Contributions of Online Visual Feedback to the Learning and Generalization of Novel Finger Coordination Patterns

Xiaolin Liu1 and Robert A. Scheidt1,2

1Department of Biomedical Engineering, Marquette University, Milwaukee, Wisconsin; and 2Department of Physical Medicine and Rehabilitation, Northwestern University Feinberg School of Medicine, Chicago, Illinois

Submitted 20 September 2007; accepted in final form 13 March 2008

Liu X, Scheidt RA. Contributions of online visual feedback to the learning and generalization of novel finger coordination patterns. J Neurophysiol 99: 2546–2557, 2008. First published March 19, 2008; doi:10.1152/jn.01044.2007. We explored how people learn new ways to move objects through space using neuromuscular control signals having more degrees of freedom than needed to unambiguously specify object location. Subjects wore an instrumented glove that recorded finger motions. A linear transformation matrix projected joint angle signals (a high-dimensional control vector) onto a two-dimensional cursor position on a video monitor. We assessed how visual information influences learning and generalization of novel finger coordination patterns as subjects practiced using hand gestures to manipulate cursor location. Three groups of test subjects practiced moving a visible cursor between different sets of screen targets. The hand-to-screen transformation was designed such that the different sets of targets (which we called implicit spatial cues) varied in how informative they were about the gestures to be learned. A separate control group practiced gesturing with explicit cues (pictures of desired gestures) without ongoing cursor feedback. Another control group received implicit spatial cueing and feedback only of final cursor position. We found that test subjects and subjects provided with explicit cues could learn to produce desired gestures, although training efficacy decreased as the amount of task-relevant feedback decreased. Although both control groups learned to associate screen targets with specific gestures, only subjects provided with online feedback of cursor motion learned to generalize in a manner consistent with the internal representation of an inverse hand-to-screen mapping. These findings suggest that spatial learning and generalization require dynamic feedback of object motion in response to control signal changes; static information regarding geometric relationships between controller and endpoint configurations does not suffice.

INTRODUCTION

The ability to learn arbitrary visuomotor transformations undoubtedly conferred evolutionary advantage to those adept at controlling the endpoint of a long stick while keeping a hungry wolf at bay. It also is useful for modern survival skills such as controlling the motion of a computer cursor while composing a research proposal. In each case, the position of the controlled endpoint is determined by coordinated activity of a large number of motor control variables (e.g., motor unit activities, multiple muscle contractions, and/or multijoint torques). Proficiency in manipulation requires the brain to learn a mapping between endpoint motion and changes in control variables: Because sticks can be different lengths and mouse sensitivities vary from computer to computer, such mappings must be adaptable.

Except in special cases, visual feedback provides the brain with the most reliable information of how objects move through space, whereas other sensory modalities provide ambiguous or otherwise limited feedback. Despite numerous studies exploring how both healthy and neurologically impaired subjects adapt to prism glasses (e.g., Held and Freedman 1963; Kitazawa et al. 1997; Redding and Wallace 1988, 1992, 1996; Rosetti et al. 1998) or rotation of cursor motion on a computer screen (cf. Krakauer et al. 2000; Tod et al. 2001; Wigmore et al. 2002), our understanding of how visual information facilitates learning of arbitrary mappings between high-dimensional control variables and low-dimensional endpoint motion is incomplete. This is due, in part, to the practical difficulty of establishing truly novel transformations; most experimental studies of visuomotor adaptation simply modify overlearned mappings through displacement, rotation, and/or dilation. The learning so characterized is incremental, not novel, because the fundamental motor control problem has already been characterized through years of practice in moving handheld objects through space.

Recently, Mosier and colleagues (2005) developed an experimental approach that circumvents this limitation. Subjects wore a data glove instrumented with a large number of bend sensors. Signals from these sensors drove motion of a cursor on a computer screen via an arbitrary linear projection (a “hand-to-screen mapping”). Under this mapping, each hand configuration (the articulation space) projected onto one and only one screen location (the endpoint space), although each screen location could be achieved using an infinite number of different hand configurations. Although the task was unlike anything subjects had experienced before (and thus they started learning it with a “blank slate”), the task was similar to all other motor tasks in that a highly redundant set of control signals defined the motion of a controlled endpoint in a low-dimensional workspace.

Mosier et al. (2005) found that by practicing the task, nearly all subjects reduced cursor positioning errors within a single experimental session lasting 1 to 2 h. A subject group deprived of visual feedback of cursor motion during target acquisition also learned, but did so more slowly than a group provided with such feedback. In both cases, learning generalized to unpracticed screen targets in trials also performed entirely without visual feedback. This ability to generalize indicated

Address for reprint requests and other correspondence: R. A. Scheidt, Department of Biomedical Engineering, Olin Engineering Center, 303, P.O. Box 1881, Marquette University, Milwaukee, WI 53201-1881 (E-mail: scheidt@ecee.org).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
that subjects had learned how information about the relative distance between screen targets could be used to constrain their choice of hand postures when presented with new targets. One could say that subjects learned a “motor representation” of the Euclidean space on which finger movements were remapped. Furthermore, the ability to generalize developed in parallel to a reduction of motion in those degrees of freedom that did not contribute to cursor motion, suggesting that the newly formed representation of endpoint space specifically constrained the control variables so as to reduce variability in degrees of freedom not contributing to task performance.

Because their subject group without ongoing visual feedback also learned to generalize, Mosier and colleagues suggested that ongoing visual feedback is not necessary for learning novel hand-to-screen mappings. This conclusion was confounded, however, by the fact that all subjects were provided visual feedback of cursor motion during error-correction maneuvers performed after each target acquisition attempt. Online feedback correction for misalignment between sensory reference frames can lead to adaptive realignment of those frames (Redding and Wallace 1996; cf. Wann and Ibrahim 1992) and, indeed, such feedback could have facilitated learning of Mosier’s novel spatial transformation in subjects otherwise deprived of visual feedback of cursor motion.

Using a similar experimental setup, we addressed two questions left unresolved by Mosier et al. (2005). First, we sought to resolve the confound by asking what visual feedback conditions are necessary for subjects to learn the spatial transformations that allow them to generalize beyond the set of training targets. Specifically, we examined whether the learning of a novel spatial relationship between highly redundant control signals and a low-dimensional space of endpoint motion might indeed depend on concurrent visual feedback of cursor motion during target acquisition. Next, we asked whether training in the reduced-dimensional endpoint space can be as effective as explicit cueing of the desired gestures themselves in training novel patterns of finger coordination. This last question was motivated by our aim to determine whether such training might have utility for promoting the acquisition of novel finger coordination patterns.

Three test groups were required to move a cursor between different sets of screen targets corresponding to different articulation space postures under a common hand-to-screen mapping. The mapping was designed so that the different target sets varied in how informative they were about the gestures to be learned. For one group, nearly all of the variance among gestures was captured by only seven independent degrees of freedom (Mosier et al. 2005). Indeed, such feedback could have facilitated learning of Mosier’s novel spatial transformation in subjects otherwise deprived of visual feedback of cursor motion.

Methods

Twenty-five adult subjects and one skilled instructor of the American Manual Alphabet (AMA) participated in this study after providing written informed consent approved by Marquette University’s Institutional Review Board. All subjects (15 females, 10 males; 23 ± 3 yr of age) self-reported to be right-hand dominant and naïve to sign language, the AMA, and to the purposes of this study. Each wore a right-handed CyberGlove (Immersion, San Jose, CA) from which 19 joint angle measurements were recorded from flexion of the phalangeal joints (proximal, middle, and distal), abduction of the thumb and fingers, and wrist flexion/extension and abduction/adduction. Glove signals were sampled at a rate of 50/ s. During the main experimental session (session 2), the glove’s 19-dimensional (19-D) signal vector was mapped onto the two-dimensional (x, y) coordinates of a computer screen using a linear transformation

\[
\begin{bmatrix}
  x \\
  y
\end{bmatrix} = \begin{bmatrix}
  a_{x,1} & a_{x,2} & \cdots & a_{x,19} \\
  a_{y,1} & a_{y,2} & \cdots & a_{y,19}
\end{bmatrix} \begin{bmatrix}
  h_1 \\
  h_2 \\
  \vdots \\
  h_{19}
\end{bmatrix} = Ah
\]

where \( p = [x, y]^T \) indicates the cursor location on the monitor (the endpoint space), \( h = [h_1, h_2, \ldots, h_{19}]^T \) is the glove signal vector (the articulation space), and \( A \) is a matrix of mapping coefficients \( [a_{i,j}] \) (the hand-to-screen mapping). Articulation degrees of freedom not contributing to endpoint motions constitute the “null space” of projection matrix \( A \), computed as \( h_N \) (Mosier et al. 2005)

\[
h_N = (I_{19} - A^+A)h
\]

where \( I_{19} \) is the 19-D identity matrix and \( A^+ = A^T(A \times A^T)^{-1} \) is the Moore–Penrose pseudoinverse of \( A \).

Experiments were conducted in two sessions performed on separate days. Session 1 collected calibration data for computing subject-specific hand-to-screen mappings that were used in session 2. All subjects used the same calibration procedure, which was based on the observation that during many object manipulation tasks (Santello et al. 1998, 2002), including finger spelling in the AMA (Jerde et al. 2003b), the vast majority of kinematic variance between static hand postures can be captured by a small number of independent principal components (PCs). Finger spelling is a skill to which most individuals are naïve and it is well suited for studying how the brain learns novel motor coordination patterns in that its kinematic demands have been well characterized (Jerde et al. 2003a,b). Although >90% of the variation in finger configurations for the static AMA characters can be captured by only seven independent degrees of freedom (Jerde et al. 2003a), our pilot studies found that in using only two PCs to define \( A \), we could titrate the amount of task-relevant visual information provided to subjects as they learned to form distinct subsets of the AMA characters (Liu et al. 2005).

Defining the hand-to-screen mapping and the three training target sets

We now describe how we defined the hand-to-screen mapping and our rationale for selecting specific subsets of AMA characters over which the three groups of test subjects trained. A skilled instructor of the AMA wore the data glove and performed five repetitions of a pseudorandom sequence of the 24 static AMA characters (all but J and Z, which require hand motion). She held each character steady as the glove signals were sampled. Finger kinematics were averaged across repetitions for each character. A data matrix \( H \) was constructed by

\[
H = \begin{bmatrix}
  Z_{1,1} & Z_{1,2} & \cdots & Z_{1,N} \\
  Z_{2,1} & Z_{2,2} & \cdots & Z_{2,N} \\
  \vdots & \vdots & \ddots & \vdots \\
  Z_{M,1} & Z_{M,2} & \cdots & Z_{M,N}
\end{bmatrix}
\]

where \( Z_{m,n} \) is the joint angle at frame \( n \) for the \( m \)-th training target. We could say that subjects learned a “motor representation” of the new hand postures represented by the data glove. This representation of endpoint space specifically constrained the control variables so as to reduce variability in degrees of freedom not contributing to task performance.

Furthermore, the ability to generalize developed in parallel to a reduction of motion in those degrees of freedom that did not contribute to cursor motion, suggesting that the newly formed representation of endpoint space specifically constrained the control variables so as to reduce variability in degrees of freedom not contributing to task performance.

We further hypothesized that feedback of cursor motion is necessary to promote generalization beyond the training set to novel endpoint targets. Portions of this work were previously presented in abstract form (Liu et al. 2005).
appending, side by side, the column vectors of average signals $h$ from each AMA character. Principal component analysis (PCA) was used to derive an ordered set of orthogonal bases spanning the space of AMA hand postures (cf. Johnson and Wischem 1998). Over 93% percent of the variance among these hand postures was represented within the subspace spanned by the first seven PCs, consistent with the previous report by Jerde et al. (2003a). We computed variance accounted for (VAF) as

$$VAF = 1 - \left( \frac{\sum \text{var}(e)}{\sum \text{var}(f)} \right) \times 100\%$$

where $J_{n \times 19}$ is the matrix containing the glove signal recordings for a given set (or subset) of hand postures and $e_{n \times 19}$ is the difference between the matrix $J$ and the hand signals matrix obtained by projecting the low-dimensional representation of those postures back into the full-dimensional space using the Moore–Penrose pseudoinverse.

We sought to identify a mapping whereby screen target locations for certain gesture sets would be more informative of the postural variation among them than for other {{gesture set}, {target set}} pairs. We empirically identified a subset of five AMA characters {A, B, C, V, W} that were nearly coplanar within a three-dimensional subspace spanned by the first three PCs and thus most of the postural variability between these characters was captured by their projection onto this plane. We then performed PCA on signals derived from only these five characters, and the projection matrix $A$ was constructed by appending, side by side, the resulting first two PCs. Consequently, any screen location could be reached by superimposing motion along the two finger coordination patterns reflecting the greatest postural variation within set {A, B, C, V, W} (Fig. 1A). Under the mapping $A$, variability in screen target locations for {A, B, C, V, W} encode >80% of the postural variance inherent to this character set, whereas screen targets encode, respectively, only 40 and 20% of the variability inherent within the sets {F, I, N, T, U} and {D, E, M, X, Y}. We also computed the amount of postural variability inherent to each character set [i.e., the square root of the mean of the diagonal of the hand signal vector covariance matrix, reported as a percentage of glove signal full-scale range (FSR)]. This measure was comparable across the three sets: {ABCVW}: 13%; {FINUT}: 12%; and {DEMYX}: 12% FSR and thus one set of gestures does not require finger configurations substantially more differentiated than the others.

**FIG. 1.** **A:** hand configuration was mapped onto a computer screen (the 2-dimensional endpoint space) through a smooth and linear mapping $A$. Representative hand postures at the center (rest) and boundaries of the cursor’s workspace are shown to the right. **B:** gestures to be learned were selected such that under the mapping $A$, screen targets encoded >80% of the variance in hand configuration for the American Manual Alphabet (AMA) character set {A, B, C, V, W}, but target locations were much less informative for the other 2 character sets: 40 and 20% of the variance in hand configuration were captured for sets {F, I, N, T, U} and {D, E, M, X, Y}, respectively (see text for details). For test subjects, movements were cued by the appearance of circular training targets (black) at screen locations corresponding to the desired hand gesture under projection matrix $A$. Generalization targets (yellow) were placed at screen positions located inside and outside of the endpoint space explored during training. The central resting point (white) corresponded to the projection of a naturally relaxed hand posture (shown in A, center). Two control groups trained with the same targets as the 80% group, but under different visual feedback conditions. Control-1 subjects were provided explicit visual cues (pictures) on most trials indicating which gesture to form. Like test subjects, the control-2 group was cued with small circular targets. Unlike test subjects, they received cursor feedback of the hand’s final configuration only at the end of movement and only if the cursor fell outside small neighborhoods surrounding the generalization targets. (Exclusion region boundaries were not visible to subjects.)
Session 1: calibrating the hand-to-screen mapping

Session 1 was intended to identify subject-specific mappings such that by bringing the cursor to screen targets corresponding to the different character sets, differing amounts of information about the hand gestures to be learned would be provided. Each of the 25 training subjects performed three repetitions of a random sequence of hand postures corresponding to the 24 static characters. Subjects were provided photographic images of the desired hand postures as formed by the AMA instructor (e.g., Fig. 1B, left). Gesture formation was coached by the investigator who ensured that there were no differences between the desired and realized gestures at the moment of data sampling. Projection matrices were constructed by applying PCA to the data from character set \{A, B, C, V, W\} as earlier. The hand-to-screen mapping varied slightly from subject to subject due to differences in hand size. Across subjects, \(A\) captured \(82 \pm 3\%\) (mean \(\pm\) SD, both here and elsewhere) of the hand signal variance among postures in the character set \{A, B, C, V, W\}. The same matrices captured \(43 \pm 5\%\) hand posture variance for \{F, I, N, T, U\} and \(23 \pm 6\%\) for \{D, E, M, X, Y\}.

Session 2: learning and generalization of the hand-to-screen mapping

Session 2 was intended to examine how the quantity and quality of visual information influences the learning and generalization of novel finger coordination patterns. Subjects were seated 0.6 m in front of a computer monitor that served as the endpoint space within which cursor motion might occur. Subjects were randomly assigned into one of three test groups or into one of two control groups; each group consisted of five subjects. No subject was informed of the relationship between hand configuration and cursor/target/picture location on the screen. Movements started from a common resting location (Fig. 1B, white dot) matching the hand configuration displayed in the center of Fig. 1A.

Each test group practiced forming one of the character sets by acquiring screen targets. A 5-mm-diameter cursor was continuously visible, providing feedback of endpoint motion within the task space. Trials began with the appearance of a 4-mm-diameter target at the screen location corresponding to a desired hand posture. Test subjects were instructed to form hand gestures so as to bring the cursor to the target as quickly and accurately as possible.

The two control groups trained on set \{A, B, C, V, W\} but were provided with qualitatively different visual feedback conditions. The control-1 group was included to examine how explicit visual cues of desired hand posture influences learning and generalization. Movements by these subjects usually were cued by the appearance of a small, 15-mm-diameter picture of the desired hand configuration at a screen location consistent with its projection under \(A\). One picture in each cycle of five training trials was replaced by a simple 4-mm-diameter target and each training posture was replaced once every five cycles. Regardless of cueing condition, control-1 subjects were instructed to form gestures corresponding to the required screen target locations. No cursor feedback was provided at any time.

The control-2 group was included to determine whether subjects might learn the hand-to-screen mapping if provided ample opportunity to observe how static hand postures projected into endpoint space. Like test subjects, the control-2 group was cued using simple targets. However, they did not receive cursor feedback during transition between targets but instead received a 2-s snapshot of final cursor location after each training trial [spatial knowledge of results (KR)\(_{S}\)] except when the cursor would have fallen within a small circular neighborhood around the generalization targets (subsequently described; Fig. 1B, right). This exclusion condition was invoked only \(5 \pm 2\) times per subject. Because \(A\) and its pseudoinverse both have 19 free parameters per endpoint degree of freedom, an optimally efficient learner would require only 19 examples of hand/cursor pairings to learn each mapping. Thus training provided control-2 subjects with nearly tenfold the theoretical minimum number of hand/cursor position pairs needed to formulate an internal representation of the kinematic transformation between hand configuration and cursor position. Control-2 subjects were instructed to minimize error between the final cursor position and its target.

In addition to the five training postures (Fig. 1B, black dots), all subjects were asked periodically to form gestures corresponding to screen locations that were not included in any training set (generalization targets; Fig. 1B, yellow dots). These targets did not correspond to any particular AMA character. In generalization trials, performance feedback was suppressed [both cursor feedback and a numerical knowledge of results (KR) score described in the following text] and subjects were instructed to form the hand posture best corresponding to the target location. The same generalization targets were used for all subjects, testing their ability to compose motor coordination patterns consistent with achieving novel objectives within the endpoint space. These trials evaluated whether learning of a novel gestural relationship between hand and amplitudem quantified by the type of visual training (explicit vs. implicit cueing of desired hand gestures) and/or by the amount of task-relevant information provided visually (80, 40, 20% VAF; online vs. terminal cursor feedback).

Training was conducted in 33 cycles of six movements each in an experimental session lasting about 1.5 h. Subjects formed each of the five training postures once along with one of the three generalization postures within each cycle. Each generalization target was presented once every three cycles. Targets and trial types were randomized within cycles. Movements were to be made within 3 s of target appearance. Subjects were instructed to hold the final hand shape until they were cued to relax. Subjects were then to bring the hand back to its resting configuration. After a brief pause, a new target was presented as the next trial began.

Performance feedback (numerical knowledge of results, KR\(_{S}\))

For test and control-1 subjects, KR in the form of a static numerical score was provided for 2 s after each training trial. This score was displayed directly above the target on the computer screen and was intended to encourage engagement in the task. Subjects were instructed to reduce the score from one trial to the next, as this would help them perform better. We computed KR\(_{S}\) as a measure of distance between the final realized posture and the desired hand signal vectors recorded during session 1. For each subject, we used PCA to decompose the hand postures recorded during session 1 into 19 ordered basis vectors. We then projected both the realized and desired hand signal vectors onto the space spanned by these eigenvectors. We computed the KR score as a weighted Euclidean norm wherein the difference between signal vectors along each eigenvector was scaled by its corresponding eigenvalue. Thus signal fluctuations contributing most meaningfully to finger spelling gestures were weighted more heavily than less meaningful fluctuations. Control-2 subjects did not receive KR\(_{S}\) feedback, allowing us to evaluate articulation space performance in the absence of all feedback driving learning of specific desired hand configurations.

Data analysis

We analyzed three aspects of kinematic performance. Articulation error was evaluated using KR\(_{S}\), which quantifies accuracy in the formation of AMA hand postures and thus may be estimated only for training trials, not generalization trials. A decrease in articulation error is expected if subjects learn the novel finger coordination patterns required by finger spelling. Endpoint space error was defined as the Euclidean distance between the target location on the computer screen and the endpoint space projection of the finger configuration realized at the end of each trial. Endpoint error was computed separately for
training trials (wherein ongoing cursor feedback was provided to test subjects) and generalization trials (wherein neither cursor feedback nor KR was provided to any subject). Endpoint errors were averaged over all the training trials in each movement cycle prior to hypothesis testing, whereas endpoint errors for generalization targets were averaged over the three generalization trials from each block of three consecutive cycles. Finally, we evaluated performance changes in degrees of freedom that did not contribute to endpoint motion using null-space variability, defined as the variability within the null-space projection of final hand postures over each cycle of training targets (cf. Eq. 2).

We then derived a number of secondary performance measures. To facilitate comparison across subject groups, we computed the cycle-average articulation error (e) over each consecutive block of three cycles. To compare the speed of learning across groups, we estimated the initial learning rate for each subject by computing the average reduction of articulation error over the first 10 trial cycles (i.e., between trial cycles 1–3 and cycles 10–12). The extent of learning for each subject was estimated as the magnitude of articulation error at the end of the experimental session (i.e., at final trial cycles 31–33). For control-1 subjects, endpoint space errors from training trials with simple targets (i.e., implicit visual cues) were averaged within each block of five cycles to evaluate whether subjects learned to associate the training postures with their corresponding screen locations. Endpoint space performance was evaluated by computing both the reduction in endpoint error over the entire experimental session (between cycles 1–3 and 31–33) and the magnitude of error at the end of training.

One-way ANOVA was conducted to evaluate the specific effects of visual training conditions on spatial learning and generalization. Post hoc Tukey t-tests were conducted to identify significant changes in performance when ANOVA revealed a significant main effect. Effects were considered statistically significant at the P = 0.05 level. Data processing and statistical testing were carried out within the Matlab (The MathWorks, Natick, MA) and the Minitab (Minitab, State College, PA) computing environments. Data values are reported as mean ± 1SD. Error bars in figures also represent ± 1SD.

Spatial learning strategies differentiated by extrapolation in generalization

Because all subjects in the study by Mosier et al. (2005) were provided cursor feedback during error-correction maneuvers following movement, it is not clear whether online visual feedback of endpoint motion was indeed necessary for them to learn how the fingers should move so that the cursor would acquire arbitrary target locations (i.e., to learn an inverse mapping $A^{-1}$). We therefore sought to determine how visual training conditions might influence generalization trial performance. We evaluated whether performance gains in generalization trials were consistent with the learning of an inverse hand-to-screen mapping $A^{-1}$ or whether the different training conditions might have promoted another form of learning, such as the formation of associations between endpoint targets and hand gestures projecting onto them (i.e., a look-up table).

We compared subject performance with a look-up table implementation wherein each generalization target evoked a hand configuration that would project onto the training target nearest to it in endpoint space. Because each subject’s projection matrix $A$ was unique, the locations of generalization and training targets varied slightly from one subject to the next. Look-up table performance was computed as the across-subjects average of the mean distance between the three generalization targets and their nearest targets on the screen (Fig. 4B; gray area indicates the predicted mean ± 1SD performance).

Because interpolation between learned associations may also lead to improved performance, we further analyzed generalization performance based on target location: Whereas generalization targets located within the region of endpoint space visited during training (i.e., interpolation targets) might be expected to demonstrate performance improvements with either the inverse mapping or look-up table learning strategies, targets outside that region (i.e., extrapolation targets) should exhibit improvement only if subjects construct an estimate of the inverse mapping $A^{-1} = dh/dq$.

The extent to which test subjects reduced articulation errors depended on the amount of task-relevant information captured by the projection matrix: ANOVA and post hoc Tukey t-test found that the reduction in articulation error over the entire session varied systematically across test groups $[F_{(2,14)} = 11.47, P = 0.002]$. Again, the 80% group reduced errors to a greater extent by the end of training than had the 40% group ($P = 0.033$) who, in turn, reduced errors more than the 20% group ($P = 0.025$). This difference was not due to the training session being too short since all subject groups appeared to approach asymptotic performance by the end of training (Fig. 2A). Furthermore, since all test subjects received the same numerical KR information, this feedback was at best a weak driving force for the training of novel articulation space coordination patterns.

Training with reduced-dimensional visual feedback can be just as effective as explicit cueing in promoting accuracy in the formation of novel hand postures (Fig. 2C). ANOVA and post hoc Tukey t-test revealed a main effect of training condition on the magnitude of articulation error at the end of session 2 $[F_{(3,19)} = 36.96, P < 0.0005]$. Whereas the 80% group had lower articulation error than that of the 40% group ($P = 0.007$), the 80% group did not differ significantly from the control-1 group ($P = 0.598$). Although the 40% group had lower error than that of the 20% group ($P = 0.075$), this trend did not quite reach statistical significance. Thus only the test group provided with the most task-relevant information attained a level of performance on par with that of control subjects by the end of training.
the difficulty in forming the different character sets. Differences in training effects were not due to differences in beginning of training (we found no differences among the test groups at the end of training). Here and in the remaining figures, error bars indicate using the final KR score. Here and in the remaining figures, error bars indicate ±1SD.

we found no differences among the test groups at the beginning of training \( (P > 0.326 \text{ in all cases}) \), the observed differences in training effects were not due to differences in the difficulty in forming the different character sets.

**ENDPOINT SPACE ERRORS.** Subjects in each test group improved their ability to place the cursor on the training targets with practice (Fig. 3A, top and B, left). Endpoint space errors (obtained by projecting the final hand postures onto endpoint space) remained consistently low throughout the experimental session for trials with explicit postural cues (Fig. 3A; Control-1). In contrast, these same subjects reduced endpoint space errors in trials cued by simple spatial targets performed without ongoing cursor feedback. This indicates that they learned to associate each desired hand gesture with a specific screen target location in accord with projection matrix \( A \) (Fig. 3B, middle). Subjects in the control-2 group also learned to place the cursor closer to the training targets with practice (Fig. 3A, bottom), even though they received final cursor position feedback only after training trials. ANOVA and post hoc Tukey \( t \)-test found a main effect of training group on the reduction of endpoint error \( [F_{(4,24)} = 32.79; P < 0.0005] \). Because reduction in endpoint error did not vary across the three test groups \( [\text{ANOVA: } F_{(2,14)} = 2.18; P = 0.10] \), we collapsed across these groups before comparing the effect of training condition on endpoint space error reduction (Fig. 3B, right). Error reduction was significantly greater than zero for test subjects and for control trials cued by simple spatial targets \( (P < 0.0005 \text{ in each case}) \), but not for control trials cued by explicit postural cues \( (P = 0.449) \).

**NULL-SPACE VARIABILITY.** Subjects in all three test groups as well as those in the control-2 group learned to reduce hand posture variability within the null-space projection of \( A \) (Fig. 3C). Hand configuration variability for control-1 subjects with explicit cueing remained low throughout the session. ANOVA and post hoc Tukey \( t \)-test found a main effect of training condition on the reduction of null-space variability \( [F_{(4,24)} = 15.57; P < 0.0005] \). The three test groups and the control-2 group reduced null-space variability more than the control-1 group \( (P < 0.023 \text{ in each case}) \), although we found no significant difference between any two test groups \( (P > 0.66 \text{ in all cases}) \). We therefore collapsed across test groups to compare the effect of training condition on the reduction of null-space variability (Fig. 3C, right). We found a greater effect for test subjects than for control-1 subjects \( (P = 0.005) \). Although control-2 subjects decreased their null-space variability to an even greater extent with practice than did test subjects \( (P < 0.0005) \) or control-1 subjects \( (P < 0.0005) \), this was largely due to the greater initial variability in null-space performance of the control-2 group at the onset of training (compare Fig. 3C, left and middle).

**Generalization trial performance**

Generalization trial performance differed markedly between subjects who trained with visual feedback of cursor motion (Fig. 4, A, top and B) and those who either were cued with pictures of hand postures to form (control-1 group) or those who received static cursor feedback (control-2 group) (Fig. 4, A, bottom and B). ANOVA and post hoc Tukey \( t \)-test found a main effect of training condition on reduction in generalization trial error from the beginning to the end of training \( [F_{(4,24)} = 6.17, P < 0.002] \). All three test groups performed significantly better than both control groups \( (P < 0.05 \text{ in all cases}) \), although we found no significant difference between any two test groups \( (P > 0.95 \text{ in all cases}) \). One-sample Tukey \( t \)-test found that all three test groups demonstrated a significant
reduction of generalization error from the beginning to the end of training ($P < 0.009$ in each case), but control groups did not (control-1: $P = 0.432$; control-2: $P = 0.100$) (Fig. 4, B and C). By the end of training, only the test groups attained a level of performance exceeding that predicted by a simple implementation of a look-up table between practiced target locations and hand configurations that would project onto them (Fig. 4B, gray band). Whereas subjects in groups provided with continuous cursor feedback clearly generalized beyond the training set with practice, subjects deprived of such feedback demonstrated no ability to generalize, even though they did in fact learn to associate screen targets with specific hand postures (Fig. 3B, middle).

**Interpolation and extrapolation in generalization**

Because interpolation between learned associations may also improve performance beyond that for the simplest look-up table implementation (described earlier), we separately analyzed generalization errors for test subject trials made to interpolation and extrapolation targets. The frequency distribution maps of Fig. 5, A–C depict the relative likelihood that subjects formed hand postures projecting onto each location in the endpoint space. Brighter highlighting indicates a greater amount of training time spent with hand configurations projecting onto that endpoint location. For the 80 and 40% groups, two of the generalization targets were within the region of...
endpoint space visited during training, whereas only one generalization target was within the visited space for the 20% group. These were considered the interpolation targets, whereas the remaining generalization targets were considered the extrapolation set (those indicated by “e” in Fig. 5, A–C). Test subjects reduced generalization error for extrapolation as well as interpolation target sets (Fig. 5D), with the reduction from beginning to end of training significantly greater than zero for both sets (one-sample t-test: $P < 0.0005$ in both cases; Fig. 5E). We found no significant difference between the two sets in reduction of generalization error (two-sample t-test: $P = 0.195$). Thus generalization trial performance is consistent with the estimation of an inverse mapping $\hat{A}^{-1}$ rather than with interpolation between learned hand/cursor associations.

**DISCUSSION**

These experiments examined how visual feedback of cursor motion and explicit postural cueing influence learning and generalization of a novel kinematic relationship between high-dimensional finger coordination patterns and object motion within a low-dimensional workspace. As originally hypothesized, learning rates depended on both the type of training provided (explicit vs. implicit cueing) and the amount of task-relevant information encoded within the spatial location of screen targets. The extent of learning varied only with the amount of task-relevant visual feedback: Those test subjects provided with the most relevant visual feedback performed as well as control subjects shown explicit visual cues of desired hand gestures. The current findings extend those previously reported by Mosier et al. (2005) in that control subjects did not generalize beyond the endpoint space explored during training when provided either explicit postural cues or terminal cursor feedback, whereas subjects provided continuous feedback of cursor motion did. Thus visual feedback of endpoint motion is indeed necessary for learning a new spatial relationship between endpoint and articulation spaces.

**Visual feedback of endpoint motion and the learning of visuomotor transformations**

The finding that online visual feedback is needed to learn a novel visuomotor mapping is not without precedence. The experimental psychology and motor psychophysics literature contain numerous studies of how motor commands adapt in response to displacements or rotational distortions of visual feedback during reaching and pointing tasks (e.g., Bock 1992; Ghez et al. 2007; Held and Freedman 1963; Held and Hein 1958; Ingram et al. 2000; Kitazawa et al. 1997; Krakauer et al. 2000; Redding and Wallace 1988, 1992; Scheidt and Ghez 2007; Southall 1962; Stratton 1897a,b). It has been noted, however, that two distinct types of responses can occur when visual perturbations are imposed suddenly, such as in the donning of prism goggles (Redding and Wallace 1996). Strategic responses include the adjustment of feedforward motor commands and/or the invocation of compensatory control actions that cause a direct and immediate reduction in performance errors (Clower and Boussaoud 2000; Redding and Wallace 1996; Welch and Sampanes 2004; cf. Mazzoni and Krakauer 2006; Redding and Wallace 1997). This response class would include, for example, the immediate elevation of limb impedance in response to suddenly imposed loads as observed in studies of mechanically perturbed reaching (cf. Milner and Franklin 2005; Takahashi et al. 2001). As discussed in the following text, it also includes the conditional execution
of learned motor responses cued by specific environmental stimuli such as those demonstrated by our two control groups. In contrast, adaptive spatial alignment is thought to result from a gradual remapping of spatial positions represented in one sensory coordinate frame (e.g., vision) onto their corresponding positions in another reference frame (e.g., proprioception) (Clower and Boussaoud 2000; Redding and Wallace 1996; Wann and Ibrahim 1992; Welch and Sampanes 2004; cf. Redding and Wallace 1997). The magnitude of this second response is typically evaluated using tests for aftereffects and generalization of learning.

Both control groups in our study clearly learned to associate gestures with each target location (Fig. 4B, middle, spatial cues). Noise in the sensory and motor systems (cf. Barlow et al. 1987; Scott and Loeb 1994) undoubtedly led to uncertainty about both the spatial location of visual targets and the postural configuration of the hand. The learned associations must therefore have distributed spatial sensitivities to have practical utility. That is, a hand posture most strongly associated with one screen target location \( q_T \) should have had nonnegligible association with neighboring locations \( q_N \) if the subject was to respond reliably. This spatial relationship, commonly called a receptive field, can be modeled by a “radial basis function” that describes the strength of association \( \alpha \) as a nonlinear function of distance \( \Delta q \) from the training target in endpoint space

\[
\alpha = e^{-|\Delta q|/\sigma^2}
\]

where \( \sigma \) represents the “width” of the association’s sensitivity function (cf. Poggio and Bizzi 2004).

It has long been recognized that parietal association cortex plays a critical role in sensorimotor integration (Lynch et al. 1977; Mountcastle et al. 1975), combining information from multiple senses for use in formulating motor behaviors (Andersen 1989). Computational simulations by Pouget and colleagues (Deneve et al. 2001; Pouget and Sejnowski 1997; see also Salinas and Abbott 1996) and neurophysiological studies from the Andersen group (Andersen et al. 1985; DeSouza et al. 2000) suggest that neurons found in the parietal cortex compute basis functions of their sensory (visual and proprioceptive) inputs as they map them onto motor responses. Pouget and Sejnowski (1997) contend that nonlinear integration of information from the different sensory modalities within parietal cortex enables the neural computation of any nonlinear sensorimotor transformation. This logically includes those needed to transform a desired displacement of the cursor into an appropriate displacement of the fingers. In the model of Pouget and Sejnowski (1997), generalization to targets within the space explored in practice may occur by interpolating between learned associations if the basis functions are sufficiently broad. However, our control subjects failed to generalize beyond their trained target sets, even though two of the generalization targets were within the space spanned by their training targets. This suggests that receptive fields for the associations learned by our control subjects were too narrow to support generalization by interpolation. That is, these subjects were not effective in importing into articulation space a representation of the spatial relationships between targets in endpoint space.

In contrast, test subjects in our study discovered how spatial relationships between endpoint targets could constrain the coordination of highly redundant control signals so as to produce accurate and efficient finger motions (i.e., they esti-
mated an inverse hand-to-screen mapping $\hat{A}^{-1}$). What aspect of visual feedback facilitated this learning? One possibility is that subjects in our experiments attempted to “solve” for each of the 38 parameters in $\hat{A}^{-1}$ [19 coefficients per endpoint degree of freedom (df)] by relating realized hand postures to either specific screen target locations (control-1 subjects) or cursor locations (control-2 and test subjects). Although control-1 subjects clearly learned to associate gestures with each target location, they could not have been expected to learn the inverse mapping because they experienced only 10 unique associations during training (5 targets $\times$ 2 df). The control-2 group also formed associations between targets and hand gestures within about 15 training cycles, but failed to generalize. This was despite observing nearly 10-fold the theoretical minimum number of hand/cursor pairings needed to solve for $\hat{A}^{-1}$. If they do so at all, subjects are highly inefficient in performing such computations.

Test subjects significantly reduced generalization error within about six trial cycles (Fig. 4B) and reduced null-space motion within about 15 cycles (Fig. 3C). We calculated that by the time these subjects learned to generalize, they had viewed at most 12-fold the number of pairings of control-2 subjects (this number being limited by the 30/s update rate of the visual display and the duration of movements). Although it is possible that additional observations facilitated spatial learning in test subjects, this explanation is not very compelling because both test and control-2 subjects observed many more times the necessary number of pairings to estimate $\hat{A}^{-1}$, yet the difference in generalization performance between the groups was categorical, not one of degree. Rather, continuous visual feedback provided a second form of information that we believe was critical for forming a motor representation of the endpoint space within which the cursor moved—gradient information relating how cursor position changes as a result of changes in hand posture. This is precisely the type of information one needs to compose both forward ($\hat{A} = dq/dh$; see Eq. 1) and inverse ($\hat{A}^{-1} = dh/dq$) models of the hand-to-screen mapping. Indeed, psychophysical evidence suggests that detection of—and online feedback correction for—misalignment between spatial reference systems is necessary for adaptive realignment to occur, whereas off-line performance corrections using knowledge of results do not suffice (Clower and Boussaoud 2000; Wann and Ibrahim 1992; for a review see Redding and Wallace 1997).

Why should visuomotor adaptation be contingent on feedback correction of alignment errors? One suggestion is that feedback mechanisms may play a fundamental role in the formation of internal models used to generate motor commands for desired movements (cf. Kawato and Gomi 1992, 1993). According to the feedback-error-learning model of Kawato and colleagues, motor commands are tuned in a supervised learning fashion using motor error information generated during movement by spinal, brain stem, and/or cerebral feedback control networks (Kawato and Gomi 1992). These feedback mechanisms provide error information—encoded in motor command coordinates, not sensory coordinates—that can be used to train the neuronal networks (inverse models) that transform desired movement trajectories into motor commands. Whether this error information is obtained during the movement (Kawato and Gomi 1992, 1993) or during corrective actions that occur afterward (Fagg et al. 1998), learning would not occur without detection and correction of performance errors because it is the error signals themselves (processed by the feedback control mechanisms) that drive learning. The cerebellum likely plays an important role in this form of supervised learning (Doya 1999; Fagg et al. 1997; Houk et al. 1996; Kawato and Gomi 1992) and convincing evidence for the existence of internal models in the cerebellum has been provided by neurophysiological studies of oculomotor control (Kawano et al. 1996) as well as by neuroimaging studies of hand movements (Kawato et al. 2003; Zimbelman et al. 2007; see also O’Doherty et al. 2003).

**Motor redundancy and the reduction of null-space variability**

Kinematic performance errors are not the only criteria likely being optimized during learning in this study. It has long been recognized that the problem of learning to control object motion in a low-dimensional space is ill-posed in that there exists an infinite set of control variable changes yielding any desired displacement (cf. Bernstein 1967; Hadamard 1902). What factors predispose the brain to select one option over all others? Todorov and Jordan (2002) and others (Latash et al. 2002; Scholz and Schoner 1999) suggest that an optimal controller should capitalize on redundancy by correcting only those motions that interfere with task performance (i.e., the “minimum intervention” principle; Todorov and Jordan 2002). Deviations in task-irrelevant dimensions should be ignored because an optimal controller has nothing to gain from correcting those deviations and because the cost of control likely includes a penalty on “effort” that would increase due to corrective action (Todorov and Jordan 2002). However, as reported previously by Mosier et al. (2005), subjects practicing with reduced-dimensional visual feedback became more proficient at solving the ill-posed problem while reducing null-space variability with similar time courses (compare Fig. 3, B and C). Since numerical KR feedback was informative of articulation space performance as well as endpoint error, it is possible that subjects used the KR score to preferentially reduce null-space variability. This seems unlikely because KR score reduction for the 20% group was only a fraction of that for the 80% group (Fig. 2B), whereas the reduction of null-space variability was similar across the three test groups. Furthermore, a similar reduction in null-space error was also observed in Mosier et al. (2005) in the absence of numerical KR feedback. It is therefore more likely that reduction in null-space variability we observed belies the influence of implicit optimizations penalizing inefficient muscle activations, excessive joint torque changes, and/or other factors related to movement energetics (cf. Scheidt et al. 2000).

Because people are adept at adjusting the relative importance of kinematic and energetic costs based on task constraints (e.g., Burdet et al. 2001; Franklin et al. 2003; Tee et al. 2004; see also Bunderson et al. 2007; Kuo 1995, 2005), the brain likely maintains separate representations of kinematic and energetic costs for later integration in the evaluation of motor performance. Moreover, since variability in task-relevant dimensions will incur predominantly kinematic costs, whereas “task-irrelevant” variability will influence only other (e.g., energetic) costs, we speculate that a solution to the ill-posed problem arises from competition between kinematic and ener-
Reduced-dimensional kinematic training

The question of how the brain uses sensory feedback to guide the resolution and/or exploitation of control redundancies has relevance for the identification of effective therapeutic approaches to the neurorehabilitation of hand and finger function. Although the ability to generate finely individuated finger movements is important for many activities of daily living such as buttoning a shirt or picking up a spoon, individuated finger motion is often compromised post-stroke. The current study describes a principled approach for training novel finger coordination patterns via object manipulation on a computer screen. In general, rehabilitation therapies are more effective when they engage the patient’s interest (cf. Bach-y-Rita et al. 2002). By combining the use of an instrumented data glove with engaging computer applications requiring target capture in a reduced-dimensional endpoint space, it may be possible to promote recovery of individuated finger motion in all the fingers in patients where residual control is preserved (cf. Jack et al. 2001; Merians et al. 2002), not just the thumb and index finger as is frequently the case clinically (Lang and Schieber 2003; but see also Raghavan et al. 2006). Although not specifically tested here, a potential benefit of this approach over explicit cueing of desired postures is that pictorial cues may be limited in the extent to which they can promote precision in individuated finger motions because differences between similar pictorial cues could evade recognition, whereas ongoing feedback of cursor position relative to a spatial target can be amplified as needed to increase the salience of postural errors. Although explicit cueing of desired movements has been shown to be detrimental to some tasks post-stroke (Boyd and Winston 2004), using implicit cues such as visual feedback of cursor motion may be beneficial in others (Patton et al. 2006). The current study suggests that to maximize training efficiency, the hand-to-screen transformation matrix should be designed so as to project as much variability among the set of desired postures as possible onto the endpoint workspace. Additional experiments are needed to determine whether reduced-dimensional kinematic training can indeed promote plastic reorganization within the brain so as to release individual finger movements from the obligatory coupling observed clinically post-stroke.

Acknowledgments

We thank S. Mussa-Ivaldi, K. Mosier, and A. Fishbach for insightful comments on a previous version of this manuscript.

Grants

This work was supported by National Science Foundation Grant BES 0238442.

References


