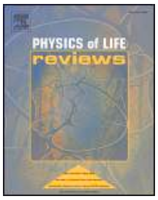
Physics of Life Reviews 51 (2024) 442–458

Contents lists available at ScienceDirect

Physics of Life Reviews

journal homepage: www.elsevier.com/locate/plrev



Review

Kinematic coding: Measuring information in naturalistic behaviour

Cristina Becchio [a,*], Kiri Pullar [a,b], Eugenio Scaliti [a,c,d], Stefano Panzeri [b,*]

[a] *Department of Neurology, University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany*
[b] *Institute for Neural Information Processing, Center for Molecular Neurobiology Hamburg, University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany*
[c] *Department of Management "Valter Cantino", University of Turin, Turin, Italy*
[d] *Human Science and Technologies, University of Turin, Turin, Italy*

ABSTRACT

Recent years have seen an explosion of interest in naturalistic behaviour and in machine learning tools for automatically tracking it. However, questions about what to measure, how to measure it, and how to relate naturalistic behaviour to neural activity and cognitive processes remain unresolved. In this Perspective, we propose a general experimental and computational framework – kinematic coding – for measuring how information about cognitive states is encoded in structured patterns of behaviour and how this information is read out by others during social interactions. This framework enables the design of new experiments and the generation of testable hypotheses that link behaviour, cognition, and neural activity at the single-trial level. Researchers can employ this framework to identify single-subject, single-trial encoding and readout computations and address meaningful questions about how information encoded in bodily motion is transmitted and communicated.

1. Introduction

In the real world, human perception, emotion, and decision-making manifest through behaviour, directing the selection of motor actions (what to do, whether and when to do it) and their execution (how to do it). Despite this inherent link to cognition [1], the characterization of behaviour has been historically overlooked in the domains of cognitive psychology and neuroscience [2].

Several interrelated reasons have been suggested for this surprising neglect: the traditional focus of psychology, inherited from philosophy, on knowledge rather than action [2]; the tacit assumption that action parameters are nonspecific to cognitive processes [3], often coupled with a reductionistic bias [4]; a robotic view of motor control as detached from decision-making [5]; the notion of motor variability as a confounding factor (for example, a potential cause of 'movement artifacts' in neuroimaging).

Countering this oversight, recent years have seen a surge in advocacy for approaches that link measurements of neural activity and cognition to behaviour [4,6–8]. Paraphrasing Krakauer and colleagues [4] – neuroscience needs behaviour. This emerging recognition has been supported, and perhaps accelerated, by recent developments in machine learning tools for automated tracking of animal behaviour, including 3D pose estimation and tracking of social interactions between animals in their environment [9–15]. Leveraging these advancements, several approaches have been proposed for jointly modelling neural activity and behaviour, enabling learning of behaviourally relevant neural dynamics [16,17].

Despite the pertinence of these advancements to human research, studies in cognitive neuroscience and psychology – from learning to social decision-making – still predominantly rely on experimental designs that minimize behaviour to a button press [18,19].

A significant barrier to adopting more naturalistic paradigms is the inherent complexity of naturalistic behaviour. Naturalistic

---

* Corresponding authors.
E-mail addresses: c.becchio@uke.de (C. Becchio), s.panzeri@uke.de (S. Panzeri).

https://doi.org/10.1016/j.plrev.2024.11.009
Received 13 November 2024; Accepted 14 November 2024

behaviour is dynamic, high-dimensional, and highly variable both within and across individuals [6]. This complexity presents challenges for both quantifying behaviour and connecting it with neural activity.

To address these challenges, here we propose an *information-based approach.* Our approach is motivated by the basic observation that the way we move carries information. This information may reflect states of the external world, task variables directly manipulated by the experimenter but also cognitive states and decision processes [3,20]. For example, when reaching for an object, movement kinematics convey information not only about the size and shape of the object to be grasped [21] but also about the agent's intentions [22,23] and expectations in grasping the object [24]. In choice paradigms, movement kinematics reveal belief tracking [25–27], evaluative biases [28], reward associations [29], confidence [30], motivation [31] and the subjective utility assigned to the chosen option [5]. Moreover, specific events within reaching trajectories such as changes in direction and velocity can be used to detect changes of mind [32,33] and changes in confidence [34].

Our proposal in this Perspective is to measure information conveyed by structured movement patterns, rather than focusing on the patterns themselves. This approach offers several advantages. One conceptual advantage is that most models in psychology and human neuroscience are already cast – either explicitly or implicitly – in terms of information processing [35–41]. An information-based approach, thus, provides a natural framework for generating and testing hypotheses relating naturalistic behaviour to cognitive and neural activity.

A second advantage is methodological. Though movement kinematics and neural activity share several characteristics, they differ in their fundamental physical units. Movement kinematics is defined by position vectors, velocities, and joint angles [42], whereas neural activity is defined by action potentials, changes in the regional blood concentrations of oxy- and deoxy-haemoglobin (BOLD signal) or spectral properties of extracellular potentials [43]. This disparity poses a challenge for direct comparisons between kinematic and neural data. An information-based approach overcomes this challenge by abstracting kinematic and neural data into a common information space. In this space, variations in movement kinematics and neural activity are projected onto a common information axis. This allows variations in behaviour to be directly related with variations in neural activity and cognitive processing at the single trial level.

In this Perspective, we propose a general conceptual framework – kinematic coding – for measuring how information about cognitive states is coded within structured movement patterns. We begin by introducing the fundamental concepts that underlie kinematic coding – kinematic encoding and kinematic readout. We then describe experimental and computational approaches for investigating kinematic encoding, kinematic readout, and their intersection. We discuss what these approaches can reveal about the mapping between information and naturalistic behaviour. Finally, we explore applications and elaborate on how kinematic coding can inform the design of new experiments and generate testable predictions linking behaviour to cognitive processes and neural activity.

2. Kinematic coding: a conceptual framework for intersecting information encoding and readout

A code serves as a method of mapping information from one space to another [44]. Neural coding refers to how information about the state of the external world is mapped onto neural population activity [38,45]. Early research in neural coding focused on how much information a neural response carries about sensory stimuli [38,40,45–48]. Over time, the emphasis has shifted towards understanding how this encoded information is transmitted and read out by downstream regions to inform behavioural outcomes [49–53]. Despite this expanded scope, in neural coding the flow of information remains confined to within individual brains. While brains are sometimes described as "interacting", direct interaction between brains does not occur. Unless interbrain neurofeedback is provided [54], information encoded in neural activity is inherently private, inaccessible to others [55].

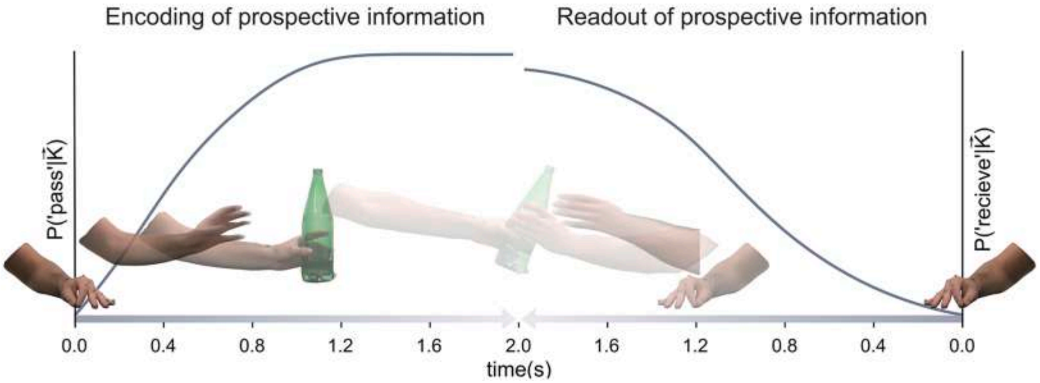


Fig. 1. | **Kinematic coding enables social transmission of information.** During social interactions, prospective information encoded in movement kinematics is read out and used to plan an appropriate action in response. Kinematic encoding estimates the posterior probability of a specific cognitive state (e.g., the intention to pass the bottle, shown on the left) from movement kinematics ($\vec{K}$). An increased posterior probability over time indicates an increased likelihood that the movement is intended to pass the bottle. As the interaction progresses, the partner reads out the encoded intention from the movement kinematics, enabling them to predict and plan an appropriate action in response (receiving the bottle). In this situation, intention can be inferred directly from kinematics of the partner's movement.

By transposing the concept of coding to naturalistic behaviour, kinematic coding moves the boundaries of information flow beyond the individual, introducing a social dimension absent in neural coding (Fig. 1). Unlike information encoded in neural activity, information encoded in behaviour is potentially available to others [3]. This opens an entirely novel set of questions regarding the social transmission of information. What information can human (and non-human) observers read in others' movements? Can information encoded in movement kinematics act as a social signal to others? Can it be used to deceive others?

Kinematic coding provides a unified framework for exploring how information about a particular cognitive dimension is encoded and read out in naturalistic behaviour. The relevance of kinematic coding for modelling naturalistic behaviour is not limited to any specific movement pattern or cognitive state. This framework can be used to measure information in eye movements [56–60], head movements [61], arm and hand movements [21,62–66], across a variety of behaviours, from crawling [67] and walking [68] to object manipulation, to racquet and ball sports [69–76]. Additionally, while kinematic coding lends itself naturally to measuring information in kinematic data, it can also be directly extended to quantify information in data obtained from other measurement modalities such as electromyography [77–79].

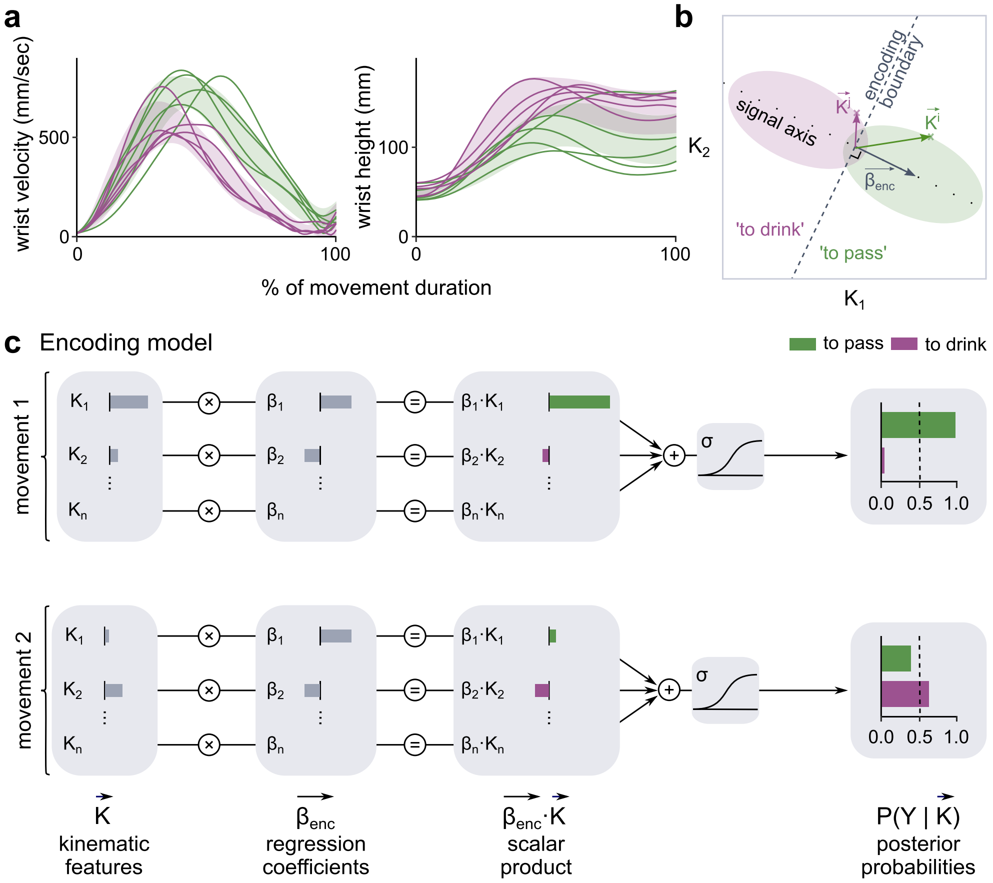


Fig. 2. | **Kinematic encoding.** (**a**) Time course of wrist velocity and wrist height for reach-to-pass and reach-to-drink movements. Thin curves show representative individual movement trajectories; shaded areas show the mean ± SD across movements. (**b**) Geometrical representation of the encoding model in a simplified two-dimensional kinematic space spanning two kinematics features $[K_1, K_2]$. The two elliptical regions represent the intention conditional probability distributions of the two features. The encoding boundary (dashed line) optimally separates 'to drink' and 'to pass' regions in the two-dimensional kinematic space. The encoding vector $\overrightarrow{\beta_{enc}} = [\beta_1, ..\beta_n]$, consisting of the regression weights, indicates the maximally discriminative axis (signal axis). Two different single-trial kinematic vectors, $\overrightarrow{K^i}$ and $\overrightarrow{K^j}$, are shown. (**c**) Block diagram illustrating a GLM encoding model using n kinematic features. The model projects the kinematic vector $\vec{K} = [K_1, ..K_n]$ of each movement onto $\overrightarrow{\beta_{enc}}$ by computing the scalar product $\overrightarrow{\beta_{enc}} \cdot \vec{K}$, which is then transformed into the posterior probability $P\left(Y \middle| \vec{K}\right)$ of state Y given kinematics $\vec{K}$. This transformation from scalar product to posterior probability is computed using a link function σ. A sigmoid function σ, corresponding to the specific case of logistic regression, is used in this illustration. The diagram shows that different features can contribute to varying degrees in the encoding of single-trial information (the product of the encoding and kinematic vectors along each feature) for different movements. For instance, K_1 makes a dominant contribution to movement 1, whereas both K_1 and K_2 contribute comparably to movement 2.

3. Kinematic encoding

Kinematic encoding refers to how information about a particular cognitive state Y is expressed (encoded) in a movement pattern $\overrightarrow{K}$. Experimentally, we can measure the strength of this encoding, i.e., mapping from Y to $\overrightarrow{K}$, by examining the extent to which variations in Y result in changes in $\overrightarrow{K}$.

This approach has been used to investigate how intentions are encoded in reach-to-grasp kinematics [80]. In everyday object manipulation, objects are reached for and grasped in ways that reflect intention – what one plans to do next with the object [3,81]. In a laboratory setting, this can be translated into a task where participants perform manipulative action sequences that differ only in their final act. For instance, participants might be instructed to reach, grasp, and lift a bottle either to pass it to another person or to drink from it. By analysing the extent to which variations in Y (intention) result in changes in $\overrightarrow{K}$ (reaching kinematics), we can measure intention encoding, i.e., the mapping from intention to kinematics.

A challenge in quantifying this mapping is the high variability in movement kinematics [82,83]. Fig. 2a shows the time course of two kinematic variables, wrist height and wrist velocity, measured during repeated reaching movements toward a bottle by different participants. These actions were performed either to pass the bottle to another person or drink from it. Despite reaching for the same bottle over a fixed distance, the kinematic trajectories for the same intention vary across trials and individuals.

To mitigate the impact of movement variability, the standard approach is to average kinematics over repeated trials and different individuals. However, motor control does not operate according to ‘averaged kinematics’. The neuromotor system is designed to facilitate the emergence of motor solutions that vary from one individual to another [84] and vary from trial to trial never repeating exactly [85]. Given this variability, we propose that kinematic encoding is best understood as probabilistic mapping. In this framework, measuring kinematic encoding is equivalent to computing the conditional probability of observing specific kinematics given a particular cognitive state (e.g. intention to pass). By applying Bayes’ rule, this probability can then be inverted to calculate the posterior probability that a movement was performed under a certain cognitive state based on the observed reaching kinematics in that trial. This posterior probability provides a measure of single-trial information encoding [82].

Encoding models for estimating posterior probabilities of Y from $\overrightarrow{K}$ based on empirical data can be computed using various parametric and nonparametric methods [86]. In a scenario where the predicted state is binary (e.g., the reaching act is performed either to pass or drink, as in the above example), simple classifiers such as Support Vector Machines [87] or regression models such as Generalized Linear Models (GLM) [88], including logistic regression [89] can be used to estimate the posterior probability of particular state given the reaching kinematics observed in that trial (single-trial movement kinematics).

Geometrically, single-trial movement kinematics can be conceptualized as a point in n-dimensional kinematic space, where each dimension represents a specific kinematic feature. The encoding boundary defines the surface within this kinematic space that best separates the two cognitive states. When GLM regression is used, the kinematic space is a n-dimensional vector space (that is, each point in this space corresponds to a vector extending from the origin to that specific point) and the encoding boundary is a hyperplane (a flat surface) (Fig. 2b). The encoding vector, orthogonal to the encoding boundary and with components equal to the regression weights of the encoding model, indicates the information axis along which changes in movement kinematics maximally discriminate between the two states (e.g., to pass or to drink). Single-trial information encoding can be obtained by projecting each movement onto the encoding vector. As shown in Fig. 2c, the encoding model estimates the posterior probability of intention for each movement in kinematic space as a function (a sigmoid function in the case of logistic regression [89]) of the scalar product between the encoding vector and the kinematic vector on that trial. Larger values of this calculation indicate a greater distance from the encoding boundary (Fig. 2b) and thus a higher posterior probability that the movement was performed with a particular intention (Fig. 2c).

This approach generalizes to cases where encoding is nonlinear – for instance, information resides in the interaction between kinematic features [53,90] – and/or the high-dimensional kinematic space is not necessarily a vector space. In such instances, the encoding model separates points in the n-dimensional kinematic space using an encoding surface that is not necessarily a hyperplane. The posterior probability for each intention increases for points farther from this encoding surface. This approach can also be extended to model scenarios where the choice is among three (ternary classifiers or regression) or more alternatives, using methods such as multiclass or softmax regression or classification [89,91–93].

The performance of the encoding model can be evaluated by integrating the single-trial information measures over all trials. For discrete state variables, classical statistical measures, such as sensitivity, specificity, and accuracy can be employed, alongside information theory measures (e.g. Shannon information) [45] (Box 1). For continuous state variables, it can be assessed using any measures that compute the model’s accuracy in estimating true values of state variables (Box 1). The performance of the kinematic encoding model provides a quantification of the overall amount of information encoded in kinematics.

The encoding model can also be used to characterize how the encoded information is represented. This involves identifying the specific kinematic features that encode information and quantifying the amount contributed by each feature. With GLM regression, the contribution of each feature to a given movement can be computed as the scalar product of the kinematic vector and the encoding vector within the subspace of the considered kinematic feature. Fig. 2c illustrates the case where different single-trial kinematic vectors (movement 1 and movement 2) are projected onto the same encoding vector. Differences in the posterior probability of Y reflect the differential contribution of individual kinematic features to intention encoding. The contribution of features across different movements can be computed by integrating such single-feature contributions. In the general case, the overall contribution of a kinematic feature can be assessed by comparing the performance of models that either include or exclude the considered feature (Box 1). Studies utilizing this approach to quantify intention encoding commonly find that, despite the high-dimensionality of kinematic data,

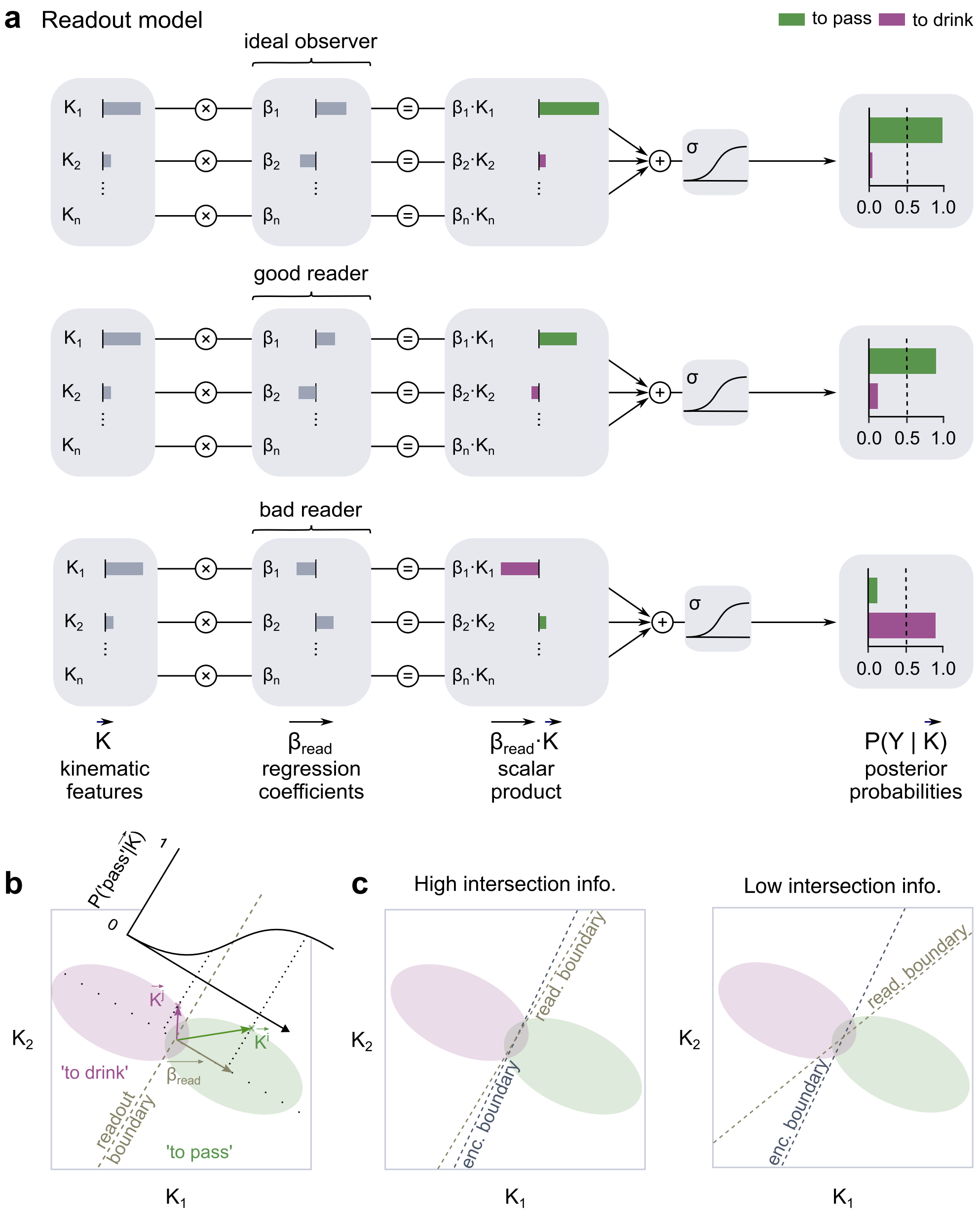


Fig. 3. | **Kinematic readout.** (**a**) Block diagram illustrating a GLM readout model using n kinematic features. The model projects each kinematic vector $\overrightarrow{K} = [K_1, ..K_n]$ onto the readout vector $\overrightarrow{\beta_{read}} = [\beta_1, ..\beta_n]$ by computing the scalar product $\overrightarrow{\beta_{read}} \cdot \overrightarrow{K}$, which is then transformed into a posterior probability $P\left(Y\middle|\overrightarrow{K}\right)$ of reported state Y given kinematics. The transformation from scalar product to posterior probability is computed using a (possibly non-linear) link function σ. A sigmoid function σ, corresponding to the specific case of logistic regression, is illustrated. (**b**) Geometrical representation of the readout model in a simplified two-dimensional kinematic space spanning two kinematics features $[K_1, K_2]$. The two elliptical regions represent the choice conditional probability distributions of the two features. The readout boundary (dashed line) optimally separates 'to drink' and 'to pass' choice regions in the two-dimensional kinematic space. The readout vector, $\overrightarrow{\beta_{read}}$ defined by the regression weights, indicates the maximally discriminative axis. Two different single-trial kinematic vectors, $\overrightarrow{K^i}$ and $\overrightarrow{K^j}$, are shown. (**c**) Geometrical representation of the intersection between encoding and readout visualized as the alignment between the encoding and readout boundaries for a given set of movements. When readout captures most of the encoded information, encoding and readout boundaries are closely aligned, and intersection information is high. In contrast, when readout is suboptimal, the two boundaries are misaligned, and intersection information is low.

intention information is typically encoded in lower-dimensional subspaces [82,83,94].

4. Kinematic readout

Having quantified information encoding, the next step is to determine whether this information is accessible to human observers. Experimentally, one approach is to ask naïve participants to distinguish between action stimuli that encode information about a particular cognitive state. When the choice is between two alternatives, the task conforms to a two-alternative forced-choice (2AFC) paradigm, requiring observers to discriminate between pairs of stimuli, differing along Y, in each trial. For example, observers might be asked to discriminate between two reaching movements, one performed with the intent to pass, the other with the intent to drink. The availability of displayed information can be manipulated through temporal and/or spatial occlusion [70,95]. For instance, temporal occlusion may be used to reveal only early movement patterns (and not the later portion of the movement) to probe the sensitivity of readout to anticipatory information [70].

Studies using this design to quantify readout of intention information in reaching movements indicate that naïve participants are able to read some, but not all, of the information encoded in moment kinematics [22,82,83]. A statistical readout model of single-trial intention choice provides a means to quantify the probability that encoded information is successfully read out.

Similar to information encoding, information readout in a single trial can be computed using Bayes' rule to determine the posterior probability that an observer judges the movement to be performed under a particular cognitive state (i.e., posterior probability of choice C) given the kinematics $\overrightarrow{K}$ observed in that trial. In cases where the readout variable is binary, a simple linear classifier (such as a GLM regression of the binary outcome) can be used to estimate the probability of reading a particular state from single-trial movement kinematics. Fig. 3b shows a schematic of a GLM readout model in a simplified, two-dimensional kinematic space. The model takes single-trial kinematics as input to estimate the probability that an observed reaching movement is judged to be performed with the intention to pass or to drink. The readout boundary defines the surface that best separates the two intention choices (to pass versus to drink). The readout vector, whose components correspond to the weights in the GLM regression model, identifies the axis orthogonal to the readout boundary along which variations in kinematics are most effective in differentiating between the two intention choices. Similar to encoding, the readout model projects each movement (point) in kinematic space onto the readout vector to compute the posterior probability of C given $\overrightarrow{K}$. As shown in Fig. 3a, this posterior probability is a function of the scalar product between the readout vector and the single-trial kinematic vector of the observed movement [82,83].

The performance of the readout model, computed by integrating the single-trial intention choice information across trials, provides a quantification of information readout. Similar to encoding, the readout model allows us to determine both the overall contribution of each kinematic feature to readout and the contribution of each feature to information readout in single trials. For a GLM readout model, the contribution of each feature to the posterior probability of intention choice is easily obtained by multiplying the readout vector by the single-trial kinematic vector within the subspace of each feature (Fig. 3a). As shown in Fig. 3a, the posterior probability of the same movement (that is, identical single-trial kinematic features) can vary significantly depending on the specific readout vector onto which the kinematics of the movement are projected.

So far, we have considered estimation of readout from binary options. In real-world scenarios, however, observers often face situations where an agent might have multiple possible intentions. For example, they might intend to pass the bottle, drink from it, or pour water into the glass. Similar to the encoding models, methods like multiclass or softmax regression can be applied to measure the likelihood of each potential intention. These methods can be further adapted to incorporate context-driven expectations [96] and increase the a priori likelihood of cognitive states more probable in a given context by adding a bias term.

Our approach also accommodates scenarios where 'behaviour affords behaviour' [97] and options interact with action dynamics [98]. In these settings, the readout state can be inferred directly from continuous, uninstructed measures of an observer's behaviour (continuous readout variables), such as eye movements [83,99–101] or muscle activity [102–104]. Research using this approach has revealed that during action observation, subjects implement motor programs equivalent to those used action execution [99,105]. For instance, corticobulbar excitability of the mylohyoid muscle involved in mouth opening is significantly greater in response to the observation of reaching actions to drink than to pour [102,106]. At the single-trial level, information readout may be inferred by isolating variations in movement kinematics that predict variations in the continuous readout variable using statistical approaches such as Partial Least Squares (PLS) regression or Canonical Correlation Analysis (CCA) (see Box 2) and then examining whether these variations encode information.

5. Intersecting kinematic encoding and readout at the single-trial level

Kinematic encoding maps information to movement kinematics. For readout, the direction of probabilistic mapping is inverse: from movement kinematics to information. Yet, kinematic readout is not simply the inverse of encoding. Human perceivers can ignore features that encode information [83] or, conversely, read features that do not encode any information [82]. Further complicating the inverse mapping, the encoded information can be misread [94].

This raises an intriguing question: how and to what extent is information encoded in movement kinematics accurately read out? In this context, we can think of the encoding model as an ideal observer capable of optimally reading all the encoded information (Fig. 3a top panel). The readout computations of this ideal observer describe the optimal computations for mapping movement kinematics to a cognitive state, mirroring the computations of the encoding model. While these computations may not match those of human perceivers, they enable us to predict the limit of behavioural performance given the code, providing an upper bound for readout.

By comparing these optimal computations with those of actual human readout, we can quantify the amount of information successfully read out and transmitted to behaviour. A good reader assigns readout weights similar to the encoding weights (as illustrated in Fig. 3a middle panel; see also Fig. 4f for a real-data example). As a result, the amount of single-trial information that is correctly read out closely matches the amount of information that is encoded and that an ideal observer would read out. Conversely, a bad reader assigns readout weights dissimilar to the encoding weights (Fig. 3a bottom panel). This leads to significant discrepancies between the information read out and the information encoded and available to the ideal observer, potentially leading to incorrect interpretations of the intended action.

In general, the amount of information encoded in kinematics that is read out depends on the intersection between encoding and readout of information. Geometrically, this intersection can be visualized as the alignment between the encoding and readout boundaries for a given set of movements. When these boundaries align, the readout is optimal, and all encoded information is accurately read out. Conversely, when the boundaries misalign, the readout is suboptimal, and the encoded information is either not read out or is read out incorrectly (see Fig. 3c).

A principled approach to quantifying how encoding and readout intersect at the single-trial level is provided by intersection information [51,107]. Intersection information is defined as the amount of information encoded in movement kinematics that is correctly read out by human observers. There are various methods to compute intersection information. The simplest method is to compute the fraction of trials in which a particular state (e.g., intention to pass) is both correctly decoded by the encoding model and correctly read out behaviourally [51]. The larger this fraction, the greater the intersection information. When computed in this way (and, in general, using linear probability models), intersection information is proportional to the alignment between the encoding and readout boundaries [51,108]. Intuitively, the more aligned these two boundaries are, the larger the intersection information. More general definitions of intersection information can be computed using information theory [51,107] (Box 1).

Comparing the amount of intersection information with the amount of encoded information (as measured by the performance of the encoding model) can be useful for disentangling the contribution of encoding and readout to less-than-perfect behavioural performance. For example, it can help determine whether less-than-perfect behavioural performance is due to a moderate amount of information being efficiently read out or a large amount of information being read out inefficiently.

6. Applications and predictions of kinematic coding

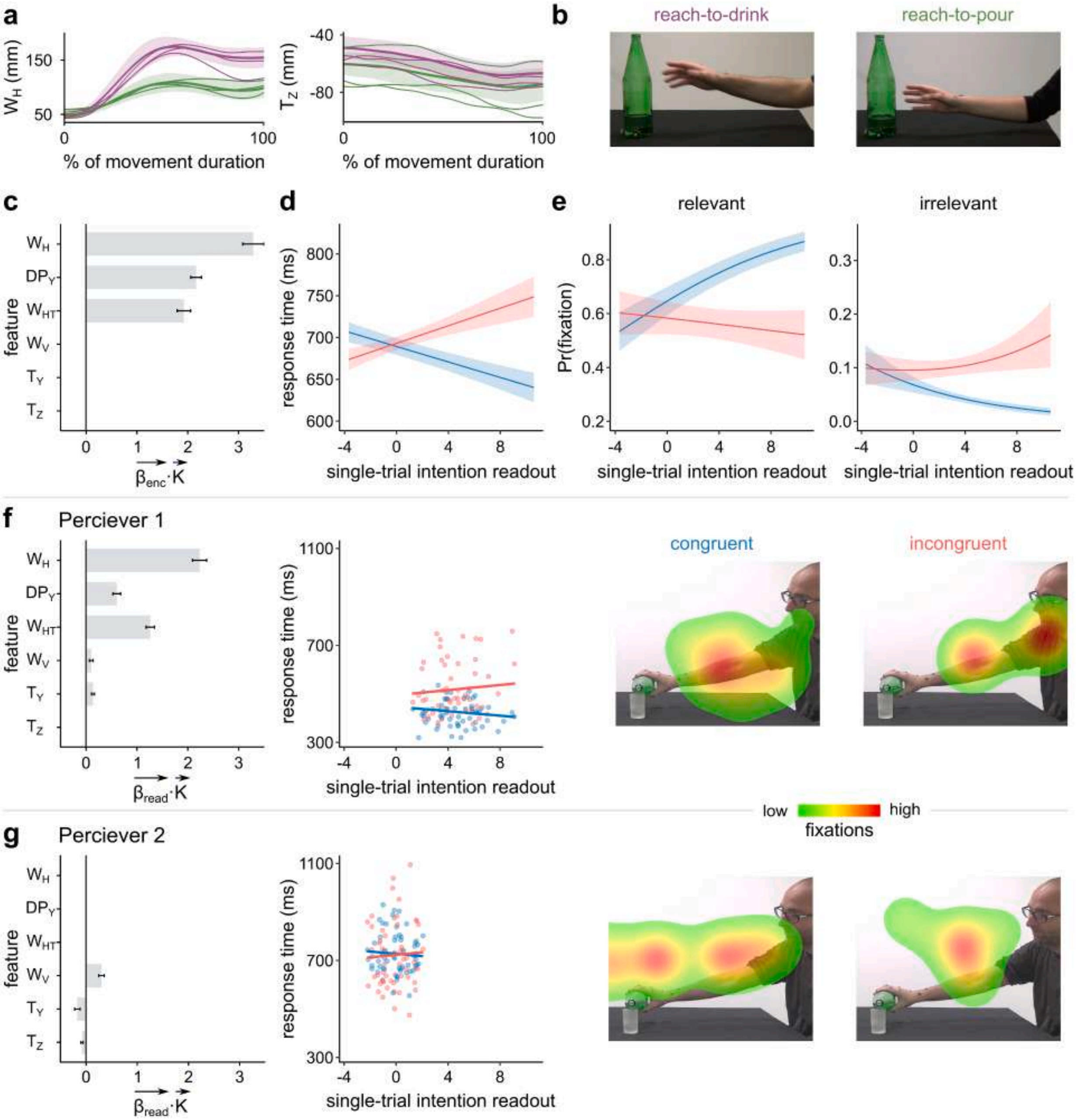
Having established a general kinematic coding framework for studying kinematic encoding, readout and their intersection, we now discuss how it can be applied to generate experimentally testable predictions.

7. Kinematic encoding and readout across multiple orders of planning

Many naturalistic behaviours consist of distinct components – behavioural syllables in animals [109], motor acts in humans [110] – organized sequentially to achieve temporally distant action outcomes. Yet, little is known about the computational principles and neural mechanisms that explain how distinct components are chained together [81]. Kinematic coding may be used to determine chaining, quantified as orders of planning [80], for sequential actions. In object manipulation, first-order planning entails shaping one's grasp in anticipation of the size [111] and shape of the object to be grasped [21]. Second-order planning entails chaining the first and second acts in the sequence by shaping one's object manipulation behaviour on second act [22]. Grasping a bottle either to pass or to drink is an example of second-order planning. Beyond the second order, third-order planning and even fourth-order planning are sometimes possible, for example, when altering one's movement in anticipation of what another person will do next with the object [112–114]. Kinematic encoding may be used to quantify orders of planning in the kinematics of a given act. The logic is straightforward: if in a manipulative action sequence, *n-th* order information is encoded in reaching kinematics, then planning (chaining) can be inferred to have occurred up to the *n-th* order. This approach can be used to quantify higher-order planning across various contexts (e.g., social vs. non-social) and populations (e.g., healthy participants vs. patients). An advantage of this approach is that it only requires analysing the initial act's kinematics, rather than the entire action sequence. This not only simplifies the analysis but also allows for the assessment of residual planning abilities in patients, such as stroke patients, who may be unable to complete a planned action sequence due to sensory or motor function loss. The same logic may be applied to measure action prediction across multiple orders of planning during action observation.

8. Revealing suboptimalities in the readout of encoded information

Commonly used trial-averaged measures such as accuracy (fraction correct), sensitivity, and psychometric functions, offer limited insight into the nature of computations involved in perceptual detection or discrimination of others' actions. Kinematic coding provides a powerful tool to 'open the black box' of trial-averaged discrimination accuracy and dissect sub-components of the readout process. Consider a hypothetical case where, despite perfect encoding, perceptual discrimination remains at or near chance level. This could indicate that the readout process is not sensitive to subtle variations in movement kinematics (kinematic blindness hypothesis). Alternatively, observers might be sensitive to variations in movement kinematics but focus on those that do not convey intention-related information (non-informative feature hypothesis). Lastly, they might focus on informative variations but fail to read the information accurately (misalignment hypothesis). Kinematic coding generates distinct and testable predictions to distinguish among these alternative hypotheses. If readout is not sensitive to variations in movement kinematics (kinematic blindness), no statistical dependency of perceptual choices on single-trial kinematics would be observed. In this case, the performance of the readout model



(caption on next page)

Fig. 4. | **Individual-specific intention readout predicts kinematic priming.** Reanalyses of data first presented in [83] **(a)** Time course of an informative kinematic feature (wrist height, W_H) and a non-informative feature (thumb z-coordinate, T_Z) during reach-to-pour (green) and reach-to-drink (purple) acts. Thin lines represent individual movement trajectories; thick lines and shaded areas depict the mean ± SD across all kinematic prime stimuli. **(b)** The kinematic primes consisted of videos of reach-to-drink and reach-to-pour movements temporally occluded at the end of the reaching phase, preventing participants from observing the second part of the action. **(c)** Contribution of individual kinematic features to encoding, computed as the scalar product between the encoding vector and the single-trial kinematic vector within the subspace of each feature (mean ± SEM across primes). **(d)** Response Times (RTs) as function of single-trial intention readout in a primed action categorization task. Higher readout led to faster RTs in congruent trials and slower RTs in incongruent trials (lines and shaded areas show mean ± SE across trials and perceivers). **(e)** Probability of initial fixation directed toward task-relevant and irrelevant regions of the probe image as a function of single-trial intention readout. In congruent trials, higher readout increased the probability of relevant fixations and decreased the probability of irrelevant fixations. In incongruent trials, the pattern was reversed: higher readout decreased relevant fixations and increased irrelevant fixations. Lines and shaded areas indicate mean ± SE across trials and perceivers. **(f)** Perceiver 1 (good reader). Left: contribution of kinematic features to intention readout shows that this perceiver assigned larger weights to informative features (conventions as in panel c). Middle: RTs plotted against single-trial intention readout. Higher intention readout yielded faster (slower) RTs on congruent (incongruent) trials. Dots represent individual trials with regression lines added. Right: fixation heatmaps for congruent and incongruent trials; the probability of initial fixations being directed to task-relevant regions is higher on congruent trials than on incongruent trials. **(g)** Perceiver 2 (chance reader) Left: contribution of kinematic features to readout shows that this perceiver read mostly non-informative features (conventions as in panel c). Middle: RTs plotted against single-trial intention readout show no significant congruency effect. Right: fixation heatmaps show no effect of congruency.

would be at chance level. If the readout process focuses on non-informative features, the weights in the readout model (e.g., the regression weights of a GLM) would be assigned to features that have zero weights in the encoding model. Finally, if there is incorrect reading of information encoded in informative features (misalignment), the weights in the readout model would correspond with non-zero weights in the encoding model, but the signs of the readout weights would not correctly align with the signs of the encoding weights.

9. Kinematic fingerprinting and beyond

Encoding (and readout) models can be fitted either to data pooled across a cohort of subjects or to individual subjects, with separate models being fitted for each participant. Single-subject fitting may be useful for investigating individual differences. Individuals exhibit variations in motor system outputs that are both consistent within a given individual and differ from one individual to another [84]. By fitting separate encoding models for individual subjects, we can detect individual-specific, idiosyncratic encoding profiles that would remain undetectable at the group level [113,115]. Extending the same logic to readout, we can capture individual-specific readout profiles [83]. The stability of these profiles over time can be tested by recomputing models using data from different sessions (at least two sessions). Using the profile from an individual in one session to discriminate among the profiles from a second session – a procedure dubbed fingerprinting in the study of functional connectomes [116] – provides a measure of identifiability. These kinematic encoding (readout) profiles can then be used predict behaviour, identify cognitive and clinical phenotypes [94,117–120] and generate models that connect behaviour to brain structure/function (e.g., individual-specific functional connectomes [116]). In a recent study, we used this approach to demonstrate that the implicit use of prospective information to anticipate others' actions is directly related to the amount of information that a specific perceiver can extract from the observed action [83] (Fig. 4). Specifically, we found that individual-specific readout computations estimated from a 2AFC task can be used to predict the implicit use of prospective information (as indexed by response times and initial fixations) by the same perceiver in a primed action categorization task performed one hour before [83]. The significance of individual differences in readout for action prediction was highlighted by showing that readout models trained on data from individual perceivers outperformed models trained on surrogate subjects — constructed by pooling data across multiple subjects and then randomly resampling them to create 'average' perceivers.

10. Testing similarity of kinematic codes

One primary advantage of kinematic coding is that encoding and readout of information are computed within a common space. This allows researchers to address meaningful questions about the similarity between encoding and readout codes. For example, a prominent hypothesis states that the same internal models used for action execution are also employed for action observation [121]. Within the framework of kinematic coding, this hypothesis can be tested by initially asking the participant to perform actions under a particular cognitive state and then testing their ability to read the same state actions performed by others. If the same models are used for action execution and observation, then we should expect individuals to read information in the same way they encode it. Montobbio and colleagues [94] used this approach to test the hypothesis that kinematic similarity [122] is important for intention reading in autism spectrum disorders (ASD) and typically developing (TD) observers. Intersecting GLM encoding and readout weights revealed that individuals with ASD extracted information from features that encode intention in ASD-specific kinematics. Similarly, TD observers extracted information from features that encode intention in TD-specific kinematics. This approach can be extended to address questions about information transmission during social interaction. For example, researchers could use an experimental diffusion chain to investigate whether the kinematic similarity of encoding facilitates the transmission of information about a particular state, such as risk [123], along the chain.

11. Tackling brain–behaviour relationships at the single-trial level

Kinematic encoding, kinematic readout, and their intersection are computed at the single-trial level and can be used to directly relate information in movement kinematics to single-trial measures of neural activity, such as electroencephalography (EEG) and magnetoencephalography (MEG) responses. In a perceptual discrimination task, this framework can be used to determine if trial-to-trial fluctuations in brain activity within a particular region correlate with variations in movement kinematics along a specific information axis, such as the intention readout axis, rather than overall variations in movement kinematics. This could help isolate brain regions that implement specific readout computations from those that are responsive to overall variations in movement kinematics. Similar logic could be applied to characterize the information flow between brain regions during motor skill learning [124,125]. Whether used to delineate the information content of specific brain regions or to track interregional communication, the dependencies identified through this approach are statistical, not causal. To infer causality, researchers could employ interventional methods (see Quantifying the effects of intervention).

12. Quantifying the effects of intervention

Evaluating changes in behaviour is a common approach for examining the effects of intervention [126]. For example, manipulation methods (e.g., brain stimulation, pharmacological agents) can be used to up- and down-regulate activity within a brain region to study how changing neural dynamics affects a particular cognitive process (e.g., [127]). One limitation of this approach is that behavioural differences after an intervention (e.g., differences between pre- and post- kinematics, differences in perceptual discrimination performance) may arise from different causes. Kinematic encoding and readout can be used to deconvolve and quantify the specific effect of intervention on these processes. Patri and colleagues [82], for example, combined kinematic readout with theta burst transcranial magnetic stimulation (cTBS) to the inferior parietal lobule (IPL) to evaluate the causal role of IPL in intention reading in a 2AFC task. Fitting readout models to intention choices expressed pre- and post-intervention revealed that transient disruption of IPL did not cause kinematic blindness, or a shift of readout weights towards non-informative kinematic features. Instead, it selectively decreased the alignment between encoding and readout, impairing the ability of observers' to correctly read the extracted information.

13. Outlook

In this article, we introduce a novel approach for understanding naturalistic patterns of behaviour through the information they convey. This framework allows for the generation of new hypotheses and experiments that link behaviour, cognition, and neural activity. Researchers can use this framework to measure how information is encoded in movement kinematics, how it is readout, and the role of these processes in the social transmission of information.

A key insight from kinematic coding is that only information that resides at the intersection of encoding and readout — the information actually read out — can be transmitted. From this perspective, the suboptimality of human readout, as highlighted by the findings presented above, limits the amount of information that can be transmitted. Paradoxically, however, it also creates opportunities for communication. Since readout is not simply the inverse of encoding and not all encoded information is read out, actors can intentionally modify their kinematics to either enhance (as in the case of demonstration [128]) or reduce (as in the case of deception [129]) the transmission of information. This concept forms the basis of sensorimotor communication [130]: humans intentionally alter their movement kinematics to regulate information transmission in social contexts. A widely held but untested assumption is that these communicative alterations influence the informativeness of the performed actions [131,132]. However, little is known about the effects of these alterations on encoding and readout computations. Do communicative alterations increase (or decrease) the amount of encoded information, or do they rather modulate information transmission by enhancing (or reducing) the readability of the encoded information? The kinematic coding framework enables us to adjudicate between these (not mutually exclusive) hypotheses by examining how the imposition of a communicative intention alters the contribution of individual kinematic features to encoding and readout.

Mechanistically, we propose three coding mechanisms by which communication could modulate information transmission: state-related variance, state-unrelated variance, and correlation between kinematic features. Given that only information at the encoding-readout intersection contributes to transmission, the features within this intersection are prime candidates for these mechanisms.

First, communication may operate by amplifying variations in features that carry state-relevant information. For example, when reaching for a bottle with the intent to pour, a person might amplify variations in intention-related features to signal the intention to pour. This amplification increases the readability of the intention, allowing an observer to anticipate and respond, for instance, by bringing their glass closer. This mechanism forms the basis intentional signalling, motionese (i.e., amplification of arm/hand gestures when talking to a child) and other forms of pedagogical demonstration [130]. A second mechanism may involve modulating variance that is orthogonal to the primary informational axis, to convey secondary, non-pragmatic information. For instance, while reaching for a bottle with the intent to pour, a person might alter features unrelated to the pragmatic task of pouring water to communicate politeness or impatience [130]. These first two mechanisms rely on altering specific features. A third mechanism may leverage correlations between features. Studies in neural coding indicate that correlations between different neural activity features can enhance readout. The redundant repetition of information across multiple features ensures that more of it is transmitted downstream in the brain [52,53]. A similar mechanism could apply to kinematics, where correlations among kinematic features could enhance readability and thus influence information transmission.

The idea that individuals alter their kinematics to regulate information transmission has profound implications for sensorimotor control. Traditional theories assume that the objective of motor control is to produce purposeful, coordinated movements. Kinematic readout introduces the possibility that movements are also planned and controlled to facilitate or even hinder readability (as in the case of deception). It seems plausible that this dual function places competing demands on the optimization of motor programs. By allowing for the quantification of encoding and readout at the single-movement level, kinematic coding equips researchers with new tools for addressing this trade-off and computing the single-trial informational costs and benefits of various motor solutions.

In presenting the kinematic coding framework, we assumed for simplicity that the encoding and readout of information in kinematics are time-invariant. Mathematically, this implies encoding or readout models have a fixed set of parameters (e.g. encoding and readout regression weights). However, naturalistic behaviour can evolve over time, such as when learning to perform a new perceptual or motor task or reconsidering a choice mid-execution, often following a structure characterized by multiple timescales [133–135]. Characterizing how these time-varying patterns of behaviour are reflected in changes in kinematic encoding and readout requires models that can learn (even in an unsupervised way) how to adjust their parameters over various timescales as a function of changes in latent variables that capture changes in internal state variables (Fig. 5). These models include input-output Hidden Markov Models, dynamic GLMs, and logistic regressions [136,137], which can adjust encoding weights or state variables across trials. Additionally, methods such as preferential subspace identification can capture within-trial dynamics shared between kinematics and state variables [138]. Extending these approaches to characterize dynamic intersection information could further elucidate how information transmission evolves over time. We anticipate that this extension will open new avenues for research in joint action [139], making kinematic coding a valuable tool for studying real-time interactions and how individuals learn to coordinate their actions with others [128].

Glossary

Canonical Correlation Analysis (CCA): A linear transformation that projects the kinematic data onto an orthogonal basis. In this transformed space, CCA identifies the directions that maximize the correlation between the kinematic data and the state variables. The components are ordered based on the strength of their correlation.

Communication: A cooperative activity where information is deliberately transmitted with the goal of ensuring that the communication partner recognizes this intention.

Encoding boundary: The boundary that optimally separates between movements generated under different cognitive states in the high-dimensional kinematic space.

Encoding model: A mathematical model used to estimate the posterior probability that a movement is performed under a cognitive state based on the single-trial movement kinematics.

Information theory: The most complete and general mathematical theory of communication, quantifying how much information stochastic variables carry about each other. It can capture the information carried by all linear and non-linear relationships between the variables.

Intention: A cognitive state directed at action, operationalized as what one plans to do next with an object in the context of a manipulative action sequence.

Kinematic encoding: The probabilistic mapping from cognitive states to movement kinematics.

Kinematic readout: The probabilistic mapping from movement kinematics to cognitive states.

Kinematic space: High-dimensional space where each dimension represents the value of a certain kinematic feature.

Movement kinematics: Quantification of how physical actions are performed by individuals, focusing on position vectors, velocities and joint angles throughout the trajectory of movements.

Partial Least Squares (PLS): A linear transformation that projects the kinematic data onto an orthogonal basis. In this transformed space, PLS identifies the directions that maximize the covariance between the kinematic data and the state variables. The components are ordered based on the strength of their covariance.

Posterior probability: The posterior probability of a state variable is the conditional probability assigned to the variable given some event.

Principal-component analysis: A linear transformation that projects the kinematic data on to an orthogonal basis. In this transformed space, PCA identifies the directions that maximize the variance in the kinematic data. The components are ordered based on the amount of variance they explain. PCA is commonly used to reduce the dimensionality of complex data while retaining the majority of the variance.

Readout boundary: The boundary that optimally separates between reported observations of different cognitive states in the high-dimensional kinematic space.

Readout model: A mathematical model used to estimate the posterior probability that a movement is reported as being performed under a cognitive state based on the reaching kinematics observed in that trial.

Transmission of information: process by which information encoded in movement kinematics is read out by another agent.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing

Box 1
Concepts of probabilistic framework

The information carried by a single-trial kinematic vector $\overrightarrow{\mathrm{K}}$ about the state Y can be quantified by the posterior probability $P\left(Y|\overrightarrow{\mathrm{K}}\right)$ of the state Y given $\overrightarrow{\mathrm{K}}$, defined by Bayes' rule:

$$\mathrm{P}\left([\mathrm{Y}={}'\mathrm{state}'] \mid \overrightarrow{\mathrm{K}}\right) = \frac{\mathrm{P}\left(\overrightarrow{\mathrm{K}}\middle|\mathrm{Y}\right)}{\mathrm{P}\left(\overrightarrow{\mathrm{K}}\right)}$$

where $P\left(\overrightarrow{\mathrm{K}}\middle|Y\right)$ is the probability that the kinematic vector $\overrightarrow{\mathrm{K}} = [K_1, ..K_n]$ was obtained under state Y. These probabilities are often computed parametrically by fitting an encoding model to the data. For a Generalized Linear Model (GLM) [88], the encoding model can be expressed as:

$$\mathrm{P}\left([\mathrm{Y}={}'\mathrm{state}']\middle|\overrightarrow{\mathrm{K}}\right) = \sigma\left(\overrightarrow{\beta}\cdot\overrightarrow{\mathrm{K}}+\beta_0\right)$$

where σ is the link function (e.g., sigmoid for logistic regression), $\overrightarrow{\beta}$ is the encoding vector containing the regression coefficients of each kinematic feature, and β_0 is the bias, kinematic independent, term. Alternatively, probability estimates can be obtained through more complex parametric models, such as artificial neural networks [140,141], or with non-parametric approaches such as Kernel or Nearest-Neighbours methods [142,143].

The single-trial information can be used for Bayesian decoding to determine the most likely state Y given single-trial kinematics $\overrightarrow{\mathrm{K}}$:

$$\mathrm{Y}^{\mathrm{dec}} = \mathrm{argmax}_{\mathrm{Y}}\ \mathrm{P}\left(\mathrm{Y}\ \middle|\ \overrightarrow{\mathrm{K}}\right)$$

The performance of the encoding model can be quantified by the fraction of correct decoding or by the difference between predicted (Y^{dec}) and actual states (Y). For example, it can be quantified as the mutual information [44,45] derived from the confusion matrix $P\left(Y, Y^{dec}\right)$, which represents the probability of decoding Y^{dec} when Y was actually present and the corresponding marginal probabilities $P(Y)$, $P\left(Y^{dec}\right)$:

$$\mathrm{I}\left(\mathrm{Y};\mathrm{Y}^{\mathrm{dec}}\right) = \sum_{\mathrm{Y},\mathrm{Y}^{\mathrm{dec}}}\mathrm{P}\left(\mathrm{Y},\mathrm{Y}^{\mathrm{dec}}\right)\log_2\left(\frac{\mathrm{P}\left(\mathrm{Y},\mathrm{Y}^{\mathrm{dec}}\right)}{\mathrm{P}(\mathrm{Y})\mathrm{P}\left(\mathrm{Y}^{\mathrm{dec}}\right)}\right)$$

An advantage of using mutual information from the confusion matrix over fraction correct decoding (which considers only the diagonal entries of the confusion matrix) is that mutual information quantifies the total information conveyed by both correct and incorrect classifications, further enabling the capture of information specifically contributed by the error pattern [144,145].

For continuous state variables, performance can be assessed based on the accuracy of the model in estimating the true values of state variables, including mean square errors, variance or deviance explained, or log-likelihoods [146,147]. The contribution of individual kinematic features to encoding can be assessed using reduced encoding models. The contribution of a kinematic feature to encoding across all movements can be computed, for example, as the Shapley value contribution [148], by comparing model performance including and excluding that feature. For a particular movement, the contribution of a kinematic feature to encoding can be quantified as its contribution to the posterior probability of that movement. For GLMs, this contribution has simple geometric interpretations (Fig. 2) and can be computed as the scalar product of the single-trial kinematic vector and the encoding vector in subspace of the considered kinematic feature.

Similarly to the encoding model, the readout model computes the posterior probability $P\left(C|\overrightarrow{\mathrm{K}}\right)$ of the observer reporting a given state C given the single-trial kinematics. The performance of the readout model can be quantified using the same tools described above.

Intersection information quantifies how much of the information about a state, encoded in single-trial kinematics, is read out to inform the observer's choices. From the above probabilities, it can be computed as the fraction of trials in which the state is both correctly encoded and correctly reported by the observer [51]. Alternatively, it can be computed as the fraction of trials in which the state is correctly encoded and correctly read out by the readout model. Intersection information, computed this way, is higher when the encoding and readout vectors are well aligned [51,53,107,108,149].

In information theoretic terms, intersection information can be computed by first constructing a trivariate probability $P\left(Y, C, \overrightarrow{\mathrm{K}}\right)$ of observing state Y and choice C in the presence of a single-trial kinematic vector $\overrightarrow{\mathrm{K}}$. Partial information decompositions are then used to quantify the shared information between state and choice that is contained in the kinematics [107]. Information-theoretic definitions capture all ways in which kinematics can be informative of states (above and beyond simply reporting the most likely alternative as decoders do), but are more laborious and data-hungry than decoding-based measures [45].

Box 2

Decomposing variance vs information in high dimensional kinematics

Kinematic data are high-dimensional, which poses challenges for estimating the probability of movements relative to different cognitive states. To mitigate this issue, dimensionality reduction techniques are often used to map the original feature set into a lower-dimensional space. Unsupervised methods such as Principal Component Analysis (PCA) [150] or Non-negative Matrix Factorization (NMF) [151] are frequently utilized for this purpose. These techniques decompose kinematic data into a lower-dimensional representation that best approximates the original data. For instance, PCA identifies an orthogonal set of (principal) components ordered by the amount of variance they capture in the kinematic data, allowing for accurate reconstruction using a small subset of high-variance components.

However, unsupervised dimensionality reduction methods may overlook components in the kinematic space that, despite capturing only a small portion of the total variance, carry significant information about a particular state. This is because these methods disregard state information during decomposition. Recent findings show that low-variance components can encode substantial information about object properties and intentions in grasping tasks [83,152]. To identify components that are highly informative about a state, even when they exhibit low variance, supervised dimensionality reduction methods can be employed. These methods involve incorporating state labels for each kinematic vector. Partial Least Squares (PLS) [153–155] and Canonical Correlation Analysis (CCA) [156] are notable examples of such approaches. Given a state variable, these methods identify a set of orthogonal components in kinematic space with the maximum covariance (for PLS) or the maximum correlation (for CCA) with the state variable. As illustrated in Fig. 5 for PLS, the component carrying the most state information may not correspond to the PCA component with the largest variance. This highlights the importance of state-informed dimensionality reduction methods such as PLS, where the first component is better aligned with the state variable.

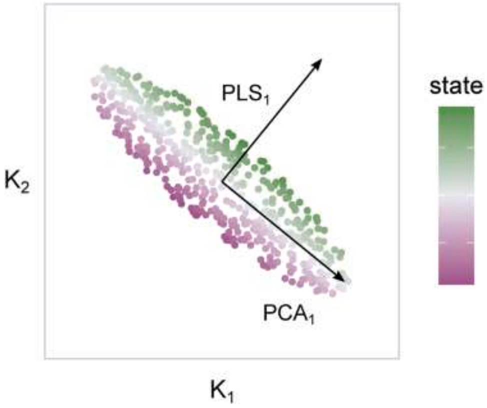


Fig. 5. | **Sensitivity of dimensionality reduction methods to state information.** PLS identifies a set of orthogonal components in kinematic space that maximize covariance with the state variable of interest. In contrast, PCA disregards state information, focusing solely on maximizing variance. Consequently, the PLS component carrying the most state information (PLS$_1$) may not align with the PCA component that captures the largest variance (PCA$_1$). Each dot (coloured by the value of the state variable) represents a movement in a simplified, two-dimensional kinematic space, spanning only two kinematic features, K_1, K_2.

interests:

Becchio Cristina reports financial support was provided by European Commission. I, the corresponding authors, serve in an editorial capacity for the journal to which I am submitting. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Funded by the European Union under Horizon Europe, grant number 101092889, project SHARESPACE. The authors thank members of their laboratories for helpful discussion.

References

[1] Gallivan JP, Chapman CS, Wolpert DM, Flanagan JR. Decision-making in sensorimotor control. Nat Rev Neurosci 2018;19:519–34. https://doi.org/10.1038/s41583-018-0045-9. 2018.

[2] Rosenbaum DA. The Cinderella of psychology: The neglect of motor control in the science of mental life and behavior. Am Psychol 2005;60:308–17. https://doi.org/10.1037/0003-066X.60.4.308 (2005).
[3] Becchio C, Koul A, Ansuini C, Bertone C, Cavallo A. Seeing mental states: An experimental strategy for measuring the observability of other minds. Phys Life Rev 2018;24:67–80. https://doi.org/10.1016/j.plrev.2017.10.002 (2018).
[4] Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron 2017;93:480–90. https://doi.org/10.1016/j.neuron.2016.12.041 (2017).
[5] Shadmehr R, Reppert TR, Summerside EM, Yoon T, Ahmed AA. Movement Vigor as a Reflection of Subjective Economic Utility. Trends Neurosci 2019;42:323–36. https://doi.org/10.1016/j.tins.2019.02.003. 2019.
[6] Datta SR, Anderson DJ, Branson K, Perona P, Leifer A. Computational Neuroethology: A Call to Action. Neuron 2019;104:11–24. https://doi.org/10.1016/j.neuron.2019.09.038. 2019.
[7] Calhoun AJ, Murthy M. Quantifying behavior to solve sensorimotor transformations: advances from worms and flies. Curr Opin Neurobiol 2017;46:90–8. https://doi.org/10.1016/j.conb.2017.08.006. 2017.
[8] Dennis EJ, Hady AE, Michaiel A, Clemens A, Gowan Tervo DR, Voigts J, et al. Systems neuroscience of natural behaviors in rodents. J Neurosci 2021;41:911–9. https://doi.org/10.1523/JNEUROSCI.1877-20.2020 (2021).
[9] Dunn TW, Marshall JD, Severson KS, Aldarondo DE, Hildebrand DGC, Chettih SN, et al. Geometric deep learning enables 3D kinematic profiling across species and environments. Nat Methods 2021;18:564–73. https://doi.org/10.1038/s41592-021-01106-6 (2021).
[10] Lauer J, Zhou M, Ye S, Menegas W, Schneider S, Nath T, et al. Multi-animal pose estimation, identification and tracking with DeepLabCut. Nat Methods 2022;19:496–504. https://doi.org/10.1038/s41592-022-01443-0. 2022.
[11] Goodwin NL, Choong JJ, Hwang S, Pitts K, Bloom L, Islam A, et al. Simple Behavioral Analysis (SimBA) as a platform for explainable machine learning in behavioral neuroscience. Nat Neurosci 2024;27:1411–24. https://doi.org/10.1038/s41593-024-01649-9. 2024.
[12] Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 2018;21:1281–9. https://doi.org/10.1038/s41593-018-0209-y. 2018.
[13] Pereira TD, Aldarondo DE, Willmore L, Kislin M, Wang SSH, Murthy M, et al. Fast animal pose estimation using deep neural networks. Nat Methods 2018;16:117–25. https://doi.org/10.1038/s41592-018-0234-5. 2018.
[14] Pereira TD, Tabris N, Matsliah A, Turner DM, Li J, Ravindranath S, et al. SLEAP: A deep learning system for multi-animal pose tracking. Nat Methods 2022;19:486–95. https://doi.org/10.1038/s41592-022-01426-1. 2022.
[15] Storchi R, Milosavljevic N, Allen AE, Zippo AG, Agnihotri A, Cootes TF, et al. A High-Dimensional Quantification of Mouse Defensive Behaviors Reveals Enhanced Diversity and Stimulus Specificity. Curr Biol 2020;30:4619–30. https://doi.org/10.1016/j.cub.2020.09.007. e52020.
[16] Sani OG, Pesaran B, Shanechi MM. Dissociative and prioritized modeling of behaviorally relevant neural dynamics using recurrent neural networks. Nat Neurosci 2024;27:2033–45. https://doi.org/10.1038/s41593-024-01731-2. 2024.
[17] Schneider S, Lee JH, Mathis MW. Learnable latent embeddings for joint behavioural and neural analysis. Nature 2023;617:360–8. https://doi.org/10.1038/s41586-023-06031-6. 2023.
[18] Maselli A, Gordon J, Eluchans M, Lancia GL, Thiery T, Moretti R, et al. Beyond simple laboratory studies: Developing sophisticated models to study rich behavior. Phys Life Rev 2023;46:220–44. https://doi.org/10.1016/j.plrev.2023.07.006. 2023.
[19] Mobbs D, Wise T, Suthana N, Guzman N, Kriegeskorte N, Leibo JZ. Promises and challenges of human computational ethology. Neuron 2021;109:2224–38. https://doi.org/10.1016/j.neuron.2021.05.021. 2021.
[20] Song JH, Nakayama K. Hidden cognitive states revealed in choice reaching tasks. Trends Cogn Sci 2009;13:360–6. https://doi.org/10.1016/j.tics.2009.04.009. 2009.
[21] Santello M, Soechting JF. Gradual molding of the hand to object contours. J Neurophysiol 1998;79:1307–20. https://doi.org/10.1152/jn.1998.79.3.1307. 1998.
[22] Cavallo A, Koul A, Ansuini C, Capozzi F, Becchio C. Decoding intentions from movement kinematics. Sci Rep 2016;6:37036. https://doi.org/10.1038/srep37036. 2016.
[23] Lewkowicz D, Quesque F, Coello Y, YN Delevoye-Turrell. Individual differences in reading social intentions from motor deviants. Front Psychol 2015;6:1175. https://doi.org/10.3389/fpsyg.2015.01175. 2015.
[24] Podda J, Ansuini C, Vastano R, Cavallo A, Becchio C. The heaviness of invisible objects: Predictive weight judgments from observed real and pantomimed grasps. Cognition 2017;168:140–5. https://doi.org/10.1016/j.cognition.2017.06.023. 2017.
[25] van der Wel RPRD, Sebanz N, Knoblich G. Do people automatically track others' beliefs? Evidence from a continuous measure. Cognition 2014;130:128–33. https://doi.org/10.1016/j.cognition.2013.10.004. 2014.
[26] Zani G, Butterfill SA, Low J. Mindreading by body: incorporating mediolateral balance and mouse-tracking measures to examine the motor basis of adults' false-belief tracking. Royal Society Open Science 2023;10:221212. https://doi.org/10.1098/rsos.221212. 2023.
[27] Zani G, Butterfill SA, Low J. Mindreading in the balance: adults' mediolateral leaning and anticipatory looking foretell others' action preparation in a false-belief interactive task. Royal Society Open Science 2020;7:191167. https://doi.org/10.1098/rsos.191167. 2020.
[28] Freeman JB, Pauker K, Sanchez DT. A Perceptual Pathway to Bias: Interracial Exposure Reduces Abrupt Shifts in Real-Time Race Perception That Predict Mixed-Race Bias. Psychol Sci 2016;27:502–17. https://doi.org/10.1177/0956797615627418. 2016.
[29] Chapman CS, Gallivan JP, Wong JD, Wispinski NJ, Enns JT. The snooze of lose: Rapid reaching reveals that losses are processed more slowly than gains. J Exp Psychol Gen 2015;144:844–63. https://doi.org/10.1037/xge0000085. 2015.
[30] Patel D, Fleming SM, Kilner JM. Inferring subjective states through the observation of actions. Proc R Soc Lond B Biol Sci 2012;279:4853–60. https://doi.org/10.1098/rspb.2012.1847. 2012.
[31] Summerside EM, Shadmehr R, Ahmed AA. Vigor of reaching movements: reward discounts the cost of effort. J Neurophysiol 2018;119:2347–57. https://doi.org/10.1152/jn.00872.2017. 2018.
[32] Barca L, Pezzulo G. Tracking Second Thoughts: Continuous and Discrete Revision Processes during Visual Lexical Decision. PLoS One 2015;10:e0116193. https://doi.org/10.1371/journal.pone.0116193. 2015.
[33] Resulaj A, Kiani R, Wolpert DM, Shadlen MN. Changes of mind in decision-making. Nature 2009;461:263–6. https://doi.org/10.1038/nature08275. 2009.
[34] Dotan D, Meyniel F, Dehaene S. On-line confidence monitoring during decision making. Cognition 2018;171:112–21. https://doi.org/10.1016/j.cognition.2017.11.001. 2018.
[35] Miller GA. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychol Rev 1956;63:81–97. https://doi.org/10.1037/h0043158. 1956.
[36] Baddeley A. Working Memory. Science 1992;255:556–9. https://doi.org/10.1126/science.1736359. 1992.
[37] Tononi G, Boly M, Massimini M, Koch C. Integrated information theory: from consciousness to its physical substrate. Nat Rev Neurosci 2016;17:450–61. https://doi.org/10.1038/nrn.2016.44. 2016.
[38] Perkel DH, Bullock TH. Neural coding. Neurosci Res Program Bull 1968;6:221–348. 1968.
[39] Searle JR. Minds, brains, and programs. Behav Brain Sci 2010;3:417–24. https://doi.org/10.1017/s0140525×00005756. 2010.
[40] MacKay DM, McCulloch WS. The limiting information capacity of a neuronal link. Bull Math Biophys 1952;14:127–35. https://doi.org/10.1007/bf02477711. 1952.
[41] Cantlon JF, Piantadosi ST. Uniquely human intelligence arose from expanded information capacity. Nat Rev Psy 2024;3:275–93. https://doi.org/10.1038/s44159-024-00283-3. 2024.
[42] Latash ML. Neurophysiological Basis of Movement. Human Kinetics Publishers; 1998.
[43] Logothetis NK. What we can do and what we cannot do with fMRI. Nature 2008;453:869–78. https://doi.org/10.1038/nature06976. 2008.
[44] Shannon CE. A Mathematical Theory of Communication. Bell Syst Tech J 1948;27:379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x. 1948.

[45] Quian Quiroga R, Panzeri S. Extracting information from neuronal populations: information theory and decoding approaches. Nat Rev Neurosci 2009;10: 173–85. https://doi.org/10.1038/nrn2578. 2009.

[46] Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D. Reading a Neural Code. Science 1991;252:1854–7. https://doi.org/10.1126/science.2063199. 1991.

[47] Borst A, Theunissen FE. Information theory and neural coding. Nat Neurosci 1999;2:947–57. https://doi.org/10.1038/14731. 1999.

[48] Pouget A, Dayan P, Zemel R. Information processing with population codes. Nature Reviews Neuroscience 2000;1:125–32. https://doi.org/10.1038/35039062. 2000.

[49] Luna R, Hernández A, Brody CD, Romo R. Neural codes for perceptual discrimination in primary somatosensory cortex. Nat Neurosci 2005;8:1210–9. https://doi.org/10.1038/nn1513. 2005.

[50] Salinas E, Sejnowski TJ. Correlated neuronal activity and the flow of neural information. Nat Rev Neurosci 2001;2:539–50. https://doi.org/10.1038/35086012. 2001.

[51] Panzeri S, Harvey CD, Piasini E, Latham PE, Fellin T. Cracking the Neural Code for Sensory Perception by Combining Statistics, Intervention, and Behavior. Neuron 2017;93:491–507. https://doi.org/10.1016/j.neuron.2016.12.036. 2017.

[52] Panzeri S, Moroni M, Safaai H, Harvey CD. The structures and functions of correlations in neural population codes. Nat Rev Neurosci 2022;23:551–67. https://doi.org/10.1038/s41583-022-00606-4. 2022.

[53] Valente M, Pica G, Bondanelli G, Moroni M, Runyan CA, Morcos AS, et al. Correlations enhance the behavioral readout of neural population activity in association cortex. Nat Neurosci 2021;24:975–86. https://doi.org/10.1038/s41593-021-00845-1. 2021.

[54] Müller V, Perdikis D, Mende MA, Lindenberger U. Interacting brains coming in sync through their minds: an interbrain neurofeedback study. Ann N Y Acad Sci 2021;1500:48–68. https://doi.org/10.1111/nyas.14605. 2021.

[55] Hamilton AFdC. Hyperscanning: Beyond the Hype. Neuron 2021;109:404–7. https://doi.org/10.1016/j.neuron.2020.11.008. 2021.

[56] Seideman JA, Stanford TR, Salinas E. Saccade metrics reflect decision-making dynamics during urgent choices. Nat Comm 2018;9:2907. https://doi.org/10.1038/s41467-018-05319-w. 2018.

[57] Salinas E, Steinberg BR, Sussman LA, Fry SM, Hauser CK, Anderson DD, et al. Voluntary and involuntary contributions to perceptually guided saccadic choices resolved with millisecond precision. eLife 2019;8:e46359. https://doi.org/10.7554/eLife.46359. 2019.

[58] Goldstein AT, Stanford TR, Salinas E. Coupling of saccade plans to endogenous attention during urgent choices. bioRxiv 2024. https://doi.org/10.1101/2024.03.01.583058. 202403015830582024.

[59] Rotman G, Troje NF, Johansson RS, Flanagan JR. Eye Movements When Observing Predictable and Unpredictable Actions. J Neurophysiol 2006;96:1358–69. https://doi.org/10.1152/jn.00227.2006. 2006.

[60] Ambrosini E, Pezzulo G, Costantini M. The eye in hand: predicting others' behavior by integrating multiple sources of information. J Neurophysiol 2015;113: 2271–9. https://doi.org/10.1152/jn.00464.2014. 2015.

[61] Vaziri-Pashkam M, Cormiea S, Nakayama K. Predicting actions from subtle preparatory movements. Cognition 2017;168:65–75. https://doi.org/10.1016/j.cognition.2017.06.014. 2017.

[62] Ansuini C, Cavallo A, Koul A, D'Ausilio A, Taverna L, Becchio C. Grasping others' movements: Rapid discrimination of object size from observed hand movements. J Exp Psychol Hum Percept Perform 2016;42:918–29. https://doi.org/10.1037/xhp0000169. 2016.

[63] Ansuini C, Santello M, Massaccesi S, Castiello U. Effects of End-Goal on Hand Shaping. J Neurophysiol 2005;95:2456–65. https://doi.org/10.1152/jn.01107.2005. 2005.

[64] Lukos JR, Ansuini C, Santello M. Anticipatory Control of Grasping: Independence of Sensorimotor Memories for Kinematics and Kinetics. J Neurosci 2009;28: 12765–74. https://doi.org/10.1523/JNEUROSCI.4335-08.2008. 2009.

[65] Santello M, Flanders M, Soechting JF. Patterns of Hand Motion during Grasping and the Influence of Sensory Guidance. J Neurosci 2002;22:1426–35. https://doi.org/10.1523/jneurosci.22-04-01426.2002. 2002.

[66] Winges SA, Weber DJ, Santello M. The role of vision on hand preshaping during reach to grasp. Exp Brain Res 2003;152:489–98. https://doi.org/10.1007/s00221-003-1571-9. 2003.

[67] Stapel JC, Hunnius S, Meyer M, Bekkering H. Motor system contribution to action prediction: Temporal accuracy depends on motor experience. Cognition 2016;148:71–8. https://doi.org/10.1016/j.cognition.2015.12.007. 2016.

[68] Summerside EM, Courter RJ, Shadmehr R, Ahmed AA. Slowing of movements in healthy aging as a rational economic response to an elevated effort landscape. J Neurosci 2024;44:e1596232024. https://doi.org/10.1523/jneurosci.1596-23.2024. 2024.

[69] Zhao Q, Wang Y, Lu Y, Luan M, Zhou C. Computational Bases of Action Anticipation Superiority in Experts: Identifying and Mapping Kinematic Invariants. ssrn 2023 doi;4591516. https://doi.org/10.2139/ssrn.4591516. 2023.

[70] Abernethy B, Zawi K, Jackson RC. Expertise and attunement to kinematic constraints. Perception 2008;37:931–48. https://doi.org/10.1068/p5340. 2008.

[71] Müller S, Abernethy B, Farrow D. How do world-class cricket batsmen anticipate a bowler's intention? Q J Exp Psychol 2006;59:2162–86. https://doi.org/10.1080/02643290600576595. 2006.

[72] Park SH, Ryu D, Uiga L, Masters R, Abernethy B, Mann DL. Falling for a Fake: The Role of Kinematic and Non-kinematic Information in Deception Detection. Perception 2019;48:330–7. https://doi.org/10.1177/0301006619837874. 2019.

[73] Aglioti SM, Cesari P, Romani M, Urgesi C. Action anticipation and motor resonance in elite basketball players. Nat Neurosci 2008;11:1109–16. https://doi.org/10.1038/nn.2182. 2008.

[74] Tomeo E, Cesari P, Aglioti SM, Urgesi C. Fooling the kickers but not the goalkeepers: Behavioral and neurophysiological correlates of fake action detection in soccer. Cereb Cortex 2013;23:2765–78. https://doi.org/10.1093/cercor/bhs279. 2013.

[75] Urgesi C, Savonitto MM, Fabbro F, Aglioti SM. Long- and short-term plastic modeling of action prediction abilities in volleyball. Psychol Res 2012;76:542–60. https://doi.org/10.1007/s00426-011-0383-y. 2012.

[76] Abreu AM, Macaluso E, Azevedo RT, Cesari P, Urgesi C, Aglioti SM. Action anticipation beyond the action observation network: A functional magnetic resonance imaging study in expert basketball players. Eur J Neurosci 2012;35:1646–54. https://doi.org/10.1111/j.1460-9568.2012.08104.x. 2012.

[77] O'Reilly D, Delis I. Dissecting muscle synergies in the task space. eLife 2024;12:RP87651. https://doi.org/10.7554/eLife.87651.4. 2024.

[78] Farina D, Merletti R, Enoka RM. The extraction of neural strategies from the surface EMG. J Appl Physiol 2004;96:1486–95. https://doi.org/10.1152/japplphysiol.01070.2003. 2004.

[79] d'Avella A, Saltiel P, Bizzi E. Combinations of muscle synergies in the construction of a natural motor behavior. Nat Neurosci 2003;6:300–8. https://doi.org/10.1038/nn1010. 2003.

[80] Rosenbaum DA, Chapman KM, Weigelt M, Weiss DJ, van der Wel R. Cognition, action, and object manipulation. Psychol Bull 2012;138:924–46. https://doi.org/10.1037/a0027839. 2012.

[81] Grafton ST. The cognitive neuroscience of prehension: Recent developments. Exp Brain Res 2010;204:475–91. https://doi.org/10.1007/s00221-010-2315-2. 2010.

[82] Patri JF, Cavallo A, Pullar K, Soriano M, Valente M, Koul A, et al. Transient Disruption of the Inferior Parietal Lobule Impairs the Ability to Attribute Intention to Action. Curr Biol 2020;30:4594–605. https://doi.org/10.1016/j.cub.2020.08.104. e72020.

[83] Scaliti E, Pullar K, Borghini G, Cavallo A, Panzeri S, Becchio C. Kinematic priming of action predictions. Curr Biol 2023;33:2717–27. https://doi.org/10.1016/j.cub.2023.05.055. e62023.

[84] Ting LH, Chiel HJ, Trumbower RD, Allen JL, McKay JL, Hackney ME, et al. Neuromechanical principles underlying movement modularity and their implications for rehabilitation. Neuron 2015;86:38–54. https://doi.org/10.1016/j.neuron.2015.02.042. 2015.

[85] Latash ML. The bliss (not the problem) of motor abundance (not redundancy). Exp Brain Res 2012;217:1–5. https://doi.org/10.1007/s00221-012-3000-4. 2012.

[86] Bzdok D, Yeo BTT. Inference in the age of big data: Future perspectives on neuroscience. NeuroImage 2017;155:549–64. https://doi.org/10.1016/j.neuroimage.2017.04.061. 2017.
[87] Drucker H, Burges CJC, Kaufman L, Smola A, Vapnik V. Support Vector Regression Machines. Adv Neural Inf Process Syst 1997;9:155–61. 1997.
[88] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33:1–22. https://doi.org/10.18637/jss.v033.i01. 2010.
[89] Padoa-Schioppa C. Primer Logistic analysis of choice data: A primer. Neuron 2022;110:1615–30. https://doi.org/10.1016/j.neuron.2022.03.002. 2022.
[90] Delis I, Berret B, Pozzo T, Panzeri S. A methodology for assessing the effect of correlations among muscle synergy activations on task-discriminating information. Front Comput Neurosci 2013;7:54. https://doi.org/10.3389/fncom.2013.00054. 2013.
[91] Chih-Wei H, Chih-Jen L. A comparison of methods for multiclass support vector machines. IEEE Trans Neural Netw 2002;13:415–25. https://doi.org/10.1109/72.991427. 2002.
[92] Lorena AC, de Carvalho ACPLF, Gama JMP. A review on the combination of binary classifiers in multiclass problems. Artif Intell Rev 2009;30:19. https://doi.org/10.1007/s10462-009-9114-9. 2009.
[93] Engel J. Polytomous logistic regression. Stat Neerl 1988;42:233–52. https://doi.org/10.1111/j.1467-9574.1988.tb01238.x. 1988.
[94] Montobbio N, Cavallo A, Albergo D, Ansuini C, Battaglia F, Podda J, et al. Intersecting kinematic encoding and readout of intention in autism. Proc Natl Acad Sci U S A 2022;119:e2114648119. https://doi.org/10.1073/pnas.2114648119. 2022.
[95] Farrow D, Abernethy B, Jackson RC. Probing expert anticipation with the temporal occlusion paradigm: experimental investigations of some methodological issues. Motor Control 2005;9:332–51. https://doi.org/10.1123/mcj.9.3.330. 2005.
[96] Koul A, Soriano M, Tversky B, Becchio C, Cavallo A. The kinematics that you do not expect : Integrating prior information and kinematics towards intention choice. Cognition 2019;182:213–9. https://doi.org/10.1016/j.cognition.2018.10.006. 2019.
[97] Tversky B. Sometimes we can see some mental states. Phys Life Rev 2018;24:88–90. https://doi.org/10.1016/j.plrev.2017.11.017. 2018.
[98] Gordon J, Maselli A, Lancia GL, Thiery T, Cisek P, Pezzulo G. The road towards understanding embodied decisions. Neurosci Biobehav Rev 2021;131:722–36. https://doi.org/10.1016/j.neubiorev.2021.09.034. 2021.
[99] Flanagan JR, Johansson RS. Action plans used in action observation. Nature 2003;424:769–71. https://doi.org/10.1038/nature01861. 2003.
[100] Flanagan RJ, Rotman G, Reichelt AF, Johansson RS. The role of observers' gaze Behaviour when watching object manipulation tasks: Predicting and evaluating the consequences of action. Philos Trans R Soc Lond B Biol Sci 2013;368:20130063. https://doi.org/10.1098/rstb.2013.0063. 2013.
[101] Houston J, Masters RE. The experimental induction of religious-type experiences. Highest state of consciousness 1972:303–21. 1972.
[102] Soriano M, Cavallo A, D'Ausilio A, Becchio C, Fadiga L. Movement kinematics drive chain selection toward intention detection. Proc Natl Acad Sci U S A 2018;115:10452–7. https://doi.org/10.1073/pnas.1809825115. 2018.
[103] Selen LPJ, Shadlen MN, Wolpert DM. Deliberation in the Motor System: Reflex Gains Track Evolving Evidence Leading to a Decision. J Neurosci 2012;32:2276–86. https://doi.org/10.1523/jneurosci.5273-11.2012. 2012.
[104] Cattaneo L, Fabbri-Destro M, Boria S, Pieraccini C, Cossu G, Rizzolatti G, et al. Impairment of actions chains in autism and its possible role in intention understanding. Proc Natl Acad Sci U S A 2007;104:11783–7825. https://doi.org/10.1073/pnas.0706273104. 2007.
[105] Giese MA, Rizzolatti G. Neural and Computational Mechanisms of Action Processing: Interaction between Visual and Motor Representations. Neuron 2015;88:167–80. https://doi.org/10.1016/j.neuron.2015.09.040. 2015.
[106] Di Dio C, Miraglia L, Peretti G, Marchetti A, Rizzolatti G. Actions chains and intention understanding in 3- to 6-year-old children. Proc Natl Acad Sci U S A 2024;121:e2317653121. https://doi.org/10.1073/pnas.2317653121. 2024.
[107] Pica G, Piasini E, Safaai H, Runyan CA, Diamond ME, Fellin T, et al. Quantifying how much sensory information in a neural code is relevant for behavior. Adv Neural Inf Process Syst 2017;30:3686–96. 2017.
[108] Yates JL, Park IM, Katz LN, Pillow JW, Huk AC. Functional dissection of signal and noise in MT and LIP during decision-making. Nat Neurosci 2017;20:1285–92. https://doi.org/10.1038/nn.4611. 2017.
[109] Markowitz JE, Gillis WF, Beron CC, Neufeld SQ, Robertson K, Bhagat ND, et al. The Striatum Organizes 3D Behavior via Moment-to-Moment Action Selection. Cell 2018;174:44–58. https://doi.org/10.1016/j.cell.2018.04.019. .e172018.
[110] Flanagan JR, Bowman MC, Johansson RS. Control strategies in object manipulation tasks. Curr Opin Neurobiol 2006;16:650–9. https://doi.org/10.1016/j.conb.2006.10.005. 2006.
[111] Ansuini C, Cavallo A, Koul A, Jacono M, Yang Y, Becchio C. Predicting object size from hand kinematics: A temporal perspective. PLoS One 2015;10:1–13. https://doi.org/10.1371/journal.pone.0120432. 2015.
[112] Ansuini C, Podda J, Battaglia FM, Veneselli E, Becchio C. One hand, two hands, two people: Prospective sensorimotor control in children with autism. Dev Cogn Neurosci 2018;29:86–96. https://doi.org/10.1016/j.dcn.2017.02.009. 2018.
[113] Cavallo A, Romeo L, Ansuini C, Podda J, Battaglia F, Veneselli E, et al. Prospective motor control obeys to idiosyncratic strategies in autism. Sci Rep 2018;8:1–9. https://doi.org/10.1038/s41598-018-31479-2. 2018.
[114] Meyer M, Van Der Wel RPRD, Hunnius S. Higher-order action planning for individual and joint object manipulations. Exp Brain Res 2013;225:579–88. https://doi.org/10.1007/s00221-012-3398-8. 2013.
[115] Turri G, Cavallo A, Romeo L, Pontil M, Sanfey A, Panzeri S, et al. Decoding social decisions from movement kinematics. iScience 2022;25:105550. https://doi.org/10.1016/j.isci.2022.105550. 2022.
[116] Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci 2015;18:1664–71. https://doi.org/10.1038/nn.4135. 2015.
[117] Cavallo A, Romeo L, Ansuini C, Battaglia F, Nobili L, Pontil M, et al. Identifying the signature of prospective motor control in children with autism. Sci Rep 2021;11:3165. https://doi.org/10.1038/s41598-021-82374-2. 2021.
[118] Cook JL, Blakemore SJ, Press C. Atypical basic movement kinematics in autism spectrum conditions. Brain 2013;136:2816–24. https://doi.org/10.1093/brain/awt208. 2013.
[119] Cook J, Swapp D, Pan X, Bianchi-Berthouze N, Blakemore SJ. Atypical interference effect of action observation in autism spectrum conditions. Psychol Med 2014;44:731–40. https://doi.org/10.1017/S0033291713001335. 2014.
[120] Montobbio N, Zingarelli E, Folesani F, Memeo M, Croce E, Cavallo A, et al. Action prediction in psychosis. Schizophrenia 2024;10:8. https://doi.org/10.1038/s41537-023-00429-x. 2024.
[121] Wolpert DM, Doya K, Kawato M. A unifying computational framework for motor control and social interaction. Philos Trans R Soc Lond B Biol Sci 2003;358:593–602. https://doi.org/10.1098/rstb.2002.1238. 2003.
[122] Cook J. From movement kinematics to social cognition : The case of autism From movement kinematics to social cognition : the case of autism. Philos Trans R Soc Lond B Biol Sci 2018;371:20150372. https://doi.org/10.1098/rstb.2015.0372. 2018.
[123] Moussaïd M, Brighton H, Gaissmaier W. The amplification of risk in experimental diffusion chains. Proc Natl Acad Sci U S A 2015;112:5631–6. https://doi.org/10.1073/pnas.1421883112. 2015.
[124] Lemke SM, Celotto M, Maffulli R, Ganguly K, Panzeri S. Information flow between motor cortex and striatum reverses during skill learning. Curr Biol 2024;34:1831–43. https://doi.org/10.1016/j.cub.2024.03.023. e72024.
[125] Celotto M, Bím J, Tlaie A, De Feo V, Toso A, Lemke S, et al. An information-theoretic quantification of the content of communication between brain regions. Adv Neural Inf Process Syst 2023;36:64213–65. 2023.
[126] Vaidya AR, Pujara MS, Petrides M, Murray EA, Fellows LK. Lesion Studies in Contemporary Neuroscience. Trends Cogn Sci 2019;23:653–71. https://doi.org/10.1016/j.tics.2019.05.009. 2019.
[127] Tunik E, Frey SH, Grafton ST. Virtual lesions of the anterior intraparietal area disrupt goal-dependent on-line adjustments of grasp. Nat Neurosci 2005;8:505–11. https://doi.org/10.1038/nn1430. 2005.

[128] McEllin L, Knoblich G, Sebanz N. Distinct kinematic markers of demonstration and joint action coordination? Evidence from virtual xylophone playing. J Exp Psychol Hum Percept Perform 2018;44:885–97. https://doi.org/10.1037/xhp0000505. 2018.
[129] Sebanz N, Shiffrar M. Detecting deception in a bluffing body: The role of expertise. Psychon Bull Rev 2009;16:170–5. https://doi.org/10.3758/PBR.16.1.170. 2009.
[130] Pezzulo G, Donnarumma F, Dindo H, D'Ausilio A, Konvalinka I, Castelfranchi C. The body talks: Sensorimotor communication and its brain and kinematic signatures. Phys Life Rev 2019;28:1–21. https://doi.org/10.1016/j.plrev.2018.06.014. 2019.
[131] McEllin L, Sebanz N, Knoblich G. Identifying others' informative intentions from movement kinematics. Cognition 2018;180:246–58. https://doi.org/10.1016/j.cognition.2018.08.001. 2018.
[132] Strachan JWA, Curioni A, Constable MD, Knoblich G, Charbonneau M. Evaluating the relative contributions of copying and reconstruction processes in cultural transmission episodes. PLoS One 2021;16:e0256901. https://doi.org/10.1371/journal.pone.0256901. 2021.
[133] Calhoun AJ, Pillow JW, Murthy M. Unsupervised identification of the internal states that shape natural behavior. Nat Neurosci 2019;12:2040–9. https://doi.org/10.1038/s41593-019-0533-x. 2019.
[134] Ashwood ZC, Roy NA, Stone IR, Urai AE, Churchland AK, Pouget A, et al. Mice alternate between discrete strategies during perceptual decision-making. Nat Neurosci 2022;25:201–12. https://doi.org/10.1038/s41593-021-01007-z. 2022.
[135] Johnson RE, Linderman S, Panier T, Wee CL, Song E, Herrera KJ, et al. Probabilistic Models of Larval Zebrafish Behavior Reveal Structure on Many Scales. Curr Biol 2020;30:70–82. https://doi.org/10.1016/j.cub.2019.11.026. e42020.
[136] Roy NA, Bak JH, Akrami A, Brody CD, Pillow JW. Efficient inference for time-varying behavior during learning. Adv Neural Inf Process Syst 2018;31:5695–705. 2018.
[137] Bengio Y, Frasconi P. An input output HMM architecture. In: Tesauro G, Touretzky DS, Leen TK, editors. Adv Neural Inf Process Syst. MIT Press; 1995. p. 427–34.
[138] Sani OG, Abbaspourazad H, Wong YT, Pesaran B, Shanechi MM. Modeling behaviorally relevant neural dynamics enabled by preferential subspace identification. Nat Neurosci 2020;24:140–9. https://doi.org/10.1038/s41593-020-00733-0. 2020.
[139] Sebanz N, Bekkering H, Knoblich G. Joint action: Bodies and minds moving together. Trends Cogn Sci 2006;10:70–6. https://doi.org/10.1016/j.tics.2005.12.009. 2006.
[140] van Gerven MAJ. A primer on encoding models in sensory neuroscience. J Math Psychol 2017;76:172–83. https://doi.org/10.1016/j.jmp.2016.06.009. 2017.
[141] Kriegeskorte N. Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. Annu Rev Vis Sci 2015;1:417–46. https://doi.org/10.1146/annurev-vision-082114-035447. 2015.
[142] Hofmann T, Schölkopf B, Smola AJ. Kernel methods in machine learning. Ann Stat 2008;36:1171–220. https://doi.org/10.1214/009053607000000677. 2008.
[143] Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Phys Rev E 2004;69:066138. https://doi.org/10.1103/PhysRevE.69.066138. 2004.
[144] Treves A. On the perceptual structure of face space. Biosystems 1997;40:189–96. https://doi.org/10.1016/0303-2647(96)01645-0. 1997.
[145] Strachan JWA, Pansardi O, Scaliti E, Celotto M, Saxena K, Yi C, et al. GPT-4o reads the mind in the eyes. arXiv:241022309v2 2024 doi:10.48550/arXiv.2410.22309 (2024).
[146] McCullagh P, Nelder JA. Generalized Linear Models, Second Edition: Chapman & Hall; 1989.
[147] Agresti A. An Introduction to Categorical Data Analysis. Wiley; 2013. Third Edition.
[148] Covert IC, Lundberg S, Lee S-I. Explaining by removing: A unified framework for model explanation. J Mach Learn Res 2021;22:1–90. 2021.
[149] Zhao Y, Yates JL, Levi AJ, Huk AC, Park IM. Stimulus-choice (mis)alignment in primate area MT. PLoS Comp Biol 2020;16:e1007614. https://doi.org/10.1371/journal.pcbi.1007614. 2020.
[150] Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. Philos Transact A Math Phys Eng Sci 2016;374:20150202. https://doi.org/10.1098/rsta.2015.0202. 2016.
[151] Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. Nature 1999;401:788–91. https://doi.org/10.1038/44565. 1999.
[152] Yan Y, Goodman JM, Moore DD, Solla SA, Bensmaia SJ. Unexpected complexity of everyday manual behaviors. Nat Comm 2020;11:3564. https://doi.org/10.1038/s41467-020-17404-0. 2020.
[153] Rosipal R, Krämer N. Overview and Recent Advances in Partial Least Squares. In: Saunders C, Grobelnik M, Gunn S, Shawe-Taylor J, editors. Subspace, Latent Structure and Feature Selection SLSFS 2005. Heidelberg: Springer Berlin; 2006. p. 34–51.
[154] McIntosh AR, Lobaugh NJ. Partial least squares analysis of neuroimaging data: applications and advances. NeuroImage 2004;23. https://doi.org/10.1016/j.neuroimage.2004.07.020. S250-S632004.
[155] Krishnan A, Williams LJ, McIntosh AR, Abdi H. Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. NeuroImage 2011;56:455–75. https://doi.org/10.1016/j.neuroimage.2010.07.034. 2011.
[156] Hotelling H. Relations between Two Sets of Variates. Biometrika 1936;28:321–77. https://doi.org/10.1093/biomet/28.3-4.321. 1936.