Motor Learning: Changes in the Structure of Variability in a Redundant Task

Hermann Müller and Dagmar Sternad

Abstract Although variability is a fundamental and ubiquitous feature of movement in all biological systems, skilled performance is typically associated with a low level of variability and, implicitly, random noise. Hence, during practice performance variability undergoes changes leading to an overall reduction. However, learning manifests itself through more than just a reduction of random noise. To better understand the processes underlying acquisition and control of movements we show how the examination of variability and its changes with practice provides a suitable window to shed light on this phenomenon. We present one route into this problem that is particularly suited for tasks with redundant degrees of freedom: task performance is parsed into execution and result variables that are related by some function which provides a set of equivalent executions for a given result. Variability over repeated performances is analyzed with a view to this solution manifold. We present a method that parses the structure of variability into four conceptually motivated components and review three methods that are currently used in motor control research. Their advantages and limitations are discussed.

Introduction: Variability, Control, and Learning

Some of the very first empirical studies in motor control have already pointed out that humans never seem to reproduce a movement in exactly the same fashion, even if they try. More than a century ago Woodworth's pioneering studies demonstrated this phenomenon at the example of simple line drawing movements (Vaillancourt & Newell, 2001; Woodworth, 1899). With the development of more sophisticated measurement techniques it has become even more evident that variability is a ubiquitous and fundamental characteristic of human performance. Two more recent edited volumes that are dedicated to

---

D. Sternad (✉)
Departments of Kinesiology and Integrative Biosciences, 266 Rec Hall, University Part, PA, Pennsylvania State University, USA
e-mail: dxs48@psu.edu

D. Sternad (ed.), *Progress in Motor Control*,
DOI 10.1007/978-0-387-77064-2_23,

variability in human movement document its fundamental importance and pervasiveness in all issues of motor control (Davids, Bennett, & Newell, 2006; Newell & Corcos, 1993). Variability has become established as a signature of skilled or rather unskilled performance. A low level of variability has been interpreted as an indicator of control but the absence of variability, stereotypy, is also a sign of disease. On the other hand, variability has been viewed as a prerequisite for flexibility and adapability, the hallmark of biological behavior. With these many perspectives, analysis of variability has become a window into many different research questions on movement generation and, accordingly, has been quantified in many different ways. The following overview aims to present one route to quantify sources of variability with the goal to better understand movement generation and learning.

For many researchers Bernstein's famous recording of hammering movements brought this problem to the fore (Bernstein, 1967). Given the many references to this example it is instructive to revisit the original "cyclogram", adapted in Fig. 1, as it illustrates the spectrum of problems that are still the focus of active research in today's motor control. In presenting these recordings Bernstein highlights that despite the same topology of the repeated hammering trajectories, there are noticeable variations across the repetitions. Assuming that this movement trace was performed by an experienced smith who aimed to hit the same target repeatedly, it may astonish how much variance was present, especially around the maximum position. Replicated by numerous studies since, it has become generally accepted that, although experienced actors show lower variability, there always remains a base level of variability even in such relatively simple and skilled movements.

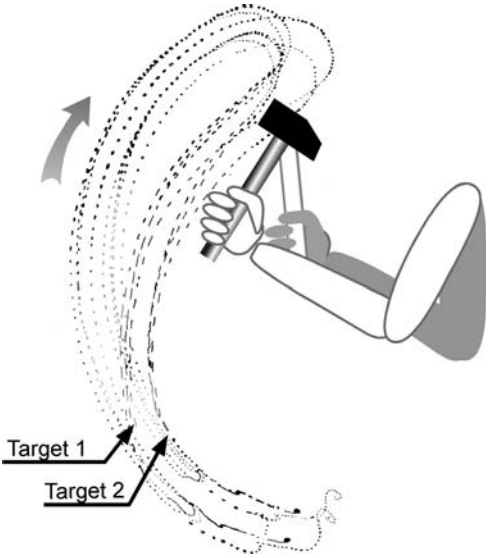

Fig. 1 Recordings of hammering movements (Bernstein, 1967); the arm and hammer movements are added for illustration

Yet, the cyclogram exemplifies more than this evident feature. Closer inspection reveals that the contact locations on the anvil cluster around two places, while the pertaining trajectories diverge and partially overlap at the apex (contact points with the anvil are not explicitly shown but they can be inferred from the fewer data points on the trajectory that indicate fast movements). It is instructive to follow two nearby trajectories from the point of contact through the loop to see that adjacent trajectories can significantly diverge at the maximum position. The same location on the anvil can be hit by very different trajectories. This is especially true if one considers that the hammering is performed by a multi-joint arm which has more degrees of freedom than necessary with respect to the result, hitting the anvil with the hammer. The lesson learnt from this observation is that variability is partly afforded by the redundancy of the task. The schematic arm in the figure illustrates this with two possible arm configurations for the same endpoint of the hammer.

This observation suggests a first entry into the problem of understanding variability: The task performance can be parsed into *execution variables*, which would be the joint angles of the arm in the case of hammering, and *result variables*, which would be the position of the endeffector. If there are more execution variables than result variables, a set of identical solutions is available to the actor. This can be expressed in a quantitative fashion if the functional relation between the execution and result variables is known. A subset of combinations of execution variables will all map onto the same result variable. In reverse direction, if there is a desired result such as an accurate hit of a target on the anvil, the possible combinations of execution variables can be calculated. Over repeated actions all elements of the solution set can be performed without changing the result. Hence, a first and fundamental step for addressing the issue of variability is the distinction into features that characterize the execution process and features that capture the result of the movement. Then, the functional relationship between execution and result level should be established where possible.

Representation of the Problem: Execution and Result Space

To further develop this distinction, let's turn to the task of dart throwing that exemplifies and at the same time generalizes this issue. For the purpose of exposition the real-life skill is simplified such that the action is confined to two dimensions and the throwing arm is fixed in space and modeled as a single-joint lever arm (Fig. 2A). The critical moment that determines the entire dart trajectory and hitting success at the dart board is the moment of dart release, specifically the angular position of the lever (θ), or equivalently the dart, and the release velocity of the dart ($\dot{\theta}$). The *execution* of the task can therefore be sufficiently described by a two-dimensional vector $\mathbf{e} = (\theta, \dot{\theta})$. The *result* can be described in terms of the distance to the center of the target (d), which is

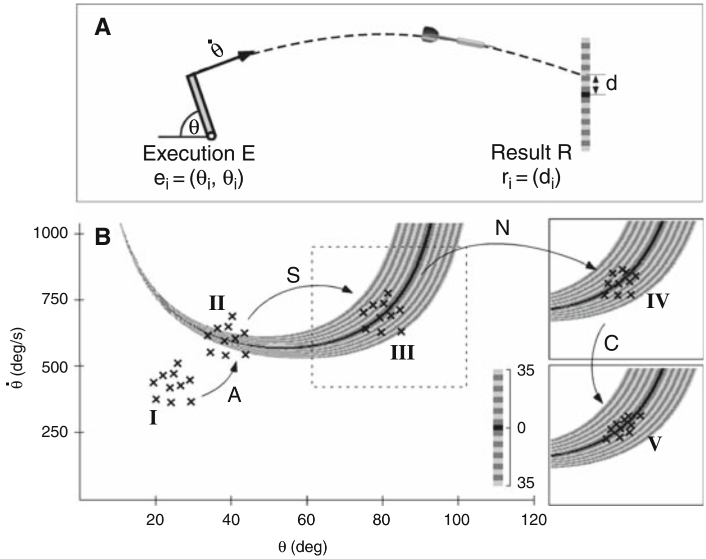

Fig. 2 The simplified dart throwing task. **A:** The action is confined to two dimensions of the sagittal plane. The position of release of the dart is expressed as the angle θ of the lever arm. At the moment of release the dart flies with a given velocity $\dot{\theta}$ in direction of the target. The success of the performance is expressed as the distance of the contact point to the center of the target *d*. **B:** Execution space with the solution manifold for the simplified dart throw. The *black line* denotes combinations of θ and $\dot{\theta}$ that achieve a hit with zero error in the center of the target ($d = 0$). The *grey shades* corespond to the rings around the bull's eye. White corresponds to those variable combinations that do not lead to a target hit. Five different hypothecial data sets are inserted which illustrate in successive fashion how the different components may contribute to improvements in performance

one-dimensional in this case, $\mathbf{r} = (d)$.[1] The task is therefore redundant in the minimal sense. If we consider a series of throws with *i* trials ($i = 1, 2, .. n$), there is a set **E** of two-dimensional execution vectors $\mathbf{E} = \{\mathbf{e}_1,\ \mathbf{e}_2,\ ..\ \mathbf{e}_n\}$ and a set **R** of one-dimensional result vectors $\mathbf{R} = \{\mathbf{r}_1,\ \mathbf{r}_2,\ ..\ \mathbf{r}_n\}$.

As repeated throws are not reproduced in identical fashion, successive trials will have variability in both the execution and the result variables, denoted by V(**E**) and V(**R**), respectively. The central question is now how result variability V(**R**) relates to execution variability V(**E**). This question has particular relevance in tasks where accuracy and the lack of variability in the result is the critical factor for successful performance. Is the variability seen over repeated performances only unwanted "noise" or is there some structure visible that potentially makes use of the redundancy of the task (Newell, Deutsch, Sosnoff, &

---

[1] Note that other researchers have referred to these two types of variables as action or performance variables on the one level versus error or task variables on the other level.

Mayer-Kress, 2006)? If a task is represented in terms of execution and result space, then variability over repeated performances can be analyzed further. In particular, we will show how the structure of variability can be decomposed into conceptually motivated components. Noise or a stochastic component is only one of the potential contributors to the dispersions seen in the output.

To begin, we model the task of interest by parsing it into result and execution variables with a functional relation such that every result $\mathbf{r}_i$ is fully determined by the execution variables: $r_i = f(e_i)$. Note that we ignore any measurement noise. For the dart throwing example this functional relation is illustrated in Fig. 2B. The two execution variables angle θ and velocity $\dot{\theta}$ span a two-dimensional space that contains all possible combinations of the two variables. This space will be referred to as *execution space*. Every throw with its two execution variables $(\theta, \dot{\theta})$ corresponds to a data point in this space. The clusters of data points in Fig. 2B, labeled by Roman numerals, refer to series of 10 throws each. All combinations of θ and $\dot{\theta}$ that lead to successful throws ($d = 0$ cm) form a subspace which will be called *solution space* or *solution manifold*. This subspace is shown as the dark line in Fig. 2B. It contains an infinite number of different combinations of execution variables, i.e., the task is redundant. In addition, the sets of solutions that lead to a given result with a constant deviation, for example $d = 3$, 6, or 9 cm, are shown as grey-shaded iso-error bands aligned with the solution manifold. Throws in the white area would not hit the dart board. The shown data points in sets I to V in Fig. 2B are fictive exemplary trial sequences.

In this representation phenomena may arise that seem non-intuitive at first sight: two data sets may have similar variability in execution space V(**E**), but may have very different means and dispersions in the result (compare sets II and III). Further, considerable variability in execution variables may lead to relatively little variability in the result (see set III). Furthermore, when practice and improvement in performance is of interest, the observed result **R**, e.g., expressed as average distance to target, may decrease while V(**E**) remains large (e.g., see sets IV and V). Given that changes in variability are at the heart of motor learning, a decomposition of variability will be proposed to shed light on these processes underlying motor learning.

Four Components Contributing to Performance Improvement

Once a task or performance is parsed into execution and result variables with redundancy across the two levels of description four conceptually different possibilities exist that can improve performance and decrease the variability in the result. Note what is described in the following is not restricted to the kinematic analysis of dart throwing, but can be applied to any task with a defined goal ranging from hammering to reaching with accuracy as the goal. Similarly postural control can be analyzed in this manner as many joints (execution level) are coordinated to obtain relative invariance of vertical orientation (result level). Note that both execution and result vectors may have more dimensions and the method can be applied in the same fashion. Similar distinctions between levels can also be made between a kinematic description of task success and redundant muscular space, or produced forces and contributing joint angles or muscles.

First Component: Approach

Figure 2B illustrates a fictive series of dart throws exemplifying possible changes in performance with practice. The data are grouped into five sets I to V, each of which contains 10 throws. When performing a new task, it is not uncommon that the initial set of trials is far away from successful performance as shown by set I. However, throughout the first trials the actor typically becomes familiar with the execution space and its mapping to results and soon finds more successful locations in the execution space, i.e., locations with combinations of variables that have zero or close-to zero error results. This putative first stage of learning is shown by the change from data set I to set II. In this data set, the mean value of the two data sets has moved onto the solution manifold. We will call this component "Approach" (A) as the data approach the solution manifold. It corresponds to what is often referred to as exploring possibilities for solutions in learning and developmental studies and will quantify such relatively large changes in execution space as often seen at initial stages of practice.

Second Component: Sensitivity

Locations on or near the manifold may differ due to how many successful solutions are adjacent. Note that the solution manifold of dart throwing is nonlinear and the widths of the iso-error bands are different for different locations on the manifold. In Fig. 2B the data sets II and III have the same dispersion, yet all throws of the series III land on the target while many throws in set II fail to hit the target. In order to reliably achieve a specific result, an actor should aim for such locations that are surrounded by a broad band of iso-error lines. Such areas of the execution space are more tolerant or less sensitive to noise in the execution variables. Finding a location in execution space with a sufficient safety margin will be quantified by the second component *Sensitivity* (S).

Third Component: Noise Reduction

In line with Bernstein's observation, many authors have subsequently corroborated that variability persists across repeated behaviors. This seemingly inevitable noise may be ascribed to lower levels of the hierarchical biological systems. Yet, good control and coordination is generally associated with smaller magnitudes of such fluctuations. In this sense, virtually every experiment on motor learning has interpreted decreases in variability over the course of practice as an indicator of learning. This decrease is generally captured by standard deviations of some performance variable, reflecting the underlying assumption that this variability is stochastic noise. Such "motor noise" has been shown to correlate with the signal magnitude, which is the essence of Weber's Law for both perception and performance (Harris & Wolpert, 1998). Figure 2B illustrates this option for performance improvement in the change from set III to set IV. While both data sets remain at the same location on the solution manifold, the magnitude of their dispersions is significantly different, leading to better results. This change in variability is referred as *Noise Reduction* (N).

Fourth Component: Covariation

A fourth and last possibility how variability in the result can decrease with practice is shown in Fig. 2B by comparing the data sets IV and V. Different from set IV the data in set V cluster in alignment with the direction of the solution manifold. The execution variables *co-vary* in a task-specific way in contrast to the data in set IV. In this case variability, or more precisely the deviations in the individual processes from the mean, show negative covariation but with a small nonlinear component. Deviations in both variables compensate for each other and more accuracy and invariance in the result is achieved. Such covariation was experimentally demonstrated in an early study by Stimpel (1933) who examined repeated throwing actions to a target and found that the dispersion in the result was smaller than expected from the dispersion measured in the execution variables (angle and velocity at release). An often cited study by Arutyunyan and colleagues showed for pistol shooting that variations in body and pistol angles compensated for each other to achieve a steady pointing position (Arutyunyan, Gurfinkel, & Mirskii, 1968, 1969). Several more recent studies corroborated support for this aspect in task performance (Cusumano & Cesari, 2006; Kudo et al., 2000; Martin et al., 2001).

The four components may be viewed as four qualitatively different kinds of interventions by the control system. To achieve changes in the component *Approach A* the average values of the execution variables must be shifted to a more successful location in execution space. Shifts of the mean along the solution manifold can further improve performance if a less sensitive location is reached (*Sensitivity S*). In contrast, the component *Noise Reduction N* does not imply a change in the mean variables but a reduction in the amount of dispersion. The component *Covariation C* summarizes all effects that arise from the task-specific modifications of the execution variables within trials.

Routes for the Quantification of these Components

How can the contribution of these components to performance be quantified? Before addressing this question, some considerations on the choice of variables should be discussed. Thus far, the discussion relied on a given description of the task where the execution and result variables were defined. Even though the variables chosen to describe dart throwing are reasonable and may appear unequivocal at first sight, there always exist alternative ways of description. For example, the execution of the throw may also be described by the acceleration-time profile or joint torques at the moment of release, or only in a different coordinate system, i.e., the release position and velocity in x-y coordinates of the sagittal plane. Besides such alternative biomechanical descriptors, one may also be interested in the muscular activation patterns of the arm movement and how they generate a kinematic or kinetic result. The operations of the present analysis method are applicable irrespective of the specific choice of coordinate systems and levels of description. Note, though, that choosing alternative variables may influence the results, as will be discussed below. Hence, the choice of the variables and the levels of description needs to be conceptually motivated or several descriptions may be contrasted (Müller, Frank, & Sternad, 2007; Smeets, 2000; Smeets & Louw, 2007).

Having identified the variables the next step is to clarify how the four conceptual components can be quantified in data to differentiate our understanding of performance or improvement in performance. To this end the variability in execution, V(**E**), will be compared with the result, **R**. As the variables for execution and result may differ in number and their dimensions, such comparison is not straightforward. Returning to the hammering example, changes in joint angles cannot be directy related to the positional accuracy of the hit. Similarly, variations in throwing angle and velocity are not simply proportional to the hitting accuracy. To solve this problem two principally different approaches are possible. First, the analysis may start by examining **E** and *projecting its effect* into the result level **R**. To this end, the functional relationship between **E** and **R** must be known: $\mathbf{R} = f(\mathbf{E})$. The evaluation will be made in the units of the result variable. We will refer to this as the result-based method. Second, it is possible to begin with the result and *project its effect back* into the execution level and subsequently evaluate the variability within the metric of the execution variables. Again, the functional relationship f must be known. The latter strategy will be discussed as an execution-based approach.

We will begin by presenting a quantification method that takes the first approach, the TNC-method (Tolerance, Noise and Covariation) by Müller and Sternad (Müller et al., 2007; Müller & Sternad, 2003, 2004a, 2004b). Subsequently, we will overview two other methods that take the execution-based approach, the UCM-analysis (UnControlled Manifold) by Schöner and colleagues and the GEM-analysis (Goal-Equivalent Manifold) by Cusumano and Cesari (Cusumano & Cesari, 2006; Scholz & Schöner, 1999). A discussion of the relative advantages and shortcomings of the three methods will follow.

Result-Based Analysis of Variability: The TNC-Method

For didactic purposes the exposition of the basic calculation steps of the TNC-method will begin with the calculation of covariation C. Using the simplified dart throwing task again as example, assume that a person performs 50 trials such that the sets of execution and result variables are $\mathbf{E} = \{(\theta_1, \dot{\theta}_1)(\theta_2, \dot{\theta}_2), ..(\theta_{50}, \dot{\theta}_{50})\}$ and $\mathbf{R} = \{d_1,\ d_2, ..\ d_{50}\}$. In order to evaluate the contribution of C in this data set a reference set is created that characterizes performance where such covariation is absent. This is achieved by randomly permutating $\mathbf{E}$ to generate a new surrogate data set $\mathbf{E}_C^0$: Keeping the values for the individual variables θ_i and $\dot{\theta}_i$ the same, their pairing in each $\mathbf{e}_\mathrm{i}$ is randomized, e.g., $\mathbf{E}_C^0 = \{(\theta_7, \dot{\theta}_{24})(\theta_{45}, \dot{\theta}_{32}), ..(\theta_{28}, \dot{\theta}_9)\}$. If the functional relationship f between $\mathbf{E}$ and $\mathbf{R}$ is known, then the results for $\mathbf{E}_C^0$ can be calculated by $\mathbf{R}_C^0 = f(\mathbf{E}_C^0)$ to obtain $\mathbf{R}_C^0 = \{d_1^*,\ d_2^*, ..\ d_{50}^*\}$. Next, the result is summarized into a single performance measure $p(\mathbf{R}_C^0)$, for example by calculating the mean distance to target over the 50 trials. Finally, the result measures of the real and the permuted data are compared to quantify the contribution of covariation ΔC. If for example the result of $\mathbf{E}$ is summarized by the mean distance of 3 cm ($p(\mathbf{R})$), and the result of $\mathbf{E}_C^0$ is 7 cm ($p(\mathbf{R}_C^0)$), their performance difference of 4 cm is attributed to covariation and called ΔC.

The same basic operations can be applied to estimate all other components. Importantly, though, the first step that creates a reference execution set that eliminates one component, $\mathbf{E}_A^0$, $\mathbf{E}_S^0$, $\mathbf{E}_N^0$ and $\mathbf{E}_C^0$, has to be implemented in different ways. The second step that calculates the respective result data sets $\mathbf{R}_A^0$, $\mathbf{R}_S^0$, $\mathbf{R}_N^0$ or $\mathbf{R}_C^0$ using the functional relation f is the same for all components. The third step that calculates a respective performance measure p of the real and reference data can take different realizations, although they should be the same to keep the components comparable. Besides the average of $\mathbf{R}$, as in the example above, one can also take the variance of $\mathbf{R}$ or any other suitable summary measure. Finally, the contribution of a single component can now be quantified by subtracting the real results with the results of the respective reference data. For example, the contributions of A, S, or N are then calculated as: $\Delta \mathrm{A} = p(\mathbf{R}_A^0) - p(\mathbf{R})$, $\Delta \mathrm{S} = p(\mathbf{R}_S^0) - p(\mathbf{R})$, or $\Delta \mathrm{N} = p(\mathbf{R}_N^0) - p(\mathbf{R})$, respectively. If all components are quantified in the same manner in result space, it is then possible to compare the relative contributions of the different components.

The TNC-method for Learning. The basic steps of this decomposition of variability can be implemented in different ways, depending on whether one aims to assess variability in a given data set or changes across data sets. In a study on motor learning, Müller and Sternad (2004a,b) developed a quantification that was tailored to examine performance changes over practice. Therefore, the reference sets chosen for comparison were the initial trials rather than creating any "virtual" sets by an elimination procedure. Further, the goal was to decompose the total change in the result into the sum of the individual components, the components *Approach A* and *Sensitivity S* were combined into one component called *Tolerance T*. As both *A* and *S* are defined as a shift in location in execution space and therefore, implicitly, require a change in the mean values of the execution variables, they may imply the same intervention of the actor. Hence, the combined component was termed *Tolerance T*, expressing tolerance to error (for a detailed exposition see Müller and Sternad, 2004b).

Once more returning to Fig. 2B asssume the data set I contains the trials at the beginning of practice and set V are the trials at the end of practice. Using the permutation method described above, ΔC is computed as the difference between the original and the permutated data for a given data set. This is done for both the original set I (ΔC_I) and the final data set V(ΔC_V). The overall contribution ΔC throughout practice is then the difference between ΔC_V and ΔC_I. In the simplified case that ΔC is zero in set I, ΔC in the final set is the overall contribution of Covariation. Note that here the comparison is made betweeen the final and the initial data. The assessment of *Tolerance T* follows a similar logic. To quantify ΔT, the mean location of set I is shifted to the location of set III, such that they only differ in their location. To this end, the means of the execution variables are shifted in space, leaving the distribution of the values unaffected. Comparing result measures in sets I and III, the effect of location/*Tolerance* ΔT can be evaluated. Finally, sets III and IV only differ in their dispersion or noise. If set III is used as the reference for set IV, then the difference in the result measures renders ΔN. Using this particular implementation, the change in performance from throughout practice $\Delta p(\mathbf{R})$ can be represented as the sum of the three components:

$$\Delta p(\mathbf{R}) = \Delta \mathrm{C} + \Delta \mathrm{T} + \Delta \mathrm{N}.$$

Exemplary Results of Virtual Dart Throwing. To illustrate the method above we show some real data from a dart throwing task where the subject's arm was positioned on a lever arm as shown in Fig. 2A. As the dart only existed as a virtual dart on a virtual projection screen, the execution and the dart trajectory were strictly confined to the sagittal plane and maximal compatibility was achieved between the model and the task. Figure 3 shows the data of one participant performing the regular overhand throw (top row) and an underhand throw (bottom row). The participant practiced both types of performance in four blocks of practice with 80 throws each. The first two panels on each row show the blocks 1 and 4, respectively. The panels on the right show the change in the result measure *d* over the four blocks, depicted by the thick solid lines. The shaded areas depict the contributions of the three components in cumulative fashion. For a given block the relative contribution of *T*, *N* and *C* is indicated by the grey shades.

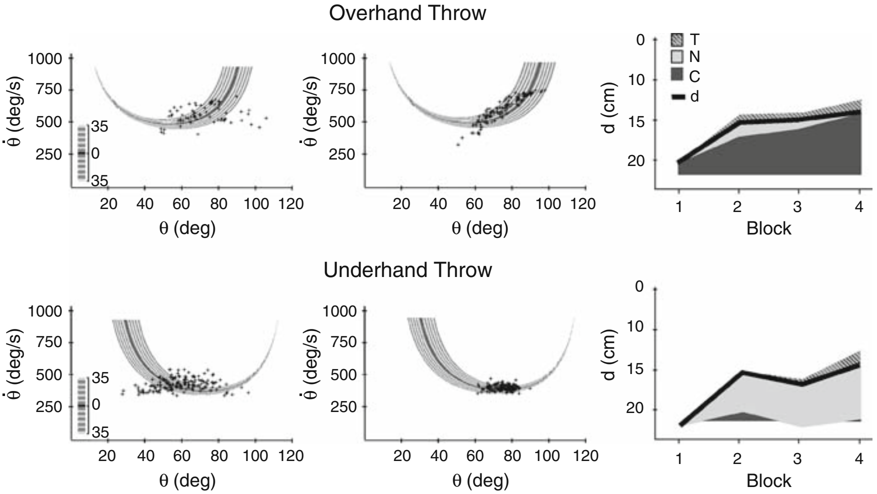

Fig. 3 Results from a single participant practicing an overhand and an underhand throw in four blocks. While both throwing tasks show improvement in the result as documented by the change in the result variable *d* on the *right panels*, the types of changes in execution differ considerably. The *left panels* on both rows show the initial performance, the *center panels* show the final performance. More explanations see text

For both throwing strategies there is a performance improvement from an average error of approximately 20 cm to 15 cm, as can be seen by the change in *d* on the right panels. Considering only the time course of improvement in *d* there are no noticeable differences in the two tasks. However, when looking at the clustering of the data in blocks 1 and 4, very obvious differences are visible. The final performance in the overhand throw shows considerable scatter, even though the scatter is largely aligned with the solution manifold. Conversely, the final performance in the underhand throw shows a tight clustering of the individual trials, again aligned with the solution manifold however at a location where the iso-error bands are relatively thin. These evident differences in performance are mirrored in very different contributions of *T*, *N*, and *C* in the two strategies. In the underhand throw it is predominantly *N* that is responsible for the good performance in block 4. In contrast, for the overhand throw it is *C* that takes the lion's share in accounting for the improved performance. *T* is negative in both conditions. In the underhand throw, the data in block 4 moved towards the right where the solution manifold is more sensitive to error. In the overhand throw the mean of the data cloud is shifted away from the solution manifold, where it is more likely to miss the target.

In sum, given that changes in variability of the execution are evaluated in result measures, e.g. distance to target, direct comparisons between components and between different strategies are possible. All features of performance are expressed in the common metric of the result measure, central to the result-based approach.

Execution-Based Analysis of Variability

A conceptually different route is to analyze the contribution of the four components in the metric of the execution variables. Two analysis methods have taken this approach: the "UnControlled Manifold" analysis (UCM) by Schöner and colleagues and the "Goal Equivalent Manifold" analysis (GEM) by Cusumano and Cesari (Cusumano & Cesari, 2006; Scholz & Schöner, 1999). Both studies developed their method in a task where a multi-joint limb points to a target and the joint angles have more degrees of freedom than the dimensions of the result, similar to the initial hammering example. The question is whether the variations in the redundant joints of the limb lead to corresponding deviations in the pointing error or whether variability has little or no effect on the constancy of the pointing. This separation of variables into joint angles and endpoint accuracy is identical to the parsing of the task in execution and result space highlighted above. However, the actual operations to analyze variability differ significantly. Both methods are grounded in the identification of the nullspace of the Jacobian, which is a well-established calculation procedure for the control of kinematic redundancy in robotics (Craig, 1986; Liégeois, 1977; Mussa-Ivaldi & Hogan, 1991).

The UCM-Method. To quantify the effect of deviations in the execution variables with respect to deviations in the result variable, a specific location in the execution space must be chosen. This is typically the mean value of the execution variables ($\mathbf{E}_M$). Assuming the result **R** is accuracy and the functional relation $\mathbf{R} = f(\mathbf{E})$ is known, the function f is linearized in $\mathbf{E}_M$. When calculating the Jacobian at this point one obtains the nullspace, i.e., the subspace within which the result remains unchanged (for an introduction see for example Craig, 1986). Conversely, all deviations in directions orthogonal to this nullspace affect the result. Scholz et al. (2000) hypothesized that execution variability is reduced in those dimensions that are sensitive to deviations, i.e., where the result is less tolerant to deviations. It is further hypothesized that such task-specific compression of variability indicates control (hence the name uncontrolled manifold). Therefore, in skilled performance variability in the task-relevant directions of the execution space should be smaller compared to directions which have only little effect on the result. The UCM-method proceeds to compare the range of variability in the two directions by taking a ratio. If this ratio is greater than 1, control is indicated (Latash, Scholz, Danion, & Schöner, 2002). Note that this comparison of variability in different directions is made within the execution space, in contrast to the result-based comparisons of the method of Müller and Sternad.

In the first step where a reference point $\mathbf{E}_\mathrm{M}$ is determined, the UCM approach typically choses the mean of the data. It is implied that $\mathbf{E}_\mathrm{M}$ is on the solution manifold, which however need not be the case. Hence, the distance to the solution manifold, estimated in the component *A* is not considered. The nullspace corresponds to a linear combination of execution variables by which the given result is achieved. Typically, neither the nullspace nor the orthogonal space are parallel to any of the axes of the execution space due to covariation and different ranges of variation between the execution variables. Hence, the operations of the UCM-analysis do not differentiate between covariation *C* and noise *N*. The component Sensitivity is also embedded in the same calculations. The calculations of nullspace are the result of taking partial derivatives with the Jacobian. The direction with the lowest sensitivity to error is the nullspace, the direction of the highest sensitivity to error is the space orthogonal to the nullspace.

A method closely related to UCM is principal component analysis (PCA), a statistical analysis procedure which shares the same basic goal of discovering predominant orientations in the variance of data in a potentially high-dimensional space. The principal components are those directions in this space in which variance is maximal, however without assuming a principal direction such as the nullspace. Hence, PCA can always be applied in an exploratory sense to any data set, as it does not require a prior model or hypothesis about relevant directions. Once principal components have been found, however, interpreting their significance remains a challenge as principal components may have no physical meaning or dimensions. The UCM approach, by contrast, links the identification of structured variance to an interpretable task-related variable. Note that besides PCA, numerous other methods to find structure in high-dimensional data have been developed and are under vigorous development, e.g., independent component analysis, information bottleneck analysis, non-negative matrix factorization and many others.

The GEM-Analysis. The GEM-approach by Cusumano and Cesari is closely related to the UCM-analysis but also offers some extensions. Identical to the UCM-approach, the so-called goal-equivalent manifold, GEM, is the nullspace that is derived at $\mathbf{E}_\mathrm{M}$ from the Jacobian of the linearized approximation of the function *f*. The complementary space is the so-called goal-relevant subspace, GRS, which is spanned by all dimensions that will negatively affect the outcome. Despite these parallel entries into the problem, the GEM approach goes further by differentiating the result of the nullspace analysis into three components: $\mathrm{V}(\mathbf{R})$ is described as the product of (*i*) the variability in execution $\mathrm{V}(\mathbf{E})$, (*ii*) the goal-relevant fraction of variability, and (*iii*) the goal-relevant sensitivity matrix (for details see Cusumano & Cesari, 2006). The sensitivity matrix contains amplification factors that express how strongly variability in GRS affects the result. This estimate corresponds to the component *Sensitivity S* in the four-component decomposition described above. The second component, the goal-relevant fraction of variability, is a factor that expresses which portion of the execution variability lies in GRS and thereby quantifies how well V(**E**) is aligned with the nullspace. If this factor is large, there is little alignment with the nullspace. While in the UCM-method the variability in the orthogonal direction is expressed relative to the variability in the nullspace, the GEM-method sets it relative to the total variance V(**E**). This quantification in execution space again comprises both the components N and C. In sum, both methods are based on a comparison of different fractions of variability within the execution space and the quantifications are based on linearizations of the functional relation f.

Discussion

Motor learning is tightly associated with changes in variability and accuracy in performance. Hence, the observed variability in behavioral measures is a suitable window for understanding motor learning. One route into this problem is the two-layered approach, including the task level and some subordinate level in which variability is observed. Several methods have been developed to describe, quantify and infer deeper insights from this representation of the problem. We conclude this exposition by addressing some critical issues as well as advantages and limitations that some of the reviewed approaches have.

Choice of Variables, Units and Metrics

As in almost all research, the choice of variables and coordinates is a deep issue upon which much of the results depend on. For the variability analyses presented here it is necessary to point out that it is by no means guaranteed that different choices of variables lead to identical results. Rather, it can be expected that dependent on the choice of parsing the levels of description different results in the variability decomposition will be obtained. How pervasive and fundamental this issue is will be illustrated by looking at a very simple example, the consistency of a sequence of steps to strides: Assuming the task is to traverse a given distance with a sequence of four steps (S_1 to S_4) the total traversed distance should be constant. Alternatively, the task may be described by two double steps: $DS_1 = S_1 + S_2$ und $DS_2 = S_3 + S_4$. The resulting movement distance can then be described in two ways: Distance $= S_1 + S_2 + S_3 + S_4 =$ constant (task description 1) or as Distance $= DS_1 + DS_2 =$ constant (task description 2). If we assume that S_1 and S_2 vary from trial to trial but covary perfectly, i.e., $DS_1 = S_1 + S_2 =$ Distance/2 for all repetitions and $DS_2 = S_3 + S_4 =$ Distance/2, the task is achieved perfectly. In task description 1 this would be the consequence of covariation C, while the success in task description 2 would be due to the absence of noise N. Both quantifications are correct within their framework of description. The principal difference lies in the fact that different elements are considered as relevant for control. This example highlights that the choice of variables is a core step in the analysis that requires care and motivation (Müller et al., 2007; Smeets, 2000; Smeets & Louw, 2007).

With the choice of variables comes another problem. It cannot always be taken for granted that the chosen execution variables have the same units. In the dart throwing example the execution space is spanned by angular position in degrees and velocity in angular velocity. How can dispersions in different units be compared? In the execution space of Fig. 3, for example, the nullspace close to a release angle of 60 degrees is parallel to the x-axis. In this case variability in the release angle can be large as it plays no role for the result; instead, variability in the release velocity orthogonal to the nullspace need to be small. From the perspective of the UCM-approach, control would be inferred because variability in the release velocity is numerically smaller than the variability in the release angle. However, such a comparison is impossible as these two variables have incommensurate units and their ranges are not comparable. Both UCM- and GEM-analysis are based on a geometric partitioning of the variability in execution space and rely on vector spaces with a homogeneous metric. Normalization of the range of variability on each dimension is clearly not an answer to the problem as it destroys the structure that is of interest. Transformations of variables and units may help but are not trivial. And yet, in many cases it is impossible to define relevant execution variables with a single metric. The TNC-method and the result-based approach in general is not limited by this problem as the contribution of components is projected into result space.

Nonlinearity and Differentiability of the Function between Execution and Result

In all three approaches the result variables must be known in their functional relationship to the execution variables. If f is not known analytically, then a (linear) estimate can be obtained by regression analysis of the data, as for example applied to the pointing task by Cusumano and Cesari (2006). This route, however, is not as satisfatory as the decomposition methods, in particular the TNC-method, extract many interesting aspects from the nonlinearities of the solution manifold. However, the function f can be established for many movement situations as biomechanical models can be developed in like fashion as for the multi-joint arm in the hammering example. For the analysis of the sway variability in postural control, for example, a rigid body model of the standing position could be used and the center of mass and its projection onto the support surface can be calculated by a function f.

One constraint for both the UCM- and GEM-approach is that this function f must be at least locally differentiable so that the Jacobian can be calculated. If this holds, then the well-developed tools of linear algebra can be applied.

However, the dart throwing example illustrates that this function is not always linear and may not even be differentiable. Further, if the result measure is discontinuous, as for example when 10 points are assigned for the bull's eye, 9 for the inner ring, etc., this constitutes a step-function which may not be differentiable for a given argument, i.e., the border between two rings, or has the value zero in all partial derivatives at all locations. Given that the TNC-method does not use linear algebra tools that are based on a metric in the execution space but takes a statistical approach it can be applied to all functions f, including nonlinear and non-monotonic functions with multiple variables. Similarly, the function may be very nonlinear such that the solution manifold is very nonlinear. This is for example the case for another throwing task, called skittles, introduced by Müller and Sternad (2004b) where the ball trajectory emulates a pendular swing. Depending on the target positions, the solution manifold can take a U-shape. Linearization as inherent to calculations of the nullspace may be problematic in such cases.

Choice of a Reference Value

All methods aim to quantify the contribution of variability components to performance. When assigning values to these components a reference should exist with zero denoting the absence of such contribution. In the UCM- and GEM-approach suitable reference values exist. For example, sensitivity is described a scalar value in the GEM-approach. UCM renders no sensitivity measure, but the amount of task specific compression is captured by a ratio, with a value of 1 representing no compression. The TNC-approach as it was summarized above in its application to learning does not provide such absolute interpretable values as the contributions are calculated relative to the initial performance. The result values are consequently relative and zero does not mean absence of tolerance but rather no change. On the other hand, the different components are evaluated in the same result units and can therefore be compared, for example N can be compared to S in their relevance to the improvement in the result. On the other hand, similar comparisons can be performed with virtual reference sets such that absolute references are given (Cohen & Sternad, 2008).

Outlook

In motor learning variability decreases with practice in almost all measures of task performance. However, this observation is so commonplace that it no longer provides information about the learning process. Further, variability is also a characteristic of all behavior, even in the most skilled and seemingly automated performance. Why does variability never go to zero even in healthy individuals? Is it simply “motor noise” that cannot be suppressed completely? We presented an approach to uncover structure in sets of data to shed light on the system processes underlying control and learning. Based on a representation of the system in two levels we identified four conceptually distinct ways how improvement in performance can be obtained due to changes in variability in execution. We presented and contrasted three methods that have been discussed in the motor control literature. Such methods are promising but also have limitations that need to be recognized before accepting their results. More research will further develop these routes and, hopefully, present more insights about the processes underlying control and acquisition of motor skills.

Acknowledgments This research was supported by grants from the National Science Foundation BCS-0450218, the National Institutes of Health R01 HD045639, and the Office of Naval Research N00014-05-1-0844.

References

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. (1967). *The coordination and regulation of movement*. London: Pergamon Press.

Cohen, R., & Sternad, D. (2008). Variability in matorlearning: Relocating, channeling and Reducing Noise. *Experimental Brain Research*.

Craig, J. J. (1986). *Introduction to robotics*. Reading, MA: Addison-Wesley.

*Cusumano, J. P., & Cesari, P. (2006). Body-goal variability mapping in an aiming task. *Biological Cybernetics, 94*(5), 367–379.

Davids, K., Bennett, S., & Newell, K. M. (2006). *Movement system variability*. Champaign, IL: Human Kinetics.

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

Kudo, K., Tsutsui, S., Ishikura, T., Ito, T., & Yamamoto, Y. (2000). Compensatory correlation of release parameters in a throwing task. *Journal of Motor Behavior*, 32(4), 337–345.

*Latash, M. L., Scholz, J. P., Danion, F., & Schöner, G. (2002). Motor control strategies revealed in the structure of variability. *Exercise and Sport Sciences Reviews, 30*, 26–31.

Liégeois, A. (1977). Automatic supervisory control of the configuration and behavior of multibody mechanisms. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-7*(12), 868–871.

Martin, T. A., Gregor, B. E., Norris, N. A., & Thatch, W. T. (2001). Throwing accuracy in the vertical direction during prism adaptation: not simply timing of ball release. *Journal of Neurophysiology*, 85, 2298–2302.

*Müller, H., Frank, T., & Sternad, D. (2007). Variability, covariation and invariance with respect to coordinate systems in motor control. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 250–255.

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, 39*, 22–33.

Müller, H., & Sternad, D. (2004a). Accuracy and variability in goal-oriented movements: decomposition of gender differences in children. *Journal of Human Kinetics, 12*, 31–50.

*Müller, H., & Sternad, D. (2004b). Decomposition of variability in the execution of goal-oriented tasks – Three components of skill improvement. *Journal of Experimental Psychology: Human Perception and Performance, 30*(1), 212–233.

Mussa-Ivaldi, F. A., & Hogan, N. (1991). Integrable solutions of kinematic redundancy via impedance control. *International Journal of Robotics Research, 10*(5), 481–491.

Newell, K. M., & Corcos, D. M. (1993). *Variability and motor control.* Champaign, IL: Human Kinetics.

Newell, K. M., Deutsch, K. M., Sosnoff, J. J., & Mayer-Kress, G. (2006). Variability in motor output as noise: A default and erroneous proposition? In K. D. Davis, S. Bennett & K. M. Newell (Eds.), *Movement system variability* (pp. 3–22). Champaign, IL: Human Kinetics.

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.

Smeets, J. B. J. (2000). The relation between movement parameters and learning. *Experimental Brain Research, 132*(4), 550–552.

Smeets, J. B. J., & Louw, S. (2007). The contribution of covariation to skill improvement is an ambiguous measure. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 246–249.

Stimpel, E. (1933). Der Wurf. In F. Krüger & O. Klemm (Eds.), *Motorik* (Vol. 9, pp. 109–138). München: Beck.

Vaillancourt, D. E., & Newell, K. M. (2001). Woodworth (1899): Movement variability and theories of motor control. In M. L. Latash & V. M. Zatsiorsky (Eds.), *Classics in Movement Science* (pp. 409–435). Champaign, IL: Human Kinetics.

Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychological Review Monograph Supplements, 3*(3), 1–119.

The references marked with an asterisk (*) are specifically recommended for further introduction or background to the topic.