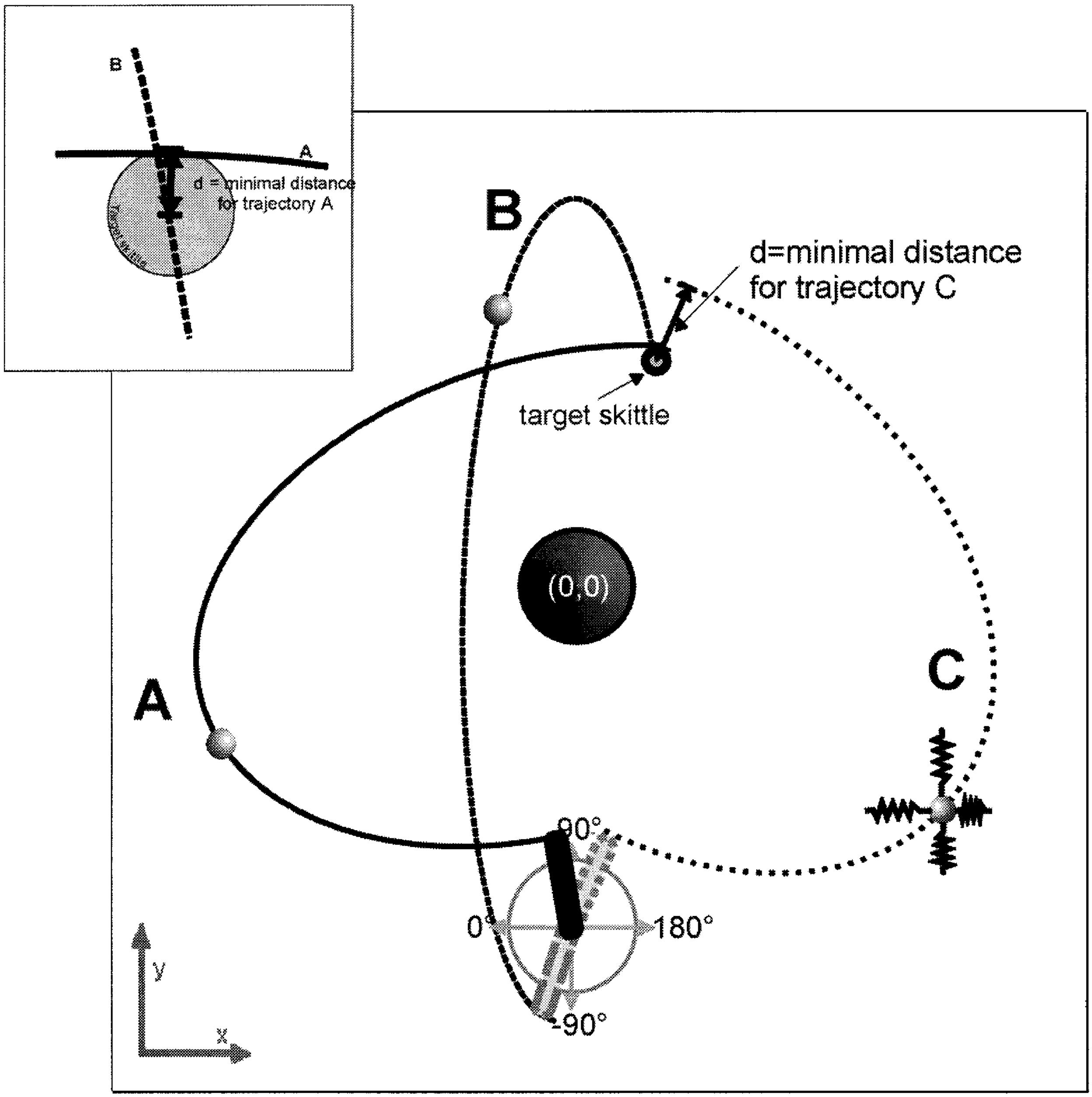

Figure 4. Three different ball trajectories produced by different release angles and velocities (A, B, C). The performance measure is the minimal distance between the trajectory and the center of the target skittle. For Trajectory C, the distance *d* is shown by an arrow. For Trajectories A and B, the distance is shown in the inset image. The area around the target skittle is zoomed (top left) as it is presented to the participants after each throw. Trajectory A has the shown distance *d*. Trajectory B hits the skittle with zero distance. The graphics on the ball on Trajectory C represent the spring-like restoring forces, which are proportional to the distance between the current excursion and the rest position (0, 0).

release velocity of the ball v. For all combinations of α and v, the result achieved in the task can be calculated using Equation 2. For a given position of the target skittle, all combinations of release parameters were entered into Equation 2, and the results were computed. In Figure 5, the results of these calculations are plotted in a 2-D task space for the target position $x = 0.35$ m; $y = 1.00$ m. The regions in white show the best solutions, where the distance to the target skittle is less than 3 cm. The two branches pertain to the two major angle ranges in which the ball can be hit. The two branches are not symmetrical because at $\alpha = 90°$ the distance to the center post is smaller than it is at $\alpha = -90°$. The effect of release velocity would be symmetrical for positive and negative values if friction were neglected, because in the calculations for the amplitudes in the x- and y-dimensions, the velocity is squared (see

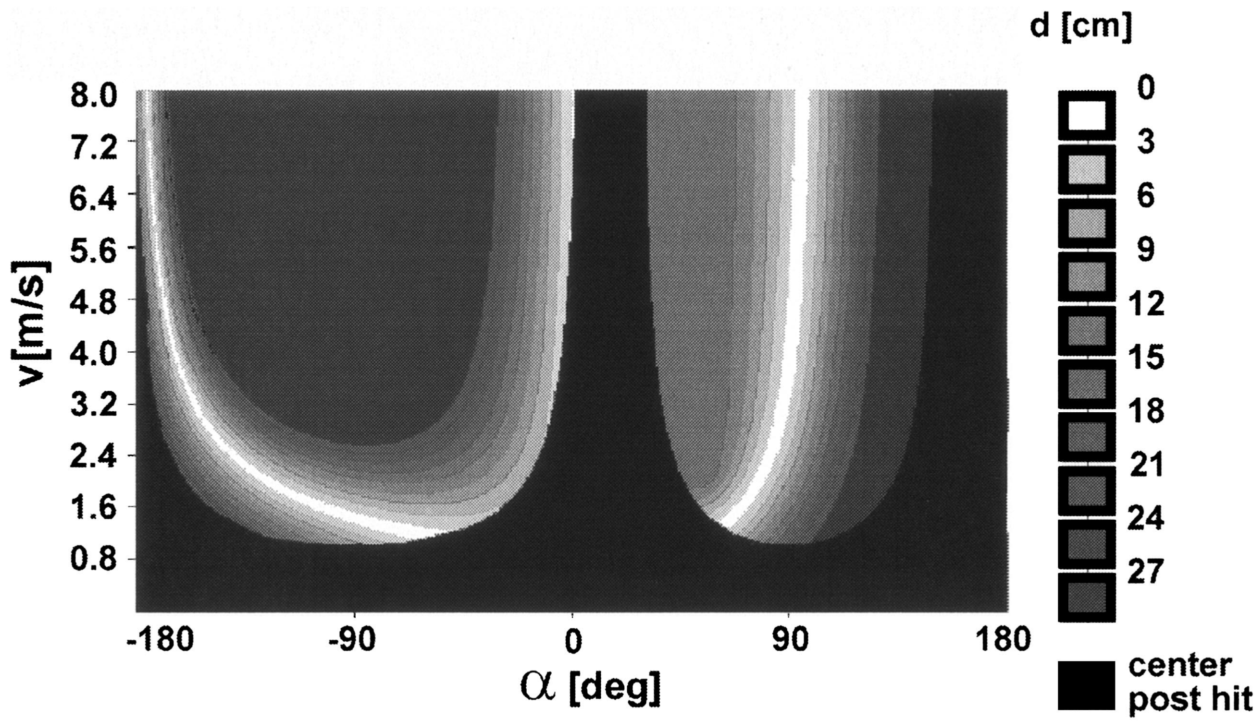

Figure 5. Task space and solution manifold of the virtual skittle task. The position of the target skittle is *x*: 0.35 m; *y*: 1.00 m. For each combination of release angles α and absolute velocities *v*, the color codes the distance *d*. The white area indicates solutions with distances less than 3 cm (solution manifold). The gray shades are codes for other distances. The black areas indicate solutions that lead to collisions with the center post.

Appendix A). Therefore, what primarily matters is the absolute value of the velocity. However, because we assumed damping of the oscillations, small differences for clockwise and counterclockwise ball trajectories arose, even for identical values of release angle and absolute velocity. Hence, the sign of velocity does matter and is determined by the case distinction of the arcsine function. These differences due to different signs of velocity are considered in the calculations of the trajectory. However, because these differences are extremely small, and their consequences for the task space are not visible at this scale, the task space in Figure 5 is only illustrated for positive velocities. The different shades of gray correspond to different degrees of performance success captured as distances *d* from the target.

It is important to note that the task space shows nonlinear bands of successful solutions, or solution manifolds. This means that for a given value of α a certain level of variance in α may lead to a noticeable change in performance, whereas the same variance in *v* produces only negligible changes in the result. Because the bands are of different widths at different locations in task space, they provide different degrees of task tolerance. In this context, tolerance captures the system's stability (i.e., resistance to variability due to noise or perturbations). Because of the inhomogeneity of the task space, task tolerance can provide a significant contribution to the change in result variability. In addition, the component noise reduction is also a possible contributor when the bandwidths are smaller than the variance of the participant. The more noise is reduced, the more likely it is that the result will remain within the desired solution area. Covariation between the execution variables can contribute when the direction of the solution manifold is not parallel to one of the axes. In this case, a change in one variable can be compensated for in the other variable. For example, in Figure 5, in the range where α is close to $-90°$, with increasing angle the corresponding velocity *v* becomes smaller. Hence, in the shown nonlinear solution manifold, covariation can also be exploited. Conversely, the right branch of the solution manifold shows very little possibility for covariation, because for a value of α close to 90°, all values of *v* produce the same result *d*. The fact that the right part of the task space does not afford the exploitation of covariation does not pose a problem because we do not assume that covariation is used every time in invariant fashion but, rather, that it is used only when the task affords it.

Quantification of the Components of Variability

Because our focus is on the decomposition of variability in the executions and results over repeated trials, the application of the method developed in this article requires sets of data—for example, a set of successive trials. If the objective is to extract changes across a series of practice sessions, two or more data sets can be formed from the sequence of practice trials. Similarly, and alternatively, two or more data sets from different individuals can also be compared. For the following exposition, we use two fictive data sets, Set A and Set B, in an arbitrary task space that has a solution manifold similar to that of the skittles task. Figure 6 accompanies and illustrates the following computational steps.

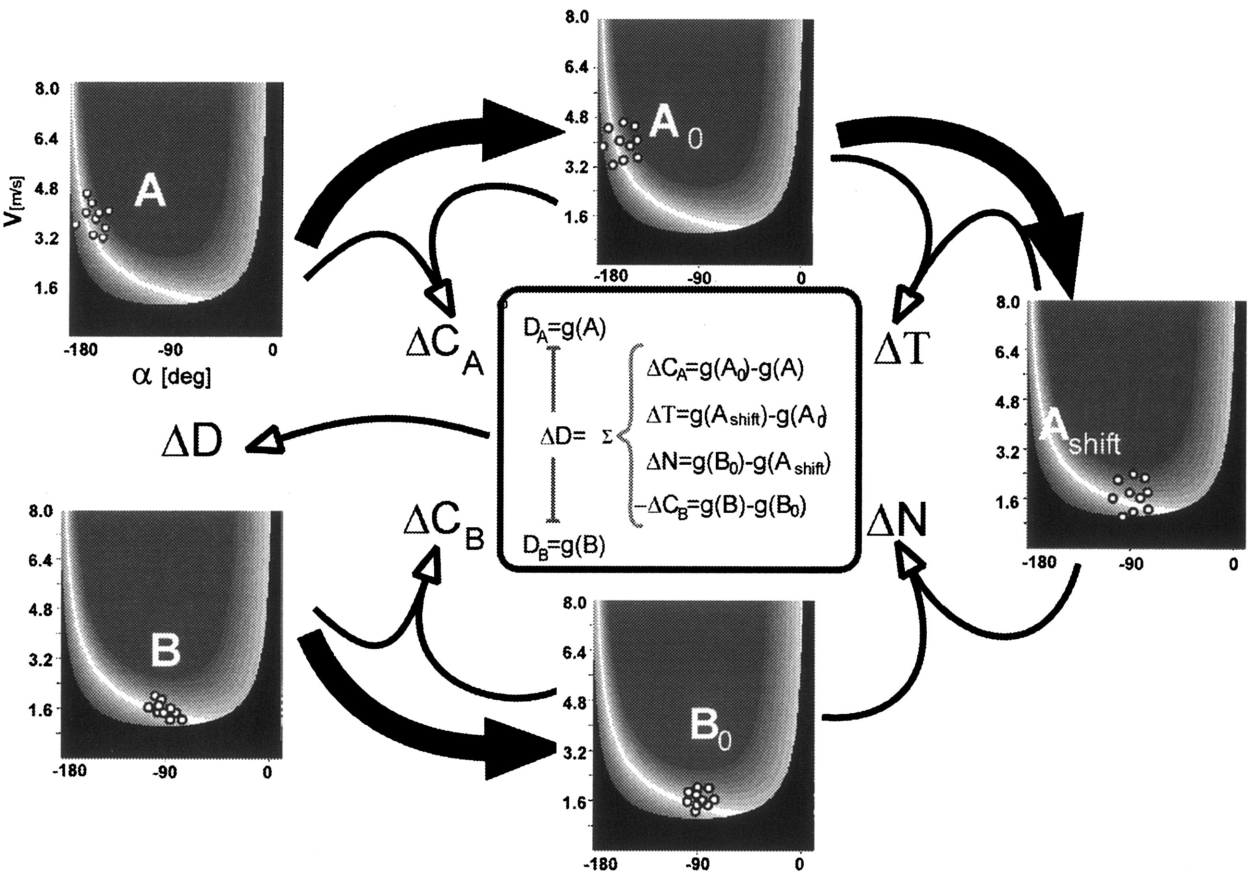

Figure 6. Illustration of the computations quantifying the contributions of the changes in performance. In each panel, several trials are plotted in task space, which is spanned by release angle α and velocity v. To compute the contributions of ΔT, ΔN, and ΔC from the original Data Set A to the second Data Set B, three more data sets are generated (A_0, A_{shift}, and B_0). The specific contributions of the components are the performance differences between data sets.

Figure 6A shows Set A, consisting of 10 data points plotted in the 2-D task space spanned by an angle α and a velocity v. Figure 6B shows Set B, which has less overall variability and a better performance result because the data points are predominantly located in the white (i.e., successful) areas. The three intermediate figures, Figures $6A_0$, $6A_{shift}$, and $6B_0$, illustrate the individual steps of the calculation that compare Set A and Set B. It is shown below that the three components of variability are additive and that they can capture the difference between Set A and Set B completely. Note, however, that the components are not independent of each other, and the sequence in which the contributions of the components are extracted is important.

A single result is captured by the vector $\mathbf{r}$, which contains all variables required to describe the movement result $\mathbf{r} = (r_1, r_2, \ldots, r_j)$. If f is the known function, $\mathbf{r}$ is calculated for each trial i of a series of trials ($i = 1, \ldots, n$) from the k-dimensional execution vector $\mathbf{e} = (e_1, e_2, \ldots, e_k)$ as

$$\mathbf{r}_i = f(\mathbf{e}_i). \tag{3}$$

Each $\mathbf{r}$ is associated with a result measure (e.g., the distance to the target d), which is calculated from the equations that describe the trajectory in the task of skittles. Because the method analyzes the variability and its consequences on the performance over a set of trials, we need a measure that expresses the collective performance. This can be, for instance, the average absolute distance to the target D calculated across n trials. In general form, D over a series of n trials is a function of $\mathbf{r}_i$:

$$D = \overset{n}{\underset{i=1}{t}} (\mathbf{r}_i), \tag{4}$$

where t denotes the function quantifying and aggregating the performance over a set of trials. In the example task of skittles, the function t calculates the trajectory's minimal distance from the target d from the spatial coordinates of the trajectories x and y (Equation 2) and averages them across a set of trials to obtain D. Hence, the performance of the set of trials can be expressed and summarized as

$$D = \overset{n}{\underset{i=1}{t}} [\, f(\mathbf{e}_i)] \text{ or } D = g(E), \tag{5}$$

where g denotes the function that combines both t and f. $\mathbf{E} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ is the set of execution vectors with results $\mathbf{r}_i$. The objective is now to calculate and decompose the difference in performance between the two sets of execution vectors of Set A and Set B, denoted by ΔD:

$$\Delta D = g(B) - g(A). \tag{6}$$

This overall difference in performance ΔD can be decomposed into three potentially unequal contributions. It is important to note that the components covariation ΔC, noise reduction ΔN, and task tolerance ΔT add up to be exactly ΔD[1]:

$$\Delta D = \Delta C + \Delta N + \Delta T. \tag{7}$$

The components' contributions are calculated from the differences across five different data sets, which are generated from the original data sets. Sets A and A_0 and B and B_0 differ only in the component covariation, because their means and variances are identical. Sets A_0 and A_{shift} differ only in their location in task space—that is, they differ in the component tolerance. Set A_{shift} and B_0 have the same means and no covariation; hence, they differ only in their amount of noise.

The first step in the calculations extracts covariation in Set A. To emphasize, covariation exists if values of the execution variables "fit" together in the sense that deviations in one variable are compensated for by deviations in the other variable. This covariation is removed by the permutation of the *n* data points of each variable in Set A separately in order to generate a data set A_0 (Müller & Sternad, 2003). In the skittles task, the variables that are permutated are release angle and velocity. The *n* individual realizations of each variable within the set are permutated such that, subsequently, velocities and angle values from different trials are paired. Due to this new randomized pairing, the target-specific covariation of velocities and angles is removed. If *d* and the collective measure *D* are calculated from the new pairs with function *g* (Equation 5), the result is changed. This transformation is symbolized by the arrow from A to A_0 in Figure 6. Because the values are only permutated, the mean values and the variance remain unchanged in the new Set A_0. Hence, the difference in performance ΔD between Sets A and A_0 can be solely attributed to covariation:

$$\Delta C_A = g(A_0) - g(A). \tag{8}$$

The subscript $_A$ is added to ΔC as because this first calculation only captures the covariation in Set A. To safeguard the calculations against numerical artifacts, the permutation should be repeated several times and the average *D* calculated over all permutations (for a more formal exposition of this method, see Müller & Sternad, 2003). In the experiment presented later, we performed 10 repetitions of the permutation.

In the same fashion, the amount of covariation is calculated for Set B, which is permutated to produce a Set B_0, which only differs from Set B in that covariation is removed. The overall contribution of covariation ΔC in the comparison of Set A and Set B is then the difference of ΔC_A and ΔC_B, as shown in Figure 6:

$$\Delta C = \Delta C_B + \Delta C_A = g(B) - g(B_0) + g(A_0) - g(A). \tag{9}$$

For a numerical illustration of these calculations, see Appendix B.

The second step is to calculate the shift in "location" of the two data sets in task space, or task tolerance ΔT. Tolerance ΔT captures what is gained by locating the data set to a different place in task space (i.e., where the region of successful solutions is larger), such that data are more tolerant to noise (structurally stable). ΔT is calculated by "shifting" Set A_0 in task space to the average location of Set B or B_0. The new data set, A_{shift}, is generated as follows: If each data point in A_0 is denoted by vector $\mathbf{a}_i$, the average of each entry of $\mathbf{a}_i$ is calculated across all *i* in A_0, $\bar{\mathbf{a}}$. For instance, the averages of all angles and all velocities of Set A_0 are calculated. In the same way, for each data point of Set B, $\mathbf{b}_i$, the averages of all angles and velocities is calculated in $\bar{\mathbf{b}}$. The difference between the mean locations is added to each data point $\mathbf{a}_i$ of A_0 to yield $\mathbf{a}_{i,shift}$:

$$\mathbf{a}_{i,shift} = \mathbf{a}_i + (\bar{\mathbf{b}} - \bar{\mathbf{a}}). \tag{10}$$

This transformation of the data corresponds to the transition from Figure $6A_0$ to Figure $6A_{shift}$. The distribution of the covariation-free Set A_{shift} is unchanged but is now at the same location in task space as Set B. The contribution of tolerance to performance *D* is obtained when *D* for Set A_0 is subtracted from Set A_{shift}:

$$\Delta T = g(A_{shift}) - g(A_0). \tag{11}$$

Now the data sets A_{shift} and B_0 are at the same location in task space, meaning that all differences due to *T* have been removed. Hence, the remaining difference in performance between Set A_{shift} and Set B_0 can be exclusively ascribed to the component noise reduction ΔN:

$$\Delta N = g(B_0) - g(A_{shift}). \tag{12}$$

If Equations 9, 11, and 12 are added, one obtains:

$$\Delta C + \Delta T + \Delta N = g(B) - g(B_0) + g(A_0) - g(A) + g(A_{shift}) - g(A_0) + g(B_0) - g(A_{shift}) \tag{13}$$

or

$$\Delta C + \Delta T + \Delta N = g(B) - g(A) = \Delta D. \tag{14}$$

This demonstrates that the three components are indeed additive and capture the difference between two sets completely. It should be pointed out that the calculations of intermediate Data Sets A_0, A_{shift}, and B_0 implicitly generate a sequence in which the components are extracted: first, covariation ΔC; second, tolerance ΔT; and third, noise reduction ΔN. The result of the decomposition is not independent from this sequence. However, any change in this sequence will yield implausible results. If, for instance, the data set were shifted first, the contribution of covariation could no longer be evaluated because the direction of the solution space may have changed at the new location. A potentially perfect covariation could no longer be captured. The comparison of the contribution of ΔN is similarly only meaningful if both data sets are at the same location. Hence, the quantification of the components is not independent of the sequence, but the suggested sequence is the only plausible one.

[1] Note that we use the notations ΔC, ΔN, and ΔT when the *change* in performance is quantified. When we refer to the three components in a general fashion, we use *C*, *N*, and *T*.

Experiment

In order to test whether these theoretically derived factors indeed provide an empirically useful method for quantifying skill improvement, we conducted an experiment in which participants performed the virtual skittles task. Participants practiced a sequence of five task constellations, which differed only in the location of the target skittle. Improvements in performance with practice were analyzed with respect to the relative contributions of the three components to the change within a constellation as well as across constellations. This yielded five different task and solution spaces, which are shown in Figure 7. Participants had to perform all five task configurations in sequence.

If the five task spaces are compared, it is evident that there are two solution manifolds at positive and negative release angles that are separated by an area of no success. The left manifold pertains to negative release angles, and the right set of bands pertains to positive ball releases angles (see Figure 4 for the angle directions). The transition between target positions therefore requires changes of the location in task and solution space (i.e., a change in the component task stability ΔT). The different task positions also have different orientations of the solution manifolds. In Target Position 1, many negative release angles lead to successful performances (i.e., these solutions are "stable" or tolerant against deviations of the release angle). Target Position 5, however, has a high tolerance against deviations in the release velocity when the release angle is close to 90°. The transition between target constellations requires also that covariation and noise reduction be used if accurate and reliable performance is to be achieved.

Method

Participants. Forty-two students from the University of the Saarland, Saarbrücken, Germany, volunteered to participate in the experiment. These 20 men and 22 women were between 17 and 53 years of age (average age = 28 years). Participants were randomly assigned to one of two groups. As a consequence, Group G1-5 comprised 10 women and 11 men and Group G5-1 consisted of 12 women and 9 men. Each group performed the same task, consisting of a sequence of five target constellations, but in opposite sequences. Prior to data collection, the participants were instructed about the purpose of the experiment, and they filled out the consent form in agreement with the regulations of the University of the Saarland.

Apparatus. Participants operated a lever arm in the horizontal plane that corresponded to the paddle movements on the visual display. This lever arm was made of light-weight aluminum and was cushioned with foam to provide comfortable support. The one end close to the elbow was poised on a needle bearing to allow free rotation in one angular dimension. The displacements of the lever arm were measured by a potentiometer with a sampling rate of 700 Hz. At the hand's end of the lever, the participant grasped a spherical wire net. By flexing the index finger he or she could close a contact switch at the tip of the lever. Opening the contact was translated into a release of the ball, and the ball trajectory was shown on a computer monitor. Figure 4 shows the workspace from a top-down view, as it was displayed to the participants. The height of the lever was adjusted for each participant so that in upright standing position his or her upper arm was vertical, and the forearm was placed horizontally on the foam-padded lever arm. The standing position of the participant was not prescribed, and he or she could stand to the right or left of the vertical fixation, depending on whether he or she aimed to hit the ball with a negative or positive release angle or whether the ball was propelled in a clockwise or counterclockwise direction around the center post. The lever arm could move an

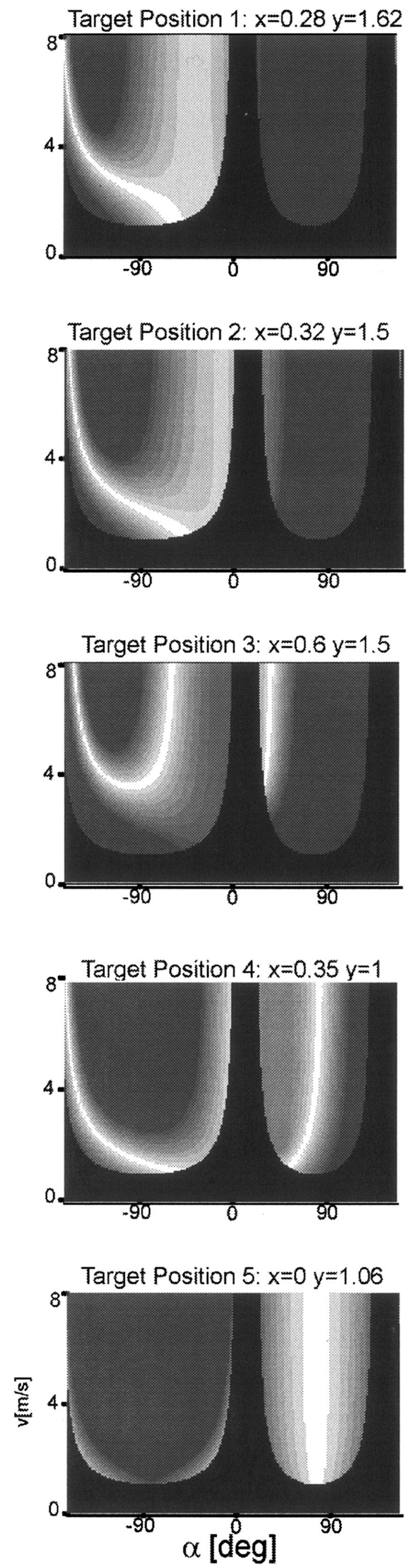

Figure 7. The solution manifolds for the five target positions used in the experiment. The different shades of gray indicate the degree of success in the performance.

almost complete 360° circle, with the exception of a 10° arc in the transition region of the potentiometer. However, this critical section was in an area that was irrelevant for the execution of the task. If the ball touched the center post, it was stopped and the distance between the post and the ball was counted as the performance measure.

The visual display was presented on a 14-in. (35.56-cm) monitor that was placed at eye height at 1.2-m distance from the participant. With this presentation, there was a 90° transformation between the behavior that took place in the horizontal plane and the vertical presentation on the computer screen. Also, the top-down view into the work space did not have the same compression as if the scene were viewed in the horizontal plane. Despite these two deviations from a realistic presentation of the task and workspace, none of the participants reported any problem with this transformation, either in the first warm-up sessions or throughout the experiment. The data collection and the calculation of the ball trajectories for the virtual display were performed by a 486 PC using Turbo Pascal software (Version 7.0). Delay between data collection and display were within the refresh rate of the computer, which was 13 ms.

Design and procedure. Each participant was presented all five target positions (T1, T2, . . ., T5) in sequence. One group of participants (G1-5) practiced the five constellations in the order from T1 through T5. The second group (G5-1) performed the same five constellations in reverse order. For each of the five target positions, the participant performed 320 trials, yielding a total of 1,600 trials. The study was conducted in four practice sessions on 4 days and all sessions had to be completed within 1 week. In each session, the participant performed 400 trials. After 320 trials, a new target position was presented. Deliberately, therefore, the transition between different constellations occurred at different times of the four practice sessions. In Session 1, the transition occurred after 320 trials; in Session 2, after 240 trials; and so forth. This overlapping organization was planned to control for a number of possibly confounding variables, such as warm-up and retention effects (see Figure 8). For the data analysis, the 320 trials of each target position were subdivided into four blocks consisting of 80 trials each. Each session lasted approximately 30 min.

In Session 1, participants were introduced to the experiment with a standardized instruction. To familiarize themselves with the apparatus, they practiced the task with a target position that was not used later in the experiment. They were told to hit the center of the target skittle with the ball in its first semicircular trajectory. Because the trajectory was generated as a damped oscillation, it would otherwise be possible to hit the pendulum strongly only once, and with the ball's slowly decreasing amplitude, it would hit the skittle in its second or third traversal around the center post.

In the experiment proper, the participants executed a series of eight throws in a self-initiated sequence but without substantial breaks. On average, one throw was performed every 4.5 s. By closing the contact switch, participants "held" the ball at the end of the lever, and the next throw could be performed at any time. The workspace shown on the monitor corresponded to a 2 × 2-m square around the center post. To provide participants with exact information as to whether and how the ball hit the target, the distance between the ball trajectory and the target skittle was enlarged after each trial, as shown in Figure 4. A presentation of 1-s duration had proven sufficient in pilot trials for participants to take in this information. Note that, although the ball might only "brush" the target skittle, the goal was to hit the center of the skittle. In addition, the

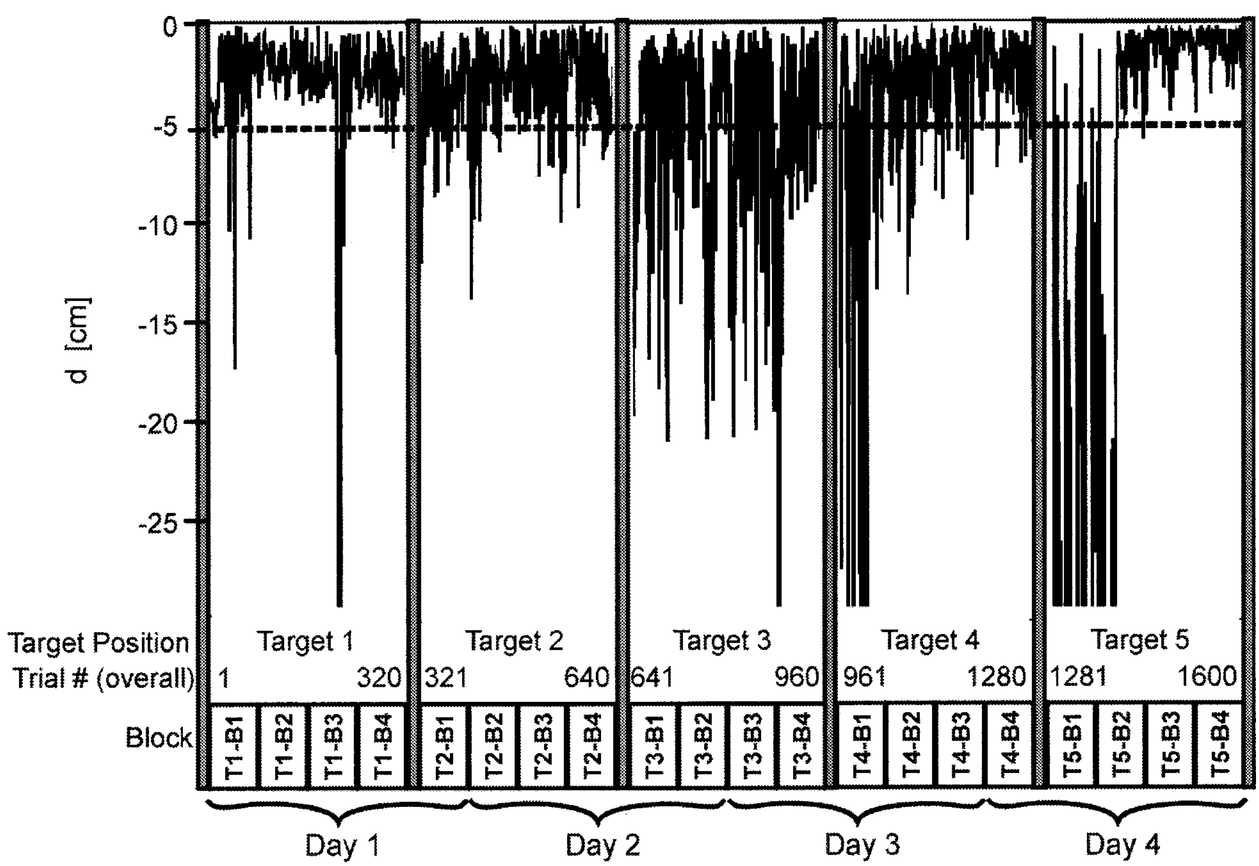

Figure 8. Time series of the performance measure *d* (minimum distance to target) of all 1,600 trials of 1 participant. The dashed line at the top of the figure represents the width of the target (5 cm). Trials with *d* in this area hit the target skittle but not necessarily in the center.

participant was given feedback on his or her performance in terms of an average score after a series of eight throws. The participants were encouraged to achieve the lowest possible average score across all series. Each practice session, with its 400 throws, lasted approximately 30 min. None of the participants reported any fatigue with this number and frequency of throws.

Data reduction and dependent measures. For each trial, the execution variables release angle α (degree) and release velocity v (m/s) were measured and used online to calculate and display the elliptic trajectory of the ball (see Appendix A). For each trajectory, the minimal distance to the target d was calculated. The distance d was defined as a negative value (e.g., an absolute distance of 5 cm is reported as –5 cm). Thus, improvements in performance in the result calculations correspond to increasing values. Note that in the experiment, participants were given feedback such that low scores represented good performance. For each series of 8 trials, D was calculated. To quantify the improvement in performance between blocks, ΔD was calculated as the difference of the D values of two blocks. To calculate ΔD in the first block of one target position, D of Block 1 was compared with a fictive D that would have been achieved if participants had continued to throw with the same average values as in Block 4 of the preceding target position.

The primary focus of the analyses was the contributions of the three factors ΔT, ΔN, and ΔC to changes in performance. For each block of 80 trials, the mean of the d values was calculated and used as performance measure D (cm) = negative of the mean distance to target. Since there were 4 blocks for each target position, this yielded a total of 20 block values for each measure per participant. The notation T2-B3, for example, refers to Target Position 2 and Block 3, combining performance in Trials 161–240. This same notation is used for both groups of participants. Note that, for instance, the trials of T2-B3 were performed at different times by the two groups. G1-5 practiced their first 320 throws in Target Position 1, whereas G5-1 had already practiced 960 trials in the Target Positions 5, 4, and 3.

Results

Descriptive results of performance. For a first impression of the change of the performance over practice, the time profile of the result measures of the 1,600 trials d is shown for 1 participant in Figure 8. The five target positions are separated by vertical bars that highlight the “perturbation” of the performance at the transition to each new constellation. The dashed line at $d = -5$ cm indicates the width of the target such that throws with d values smaller than -5 cm hit or at least brushed the skittle. d was subject to large fluctuations, and the amount of variation for different target positions was between 5 cm in T1 and 20 cm in T3. It is therefore difficult to infer improvements in performance directly across constellations, especially because the participant began with very good performance at T1. In T5-B1, the target was initially never hit because the participant first continued to throw the ball with a strategy that was only successful for T4. Only from T5-B2 onward was a better solution discovered that lead to consistently good performance.

Figures 9A and 9B summarize the progress in the performance score of all participants but separately for each of the two groups. The average D for all participants for one block was calculated and plotted against blocks. For G1-5, performance improvements can be seen within one task constellation, but each time the target position was changed, there was a marked decrease in performance. In contrast, for G5-1, the performance improvements can be seen both within and across task constellations.

If D became larger (i.e., less negative), the throws became more accurate, and the minimum distance to the target skittle decreased with increasing practice. To evaluate whether there was indeed improvement, we computed ΔD across blocks, testing whether the cumulative change from B1 to B4 was indeed greater than zero within each target position. Ten one-sample t tests were performed on the ΔD values, pooled for each group and target position, and these were compared with zero. An adjustment of the level of significance was not necessary because (a) these analyses did not test the hypotheses but, rather, whether the central assumptions for the decomposition method were satisfied, and (b) the 10 hypotheses are AND connected such that the prerequisites are only satisfied if all 10 tested cases are significantly different from zero. An adjustment of the p level was only necessary if the partial hypotheses were OR connected. It should also be mentioned that the assumption of normality was satisfied for each t test. The violation of a normal distribution, mentioned below, only arose when all cases were pooled. Therefore, the variances of the individual cells were not homogeneous, as can be easily appreciated from inspection of Figure 8, and this violated the assumption of homogeneity of variance for conducting an analysis of variance (ANOVA). For these reasons, a set of 10 t tests was preferred to a three- or four-way ANOVA.

With one exception, it could be verified that practice indeed lead to significant improvements. The p values for each block are given in Figures 9A and 9B. The exception was seen in G5-1 ($p = .078$) when participants practiced T1 after T2. This absence of improvement can be explained in light of the fact that for T1 and T2, the task and solution space have very similar characteristics. The participants who performed T2 already showed such high performance scores that improvements in T1 were only moderate. This interpretation was also supported by the results of G1-5, who practiced the two constellations in reverse order. Improvements in T2 were also only small, because participants had already reached a high level of performance in T1.

Performance change in task space. A more informative depiction of the changes across practice of the five target positions is obtained when the data points are plotted in task space spanned by release angle and velocity. Figure 10 gives an impression of the execution over the course of practice for the same participant from Group G1-5 shown in Figure 8. Each of the nine panels depicts the data points of one block of 80 trials for the target positions T3 (B4), T4 (B1, B2, B3, B4), and T5 (B1, B2, B3, B4). The sequence of nine blocks illustrates how the location and size of the data cloud changes with the transitions across the constellations. In the first transition, from T3-B4 to T4-B1, it is evident how the participant very quickly finds better angle–velocity combinations directly after the transition to the new target position T4-B1. The white arrow points to the location of the first trials in the block, showing that only few trials are needed to relocate the execution variables to the new solution space of T4. In the subsequent T4-B2 and T4-B3, the variance decreases. In addition, the data cloud becomes slightly more elliptic to better align with the horizontal direction of the solution manifold. With the change to T5, a new solution becomes necessary, and the data immediately become more scattered in both angle and velocity, evidently reflecting the search for a new solution. This corresponds to the very bad performance reflected in low d values in Figure 8. In contrast to T4-B1, a shift of the data cloud against the gradient becomes necessary. The white arrow points again at the first data points of the block. By searching for new angle–velocity combinations, the

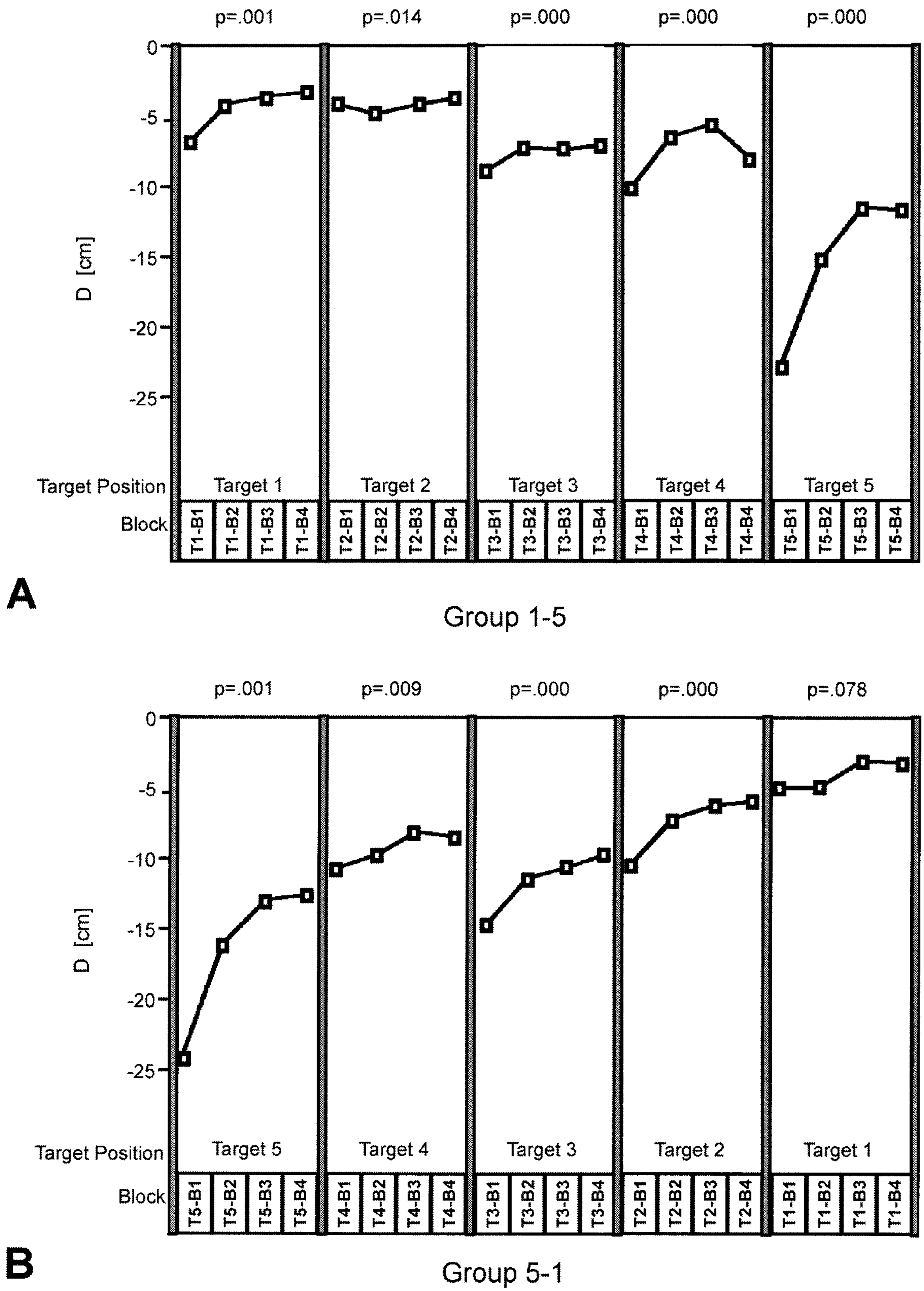

Figure 9. Time course of the performance measure *D* (distance to target). The data points represent the average score of all participants in one block. A: Group G1-5. B: Group G5-1.

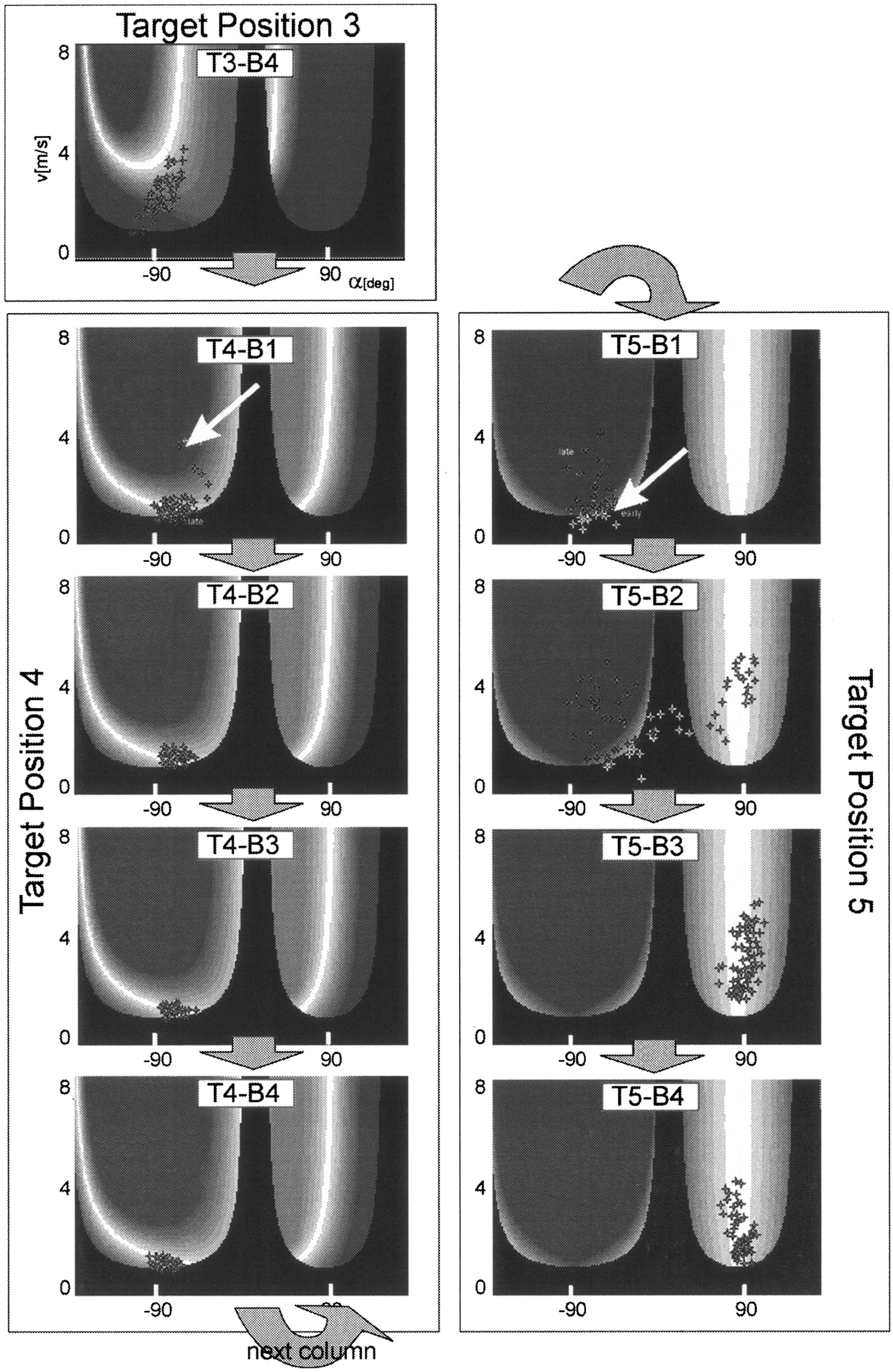

Figure 10. Presentation of the data in task space of the same participant as in Figure 8. The nine panels display the nine blocks from T3-B4 to T5-B4. Each data point is the execution variables of one trial. The solution manifold is the calculated one as shown in Figure 7. The arrows point out the transition from one target position to the next.

participant first performed worse trials in the dark area between the two manifolds. In T5-B2, the search has become successful, and trials are increasingly performed with correct angle and velocity combinations. In T5-B3 and T5-B4, the variance is reduced, and the shape of the data cloud gets adjusted to the shape of the solution manifold.

Note that the search for a new solution manifold, as seen in the transitions from T3-B4 to T4-B1, leads to large values of ΔT. This large ΔT, however, is a relocation from unsuccessful to successful solutions that is conceptually different from optimizing a solution by placing it into an area that affords higher resistance to perturbations (i.e., more tolerant areas). Yet, as can be appreciated in the transition from T3-B4 to T4-B1, this search occurs relatively quickly. In contrast, in the transition between T4-B4 and T5-B1, the search lasts longer. In the current sequence of target constellations, this means that ΔT after a target-position change has to be interpreted with caution.

Components of variability. Figure 11 shows the relative contribution of the three components ΔC, ΔT, and ΔN. The coordinates are almost identical to those for Figures 8 and 9 (note that the range of *D* values is slightly larger), and the solid black lines show the same *D* values across the blocks as in Figure 9. Of interest are the shaded areas that represent the relative contributions of the three components to performance improvements from one block to the next. The contributions of the components were graphed in a cumulative fashion beginning from below with ΔT, followed by ΔN and ΔC. The dark grey at the bottom indicates ΔT, the light grey indicates ΔN, and the hatched areas indicate ΔC. The lower border of the shaded area is determined by the performance that would have been achieved in the new target position if the release parameters had not been changed from the previous block and none of the three components had been utilized. From this bottom line, the contributions are graphed in a cumulative fashion, and the sum of all components produced the actual performance (i.e., the *D* values shown as solid lines). Two things need to be pointed out: (a) There is no baseline for the very first blocks (i.e., T1-B1 for G1-5 and T5-B1 for G5-1) because there is no reference value from preceding performance. This is indicated by the fuzzy lower border. These values were not included in the following statistical analyses. (b) Sometimes, the magnitudes were negative. This lead to the fact that the upper border was sometimes higher than the overall performance, as for instance is the case for ΔT in T3-B1. Hence, the contribution of the following component ΔN was negative and could not be shown as an area but can be induced when there is a gap—in this case, between the dark grey and the hatched area.

Qualitative inspection of these illustrated results shows that the component ΔT dominated in almost all task constellations for both groups. It is also clear that the component ΔN showed the largest change within a target constellation. At the beginning of practice of one target position, this component was irrelevant or would even bring about a decrement in performance (e.g., T3-B1 in Figure 11B). When the target position was changed, it became necessary to explore new and stable solutions, which then necessarily increased the variance. Only when a new solution was found could the variance be reduced. The component ΔC is represented by the hatched area. It evidently played a relatively subordinate role in this task.

To quantify the sequential effect of target positions, the data of all participants of one group was pooled, and the evolution of one execution variable across the five constellations was summarized in a histogram. Figure 12 shows five histograms of α for each of the two groups. For each group, considerable changes in the frequencies of α were evident at the transitions between the different target positions. Note that the temporal sequence of the practice sessions for G1-5 is from top to bottom, whereas for G5-1 it is from bottom to top. From a juxtaposition of the constellations for the two groups, it is evident that the temporal succession of the target positions mattered and lead to choices of different solution areas. For instance, T4 had two equally large solution manifolds, one in the area of release angle 90° and one at −90° (Figure 7). In G1-5, the participants were confronted with this constellation after they had collected experience in T3 that α of −90° produced reliable results. Correspondingly, participants chose this option. This was different for G5-1. Before T4, participants performed in T5, where only solutions with α of 90° were possible. Hence, in T4, participants of G5-1 continued to choose solutions with α of 90° more frequently than α of −90°. This comparison indicates hysteresis.

Obviously, the changes in release angle were accompanied by changes in release velocity, as shown in Figure 10. However, these changes were in a smaller range, between 1 m/s and 5 m/s, and changes were less obvious in a histogram.

Statistical confirmation of contributions of ΔT, ΔN, ΔC. The above results verified what was conceptually prepared for in the introduction: (a) For each target position, there were significant improvements in performance across the 320 practice trials. (b) The changes in performance within and at the transition between target positions were captured by three different components of variability. (c) These components contributed unequally and explained different aspects of the observed changes in performance. Although the analyses above demonstrated the theoretical framework by example, it remains to be statistically confirmed that all three components contribute to performance improvement, or ΔT, ΔC, and ΔN are greater than zero (Hypothesis 1), and the relative contribution of the components changes in the course of practice (Hypothesis 2). If these hypotheses are supported, the decomposition of variability yields helpful explanatory information to understand improvements in performance and can be generalized to other tasks.

To test the hypotheses, we aggregated the data of ΔT, ΔC, and ΔN for each block over all participants. To check the prerequisites for an ANOVA, we tested the distribution properties of each of the 40 data sets by the Kolmogorov–Smirnov test with Lilliefors adjustment. In 33 of the 40 cells, a *p* value of .001 was obtained, which signaled that the assumption of a normal distribution was violated. As evident from Figure 11, the data in the individual cells indeed have different means and (not shown in Figure 11) also different variances, which prohibited the application of parametric tests like ANOVAs. For this reason, the data were transformed to a different scale neglecting the absolute magnitudes. For the test of Hypothesis 1, the data were transformed to a nominal scale; for testing Hypothesis 2, the data were transformed to an ordinal scale.

Testing Hypothesis 1: ΔT, ΔC, *and* ΔN *are greater than zero.* A component contributes if it leads to a positive change more often than a negative change. The null hypothesis is that both positive and negative values occurred with 50% probability. Using a bino-

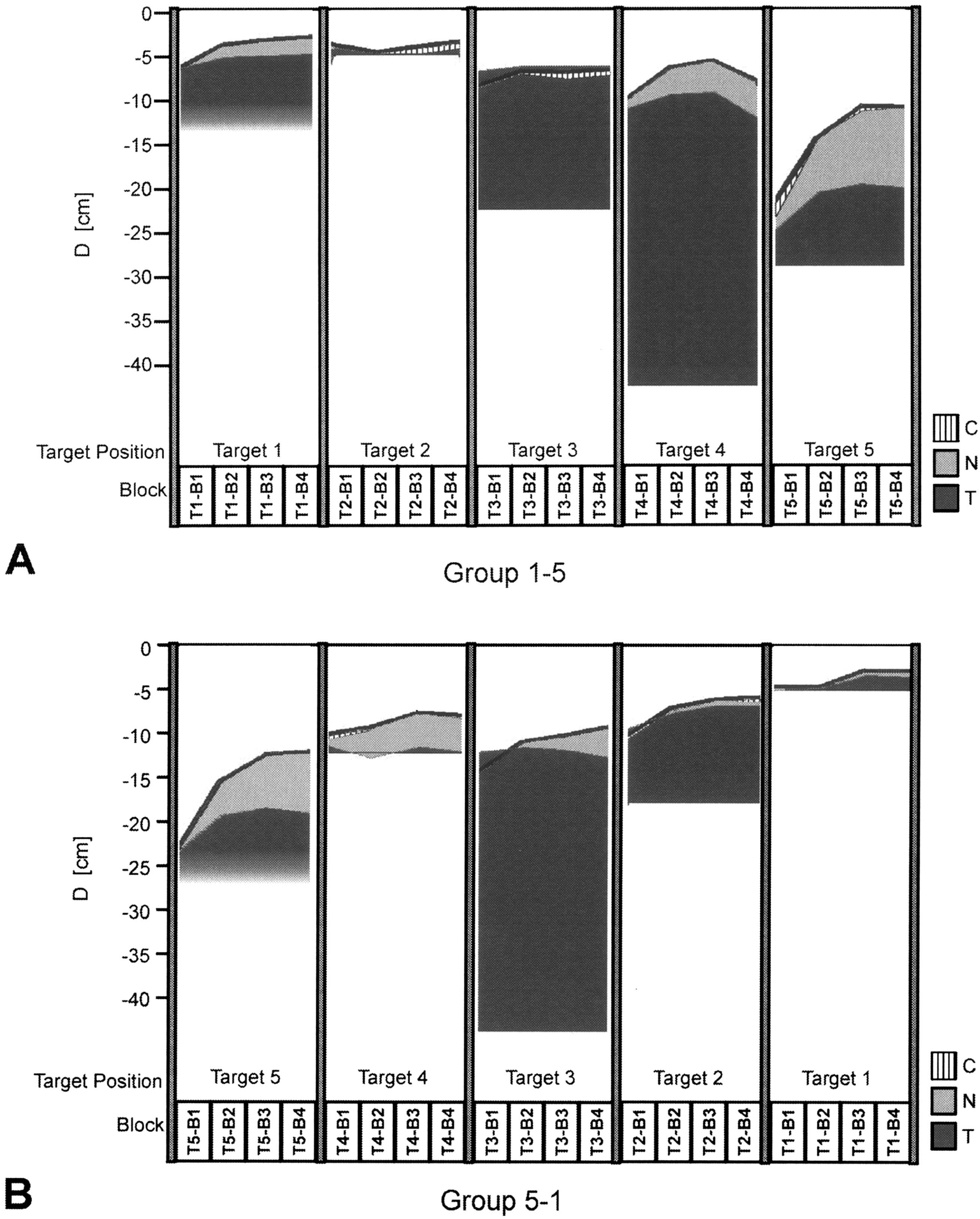

Figure 11. Contributions of the three components to the change in performance for the two groups: G1-5 (A) and G5-1 (B).

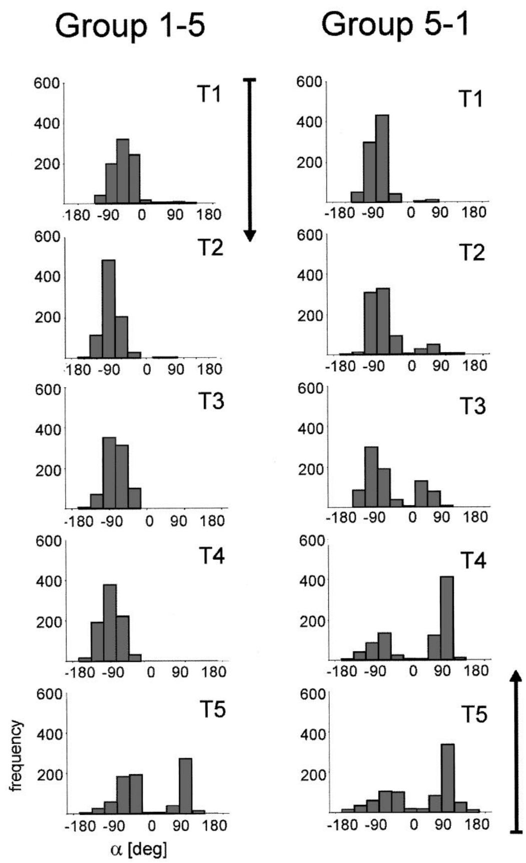

Figure 12. Frequency distributions of the release angles α for all five target positions for both groups. The data of all participants for one group are pooled.

mial test, we evaluated whether the empirical frequencies differed from this expectation. Results showed that for all three components the positive contributions occurred more often than negative ones. For ΔT and ΔN, this was highly significant, with p values greater than .0001. These results are listed in Table 1. Moreover, to achieve a positive overall effect, the positive contributions have to be at least as large as the negative ones, even if there were more negative contributions. If this condition is not satisfied—as, for instance, with three small positive contributions in combination with a single but large negative value—this leads to a decrement in performance. When the averages of the absolute positive and negative contributions were compared, ΔT and ΔN gave a clear picture. The positive contributions were always larger than the absolute values of the negative ones. This was confirmed by t tests ($p < .001$). Table 1 shows that the difference of the means was also consistent with this expectation. However, for ΔC, the argument must be moderated because, as already shown in Figure 11, the average negative contribution was indeed slightly higher than the average positive contributions. However, this difference was not significant in a t test ($p = .377$).

To summarize, all three components lead to positive performance increments rather than to decrements. For ΔT and ΔN, performance increased in an absolute sense. For ΔC, this could not be established. As already shown in the descriptive data presentation, the contributions of ΔT and ΔN were unambiguously present, wheras ΔC was subordinate for the present task.

Testing Hypothesis 2: Changes of relative contribution of ΔT, ΔC, *and* ΔN. An evaluation of the changing prominence in contribution of the three different components faces the problem that, depending on the target constellations and their sequence, the quantitative amounts cannot be compared directly. A change in performance due to ΔN by 3 cm in one condition cannot be evaluated as more important than a change due to ΔT by 1 cm in a different constellation. What can be compared, though, are the changes in the same target positions. Therefore, the comparisons of the relative contributions of ΔN, ΔT, and ΔC were performed “locally” and within each participant. To this end, the rank order in which the three components contributed to the improvement was determined. The component that had the largest contribution was given Rank Order 1, the one with the least contribution received Rank Order 3. For example, Participant 9 has the following values for the change from T4-B4 to T5-B1: $\Delta T = 8.58$ cm = Rank 1, $\Delta N = 1.75$ cm = Rank 2, $\Delta C = -0.08$ cm = Rank 3. Subsequently, the values for all participants in all 40 blocks were summed for each block. Figure 13 summarizes the frequency of the individual factors that held Rank 1 in each of the four blocks. The differences between the components were statistically confirmed with a Pearson χ^2 test, $\chi^2(6, N = 839) = 98.99$, $p = .0001$.

Figure 13 illustrates the fact that, in Block 1, ΔT was utilized most. In the subsequent blocks, ΔT diminished in importance, and ΔN and ΔC started to play a more prominent role during the later part of practice.

Discussion

In goal-oriented tasks, such as grasping an object or hitting a target, the quality of performance is defined by the accuracy and

Table 1
Results of Tests of Hypothesis 1

Component	N	Percentage of values: Positive	Percentage of values: Negative	p	Mean absolute value of contributions: Positive	Mean absolute value of contributions: Negative
ΔT	798	39.6	60.4	.000	1.97	6.12
ΔN	798	41.4	58.6	.000	1.27	2.31
ΔC	798	46.8	54.2	.009	0.73	0.67

Note. N is the number of estimates for the three components. The p values are for a one-tailed binomial test.

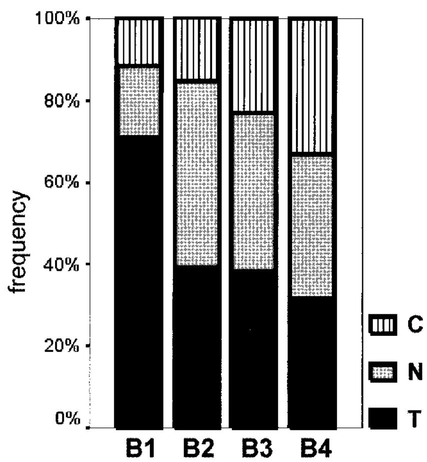

Figure 13. Relative number of Rank 1 (in percent) of the individual components when their relative contribution to performance improvement was computed. The data of the individual blocks were averaged across all five target positions and both groups.

replicability of the result over repeated attempts. The result of the performance—defined, for example, as how safely one grasps a glass or how accurately one hits a target—is highly dependent on the accuracy and consistency of the execution. Execution refers to how the many involved joints are configured in the grasping action or with what velocity, angle, and position the ball is released to hit the skittle. Different combinations of these execution variables can lead to the same result. This observation has often been invoked in the redundancy problem and poses one of the core questions in motor control and learning. Which solutions are chosen from the many possible ones to obtain a desired result, or, conversely, which solutions are avoided to achieve a high probability of success? This article proposes that the analysis of variability of both the result and the execution and particularly their relation can provide insight into this question.

The present study examined the virtual hitting task of skittles, in which the result was quantified by the deviations of the ball trajectory from the target skittle—that is, some "error" measure. The execution was defined by the two variables that completely determine the ball trajectory: angle and velocity at ball release. Redundancy in the task is given by the fact that different combinations of the two execution variables are not necessarily reflected in different results. In order to understand how actors control such goal-oriented behaviors, it is informative to look at both levels of the performance. The relation between execution and result variables is captured and visualized in the so-called "task space" that is spanned by the execution variables. Solutions with different levels of success, as described by the distance from the target, define different values in this space. The skittles task was chosen because the solution manifolds were nonlinear, and for different target locations there were two separate subsets. As such, the development of performances over the course of practice and adaptations to new target locations provided interesting challenges to the actor and potentially interesting features for the analysis. The presented method showed that a change in result variability could be decomposed into three orthogonal components, referred to as tolerance *T*, noise reduction *N*, and covariation *C*.

The experiment produced a number of significant results. Overall, the three components *T*, *N*, and *C* contributed in a positive fashion to the improvement in the result. Further, the three factors contributed in a unequal fashion to the change in the result. It was shown that the factors *T* and *N* were dominant in the improvement of performance, whereas *C* was not as strongly utilized in the current task. Over the course of practice of one target position, it was *T* that was the first factor to be used to obtain an increment in performance. This meant that new sets of angle and velocity combinations were explored to find a larger region of the solution manifold that provided a larger tolerance for error. Over the four successive blocks, *C* became more prominent, and *N* remained a relatively invariant but nevertheless important factor.

When target positions changed, persistence in the types of solution was noted. For instance, when a new target location made a qualitative change in the release angle necessary, participants first tended to persist with solutions similar to the ones from the previous target position. This effect of history was most evident in the comparison of two groups of participants who practiced the sequence of five skittle targets in opposite sequence. The same target was performed with qualitatively different solutions, depending on the preceding target. Transitions between solution types were observed at different times, indicating hysteresis. However, this persistence could also have been due to a rule-based behavior. A deeper understanding of the possible mechanisms underlying this observation has to be relegated to future investigations.

Noise reduction. That a decrease in noise or stochastic fluctuations around the mean performance was recorded as a function of practice may at first sight be regarded as a trivial and well-documented result. Many studies on motor learning have focused on the understanding of the time course of improvement in a motor task over practice, most prominently by analyzing the standard deviations of the result measure in different estimates, ranging from absolute error to variable error. Most frequently, these studies have reported a largely monotonic decrease of the variability in the result. Recently, the question of whether this time course is of exponential nature or indicative of a power law has been discussed (Newell, Liu, & Mayer-Kress, 2001). With a view toward the distinction into the three components, however, this observation in the result variables does not necessarily reflect the same changes in the execution variability. If one, for instance, considers the transition from T5-B1 to T5-B2 in Figure 10, this shows a case in which the dispersion in the execution variables increases with improving results. In contrast, the transition from T5-B3 to T5-B4 shows the dissociation between execution and result variability such that dispersion decreases without directly improving the result. At Target Position 5, the solution manifold runs parallel to the *y*-axis (velocity). With a release angle of 90°, the release velocity can take on almost any value, and the actor still hits the skittle. A decreased variability in velocity has almost no consequences for the result variability.

Tolerance. While there is a host of studies that have analyzed different aspects of variability, the majority of them have focused solely on the stochastic component of variability. The component tolerance *T* has not been recognized in the literature before. This aspect of performance change can only be appreciated when both execution and result variables are considered and their relation is analyzed in task space. The present results show that the increasing use of T is an essential contributor to performance improvement. The early phase of practice of a specific task variant is typically characterized by a large contribution of T. But a caveat is necessary for the interpretation of such large initial changes: The quantification of ΔT cannot tease apart whether large values are due to improvement of tolerance or to the searching of the task space. The latter can, in some cases, confound the core conceptual content of tolerance. In later phases of practice, when the average execution variables already achieve successful solutions, the nonzero ΔT that we found in the results expresses the more subtle exploitation of tolerant areas in task space.

If one considers the fact that when a successful location in task space is sought, the variance necessarily increases as well and, most markedly, the contribution of noise N increases. This explains the increase in N in the early practice phases. This observation is in line with the often-stated phenomenon that a functionally positive side effect of variability is the exploration of the workspace (Newell & Corcos, 1993). The increase of dispersion, as seen in T5-B2 in Figure 10, is a good example for this. The large scatter of data points reflects the exploration of the new task demands. Note that when seeking new solutions, the actor often purposefully varies the solutions, and the registered variability therefore is only partially “noise,” in the sense of uncontrolled fluctuations. However, the presentation of task and solution space provides a tool to dissociate exploratory “variability” from uncontrolled “motor noise.”

The present findings confirm the interpretation of one of the results in a previous study on dart throwing (Müller, 2001). Using the same decomposition method as in the present study, the data showed that T played only a minor role. As explanation, it was suggested that the participants were in the stable region of the solution manifold from the very beginning such that no more shifts to more optimal locations could be registered. It was concluded that in this more practiced task, this factor was already exploited before data collection began.

The task space identifies regions of different size and shape corresponding to different degrees of success in the solutions. In fact, the task space, with its bands of solutions of different success, can be viewed as a gradient field. If each trial is a data point, the sequence of data points can be viewed as a search through this gradient field until the best location in the solution manifold is found. The best location is defined not only by the minimal deviation from the target but also by the width of the solution. A region in the solution manifold with larger width allows more variability with a lower probability of inducing a decrement in the result. The search through this gradient field was highlighted best in the current results when transitions between different skittle constellations that had different types of solution manifolds were investigated. In some transitions, the sequence of data followed the gradient, as observed in the exemplary data of 1 participant shown in Figure 10. In the transition from T3-B4 to T4-B1, the data points followed the gradient. Conversely, the transition between T4-B4 and T5-B1 required a change in task space, with the crossing of an area of unsuccessful trials to exploit tolerance afforded by the right branch. Figure 10 suggests that the gradient-following search leads to success more directly and requires less practice trials. This raises the more general question of to what degree practice is a gradient-sensitive process to the “minimum” in the solution manifold. The present data, however, do not allow a conclusive evaluation of such differences between the gradient-following and the gradient-climbing search.

Covariation. The results of the present task show that the contribution of covariance is rather modest, contrary to expectation. Yet, this result should not downplay the fact that this component can have a marked influence on the improvement in other tasks. In fact, there are numerous studies that have examined covariation between task dimensions and which are listed above. While all of these studies examined covariation between variables, the examination was often limited to the study of a linear covariation or correlation between two variables. When more than two variables were involved, and when the relation was no longer linear, this method was no longer applicable. As a solution, a number of methods and transformations were used that are only marginally adequate. For instance, the logarithmic transformation to linearize exponential data is confounded with many problems (Elzinga, 1985; see also Müller & Sternad, 2003). Also, the application of the method of determining the uncontrolled manifold has prerequisites to the data, such as validity of a linear approximation and equidistance in all dimensions, which limit the applicability of the method (Scholz & Schöner, 1999). Only a few recent studies have used randomization techniques that provide an unbiased approach for this quantification (Kudo et al., 2000; Müller, 2001; Müller & Loosch, 1999). The present analyses also applied this randomization method to quantify the covariation between the two execution variables with their nonlinear relations. Note that this method is also applicable for more than two variables. For a formal development and proof of this method, see Müller and Sternad (2003).

Independent of the methodological issues raised in Müller and Sternad (2003), all of these studies reviewed above agree that covariation is a decisive factor for the achievement of high scores in goal-oriented tasks and that this factor grows in importance with practice. Müller (2001) gave an overview of experiments in which the proportion of covariation in the performance was quantified. For instance, for a basketball freethrow, it was shown that in expert performance the variability of the result was only 20% of the variability that would have been expected from the variability recorded in the execution. In other words, with the use of covariation, variability was reduced by 80%. Aside from this outstanding example for the use of covariation, it can also be seen that the extent of covariation can be very different for different tasks and levels of expertise on performance. That C is small in the skittle task may be due to the fact that C requires a longer practice time to be exploited sufficiently.

To conclude, in this article we present a new quantitative method that can parse the variability observed over repeated trials into three components. This method is not confined to the specific example analyzed in this study but has many different applications, wherefore the method has been laid out in a more generalized fashion. Another example is variability in the redundant limb in the performance of a reaching or pointing task. The execution variables are the joint angles, and the result variable is the accuracy of the pointing. Prerequisites for the application of the method are that the numerical relation f between the execution and the result in a single trial is known and that the result can be captured in some measure D that quantifies the success of a set of trials. Although the analysis of the present task focused on single data points in repeated performances, the method can be similarly applied to evaluate movement trajectories if a suitable measure for good and reliable execution can be found, such as, for instance, a root-mean-square error. In a more general sense, the method assumes a theoretical status in that it proposes that variability in execution is explored and exploited for the achievement of a successful solution and that three conceptually different means for this exploitation are available to the actor. As such, this article presents a first methodological step to facilitate further empirical work that should test this method in different experimental paradigms and with different questions—for example, long-term acquisition and retention or the adaptation to changed task spaces with different degrees of complexity. With a broader database, this method will advance toward a theoretical framework for the understanding of how the central nervous system deals with the seemingly inevitable noise in both the system itself and its relation to the environment.

References

Abbs, J. H. (1986). Invariance and variability in speech production: A distinction between linguistic intent and its neuromotor implementation. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 202–225). Hillsdale, NJ: Erlbaum.

Abbs, J. H., Gracco, V. C., & Cole, K. J. (1984). Control of multijoint coordination: Sensorimotor mechanisms in speech motor programming. *Journal of Motor Behavior, 16,* 195–231.

Arutyunyan, G. H., Gurfinkel, V. S., & Mirskii, M. L. (1968). Investigation of aiming at a target. *Biophysics, 13,* 536–538.

Arutyunyan, G. H., Gurfinkel, V. S., & Mirskii, M. L. (1969). Organization of movements on execution by man of an exact postural task. *Biophysics, 14,* 1162–1167.

Bernstein, N. A. (1935). Ocerki po fiziologii dvizenij i fiziologii aktivnosti [Essays on the physiology of movements and physiology of activity]. *Archiv Biologiceskich Nauk, 38*(1), 1–34.

Bernstein, N. A. (1967). *The coordination and regulation of movements.* London: Pergamon Press.

Bernstein, N. A. (1975). *Bewegungsphysiologie* [Physiology of movement]. Leipzig, Germany: Barth.

Bootsma, R. J., & van Wieringen, P. C. W. (1990). Timing in attacking forehand in table tennis. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 21–29.

Cordo, P. J. (1988). Kinesthetic coordination of a movement sequence in humans. *Neuroscience Letters, 92,* 40–45.

Cordo, P. J. (1990). Kinesthetic control of a multijoint movement sequence. *Journal of Neurophysiology, 63,* 161–172.

Crossman, E. R. F. W., & Goodeve, P. J. (1983). Feedback control of hand-movement and Fitts' Law. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 35*(A), 251–278.

Darling, W. G., Cole, K. J., & Abbs, J. H. (1988). Kinematic variability in grasp movements as a function of practice and movement speed. *Experimental Brain Research, 73,* 225–235.

Darling, W. G., & Cooke, J. D. (1987a). Changes in the variability of movement trajectories with practice. *Journal of Motor Behavior, 19,* 291–309.

Darling, W. G., & Cooke, J. D. (1987b). Movement related EMGs become more variable during learning of fast accurate movements. *Journal of Motor Behavior, 19,* 311–331.

Darling, W. G., & Stephenson, M. (1993). Directional effects on variability of upper limb movements. In K. M. Newell & D. M. Corcos (Eds.), *Variability and motor control* (pp. 65–88). Champaign, IL: Human Kinetics.

Drill, R. (1933). Der Hammerschlag [On the action of hammering]. In O. Klemm & F. Krüger (Eds.), *Neue psychologische Studien: Motorik* [New psychological studies: Motor control] (Vol. 9, pp. 142–163). Leipzig, Germany: Leipzig University Press.

Elzinga, C. H. (1985). A note on estimation in the power law. *Perception & Psychophysics, 37,* 175.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47,* 381–391.

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology, 67,* 103–112.

Fullerton, G. S., & Cattell, J. M. (1892). *On the perception of small differences* (Vol. 2). Philadelphia: University of Pennsylvania Press.

Gutman, S. R., Latash, M. L., Almeida, G. L., & Gottlieb, G. L. (1993). Kinematic description of variability of fast movements: Analytical and experimental approaches. *Biological Cybernetics, 69,* 485–492.

Harris, C. M., & Wolpert, D. M. (1998, August 20). Signal-dependent noise determines motor planning. *Nature, 394,* 780–784.

Higgins, J. R., & Spaeth, R. K. (1972). Relationship between consistency of movement and environmental condition. *Quest, 17,* 61–69.

Hughes, O. M., & Abbs, J. H. (1976). Labial-mandibular coordination in the production of speech: Implications for the operation of motor equivalence. *Phonetica, 33,* 199–221.

Kudo, K., Tsutsui, S., Ishikura, T., Tomoki, I., & Yamamoto, Y. (2000). Compensatory coordination of release parameters in a throwing task. *Journal of Motor Behavior, 32,* 337–345.

Latash, M. L., Scholz, J. F., Danion, F., & Schöner, G. (2001). Structure of motor variability in marginally redundant multi-finger force production tasks. *Experimental Brain Research, 141,* 153–165.

Loosch, E. (1990). *Ganzheitsprinzip und Variabilität in der Motorik* [Variability and the "whole" in the control of movements] (Vol. 5). Kassel, Germany: Gesamthochschule-Bibliothek.

McDonald, P. V., van Emmerik, R. E. A., & Newell, K. M. (1989). The effect of practice on limb kinematics in a throwing task. *Journal of Motor Behavior, 21,* 245–264.

Meyer, D. E., Smith, J. E. K., Kornblum, S., Abrams, R. A., & Wright, C. E. (1989). Speed–accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. In M. Jeannerod (Ed.), *Attention and performance XIII: Motor representation and control* (pp. 173–226). Hillsdale, NJ: Erlbaum.

Müller, H. (2001). *Ausführungsvariabilität und Ergebniskonstanz* [Variability and result invariance in the execution of goal-oriented movements]. Lengerich, Germany: Pabst Science.

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. (2003). A randomization method for the calculation of covariation in multiple nonlinear relations: Illustrated at the example of goal-directed movements. *Biological Cybernetics, 89,* 22–33.

Newell, K. M., & Corcos, D. M. (1993). Issues in variability and motor control. In K. M. Newell & D. M. Corcos (Eds.), *Variability and motor control* (pp. 1–12). Champaign, IL: Human Kinetics.

Newell, K. M., Liu, Y.-T., & Mayer-Kress, G. (2001). Time scales in motor learning and development. *Psychological Review, 108,* 57–82.

Newell, K. M., & Vaillancourt, D. E. (2001). Woodworth (1899): Movement variability and theories of motor control. In M. L. Latash & V. M. Zatsiorsky (Eds.), *Classics in movement science* (pp. 410–435). Champaign, IL: Human Kinetics.

Newell, K. M., van Emmerik, R. E. A., & Sprague, R. L. (1993). Stereotypy and variability. In K. M. Newell & D. M. Corcos (Eds.), *Variability and motor control* (pp. 475–496). Champaign, IL: Human Kinetics.

Plamondon, R., & Alimi, A. M. (1997). Speed/accuracy trade-offs in target-directed movements. *Behavior and Brain Sciences, 20,* 279–303.

Scholz, J. P., Danion, F., Latash, M. L., & Schöner, G. (2002). Understanding finger coordination through analysis of the structure of force variability. *Biological Cybernetics, 86,* 29–39.

Scholz, J. P., & Schöner, G. (1999). The uncontrolled manifold concept: identifying control variables for a functional task. *Experimental Brain Research, 126,* 289–306.

Scholz, J. P., Schöner, G., & Latash, M. L. (2000). Identifying the control structure of multijoint coordination during pistol shooting. *Experimental Brain Research, 135,* 382–404.

Stimpel, E. (1933). Der Wurf [On throwing]. In F. Krüger & O. Klemm (Eds.), *Motorik* [On motor control] (Vol. 9, pp. 109–138). München, Germany: Beck.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience, 5,* 1226–1235.

van Emmerik, R. E. A. (1992). Kinematic adaptations to perturbations as a function of practice in rhythmic drawing movements. *Journal of Motor Behavior, 24,* 117–131.

Vereijken, B., van Emmerik, R. E. A., Whiting, H. T. A., & Newell, K. M. (1992). Free(z)ing degrees of freedom in skill acquisition. *Journal of Motor Behavior, 24,* 133–142.

Voigt, E. (1933). Über den Aufbau von Bewegungsgestalten [On the structure of movement gestalts]. In O. Klemm & F. Krüger (Eds.), *Neue psychologische Studien*: *Motorik* [New psychological studies: Motor control] (Vol. 9, pp. 1–32). Leipzig, Germany: Leipzig University Press.

Vorro, J. R. (1973). Stroboscopic study of motion changes that accompany modifications and improvements in throwing performance. *Research Quarterly, 44,* 216–226.

Woodworth, R.S. (1899). The accuracy of voluntary movement. *Psychological Monographs, 3*(2), 1–119.

Worringham, C. J. (1991). Variability effects on the internal structure of rapid aiming movements. *Journal of Motor Behavior, 23,* 75–85.

Worringham, C. J. (1993). Predicting motor performance from variability measures. In K. M. Newell & D. M. Corcos (Eds.), *Variability and motor control* (pp. 53–63). Champaign, IL: Human Kinetics.

Appendix A

Physical Parameters of the Experimental Task and Calculation of the Release Parameters α and v

The parameters that were used for the simulation of the present experiment were as follows:

Mass of the ball:	$m = 0.1$ kg
Spring constant:	$k = 1$ N/m
Relaxation time:	$\tau = 20$ s
Frequency:	$\omega = \left(\frac{k}{m} - \frac{1}{\tau^2}\right)^{0.5} = 3.16$ rad/s
Rest position of the pendulum:	$x_0 = 0$ m; $y_0 = 0$ m
Position of the center post:	$x_0 = 0$ m; $y_0 = 0$ m
Radius of the center post:	0.25 m
Axis of rotation of the paddle:	$x_{P0} = 0$ m; $y_{P0} = 1.5$ m
Length of the paddle:	$\ell = 0.4$ m
Radius of the ball:	0.05 m
Radius of the target skittle:	0.05 m
Position of the target skittle:	varies depending on the constellation.

The coordinates (x, y) of the ball at release result from the angle of the paddle (α) and the velocity of the ball at release (v) as follows:

$$x = x_{P0} - l\cos(\alpha), \tag{A1}$$

and

$$y = y_{P0} + l\sin(\alpha). \tag{A2}$$

The energy in the x-direction (E_x), consisting of a kinetic and a potential component, can be formulated as follows:

$$E_x = 0.5(mv_x^2 + kx_r^2), \tag{A3}$$

where v_x is the x-component of the velocity at release, and x_r is the x-coordinate at release.

The amplitude in the x-direction (A_x) is

$$A_x = \left(\frac{2E_x}{k}\right)^{0.5}. \tag{A4}$$

The phase difference φ_x is calculated as follows, using case distinctions for the arcsin function:

$$\varphi_x = \arcsin\left(\frac{x_r}{A_x}\right). \tag{A5}$$

The calculations of E_y, A_y, and φ_y are performed in analogous fashion. The values A_x, A_y, φ_x, and φ_y are inserted into Equation 2 in the text.

Appendix B

Numerical Example of the Calculation Steps

A numerical example explicitly demonstrates the calculation steps and shows that the method can be generalized from the two-to-one mapping illustrated in Figure 6 to more complex many-to-many mappings. The example has three execution variables, x, y, and z, and two result variables r_1 and r_2. The relation between execution and result is defined as follows:

$$r_1 = y - x, \tag{B1}$$

and

$$r_2 = z - x. \tag{B2}$$

The success of performance in the single task criterion D is calculated over a data set with n trials ($i = 1, \ldots, n$) as

$$D = \frac{1}{n} \sum_{i=1}^{n} |r_{1i}| + |r_{2i}|. \tag{B3}$$

The exemplary Data Set A consists of three trials and Data Set B consists of four trials (indicated by bold figures in Table B1). This should illustrate that the number of executions in the individual sets can be different as long as the success measure is defined independently of the number of values, by its mean value.

Step 1: From Data Set A, we generate the Data Set A_0 by permuting the sequence of values within the components. To exclude chance effects in this permutation, either the set of all value combinations is calculated or the permutation is repeated a sufficient number of times. In this numerical example, the permutation has been performed twice, whereas we used 10 repetitions in the experiment. Here, we only indicate this by repeating the permutation twice. The two resulting sets were added to generate $D(A_0)$ with Equation B3. Consider that the Sets A and A_0 have the same mean and variance because they consist of the same values. They only differ in covariation. Assuming the permutation is repeated sufficiently often, A_0 has zero covariation.

Step 2: Set B_0 and $D(B_0)$ were generated from Set B in the same manner.

Step 3: Set A_{shift} is obtained when the difference between the mean values of Set A and Set B are added to each component of A_0. The x-values of Set A_{shift} are obtained by adding 60 to each data point, 50 to obtain y-values, and 70 to obtain z-values. Hence, A_{shift} and B_0 only differ in the magnitude of their variance.

Step 4: $D(A_{shift})$ and $D(B_0)$ are obtained from Equation B3. Their difference is only noise, as tolerance and covariation has been extracted.

Step 5: The contributions of covariation ΔC, tolerance ΔT, and noise reduction ΔN are calculated as follows:

$$\Delta C = D(B) - D(B_0) + D(A_0) - D(A)$$

$$= 15.00 - 10.38 + 10.50 - 1.50 = 13.62,$$

$$\Delta T = D(A_{shift}) - D(A_0) = 14.17 - 10.50 = 3.67,$$

Table B1
Numerical Examples for the Calculation Method

Data set and trial	x	y	z	r_1	r_2	D
A						1.50
1	**11**	**12**	**13**	1	2	
2	**21**	**22**	**23**	1	2	
3	**31**	**32**	**33**	1	2	
A_0						10.50
1′	11	32	13	21	2	
2′	31	12	23	−19	−8	
3′	21	22	33	1	12	
1″	31	12	33	−19	2	
2″	21	22	13	1	−8	
3″	11	32	23	21	12	
A_{shift}						14.17
1′	71	82	83	11	12	
2′	91	62	93	−29	2	
3′	81	72	103	9	22	
1″	91	62	103	−29	12	
2″	81	72	83	−9	2	
3″	71	82	93	11	22	
B_0						10.38
1′	80	74	95	4	15	
2′	82	70	91	−12	9	
3′	79	73	94	−6	15	
4′	83	71	92	−12	9	
1″	82	70	92	−12	10	
2″	79	73	91	−6	12	
3″	80	71	94	−9	14	
4″	83	74	95	−9	12	
B						15.00
1	**79**	**70**	**91**	−9	21	
2	**80**	**71**	**92**	−9	21	
3	**82**	**73**	**94**	−9	21	
4	**83**	**74**	**95**	−9	21	

and

$$\Delta N = D(B_0) - D(A_{shift}) = 10.38 - 14.17 = -3.79.$$

Note that the contributions can be negative, indicating that there is an increase in noise. The following calculations confirm that the three components are additive and capture the changes in performance completely:

$$\Delta D = D(B) - D(A) = \Delta C + \Delta T + \Delta N = 13.62 + 3.67 - 3.79 = 13.50.$$

Received October 30, 2002
Revision received August 13, 2003
Accepted September 5, 2003 ■