NEW TECHNIQUES TO FACILITATE LONGITUDINAL VIDEO-BASED
DIGITAL TRACKING OF INFANT DEVELOPMENT

By

RICA RAI

A thesis submitted to the

School of Graduate Studies
Rutgers, The State University of New Jersey

In partial fulfillment of the requirements
For the degree of
Master of Science

Graduate Program in Psychology

Written under the direction of
Elizabeth Torres

And approved by

_________________________________________

_________________________________________

_________________________________________

New Brunswick, New Jersey

January 2022
ABSTRACT OF THE THESIS

New Techniques to Facilitate Longitudinal Video-Based Digital Tracking of

Infant Development

By RICHA RAI

Thesis Director:

Elizabeth Torres

In the last few decades, the rate of neurodevelopmental disorders has increased. However, current criteria to detect developmental differences rely on social and emotional parameters determined by observation alone. This reliance on observation restricts the description of movements to those that are less ambiguous, thus leaving out important components of the continuous stream of motions that make up natural behaviors. Movements that occur spontaneously and lend fluidity to social behavior occur largely beneath awareness and escape the naked eye of the observer. Furthermore, they are present since birth, unlike the more reliable social movements that appear later in life, when an infant’s motor system has matured. We here take the approach of examining movements since birth and tracking their maturation, to determine natural ranges of typical neurodevelopment. We leverage off-the-shelf technology and new advances in computer vision to bring the parents into the research loop, empowering them as active contributors to cognitive developmental research. Two families submitted
weekly videos of their infants starting at the age of 10 days, and continuing for 17 weeks, both infants comprise a longitudinal data set tracking the infants over the first 17 weeks. In addition, videos of infants were acquired from open access sources, which when combined with the longitudinal set, comprise a cross-sectional dataset. The videos were processed through OpenPose, and using these videos, we develop proof of concept that simple video-based assessment is possible from the comfort of the home. This new model can provide both longitudinal and cross-sectional data, to shed light on the evolution, since birth, of basic building blocks that will make up social motor behavior, possibly including atypical patterns detectable years later. Our work introduces new methods of data acquisition and analysis that can advance the study of neuromotor control in neonates and provide the means for much earlier detection of neurodevelopmental derail.
Acknowledgements

I would like to thank my mentor Dr. Elizabeth Torres, my committee members, Dr. Thomas Papathomas and Julien Musolino, and all my fellow lab mates in the Sensory Motor Integration Lab for their continued support.
# Table of Contents

Abstract .............................................................. ii  

Acknowledgements ................................................ iv  

List of Illustrations .............................................. vi  

1. Introduction ...................................................... 1  

2. Methods .......................................................... 10  

3. Results .......................................................... 33  

4. Discussion ....................................................... 49  

References .......................................................... 65
List of Illustrations

Figure 1. Neurodevelopmental Derail Tracked by Wearable Sensors.............. 2
Figure 2. General Movements (GM) Theory and New Joint Decomposition..... 4
Figure 3. Pipeline for data acquisition.................................................. 14
Figure 4. Pipeline to Transform Data from Continuous Analog Data to Digital
MMS Derived from Pixel Trajectories.................................................. 19
Figure 5. Stochastic Analyses Interpretability and Inferential Statistics......... 21
Figure 6. Gamma Moments parameter space. Pipeline to represent change..... 27
Figure 7. Pipeline to obtain the critical points profile.................................. 29
Figure 8. Cross-sectional Analyses.......................................................... 31
Figure 9. Video Data Validation............................................................... 34
Figure 10. Sum of Movement Excursions (mm) per Day.............................. 35
Figure 11. Cumulative Sum of Movement Excursions (mm) Over Time......... 36
Figure 12. Longitudinal Bar Plots of Quadrant Occurrences – Baby B1........ 38
Figure 13: Longitudinal Bar Plots of Quadrant Occurrences – Baby B2......... 39
Figure 14. Superimposed Critical Point Profiles of B1 and B2...................... 41
Figure 15. Quadrant ratio to track maturation in Babies 1 and 2................. 42
Figure 16. Average Signal to Noise Ratio................................................ 43
Figure 17. Tracking B1 and B2 Longitudinally Through the Empirically Estimated
Gamma Moments Space........................................................................ 45
Figure 18. Tracking Cross-Sectional Data – Identifying Self-Emerging Patterns of Similarity…………………………………………………………………………………….. 47

Figure 19. Maximum and Minimum Cross-Sectional Critical Points……………….. 48

Figure 20. Minimum and Maximum Critical Points Summary…………………….. 48

Figure 21. Order of Maturation…………………………………………………………….. 54

Figure 22. Informative Poses for the Evaluation of Motor Skills Acquisition…….. 61
1. Introduction

Physicians and infant specialists frequently use the neonate’s earliest motor patterns to gauge and predict the intactness of the neurological systems. Among these clinical tests, many are currently in use, e.g., Prechtl’s General Movement Assessment, the Abnormal Involuntary Movement Scale (AIMS), to name a few [1-3]. These tests rely on subjective observation by a well-trained clinician who uses standardized criteria, validated by others in his/her field. They have successfully identified abnormalities in neurological development and growth during infancy [4-8], however, they are limited in their ability to see past what is visible to the naked eye. The use of technology to track biorhythmic bodily motions from birth offers the potential to bring a higher level of precision to complement current clinical criteria. It also offers new avenues to theorize about motor organization and emerging neuromotor control, which lend to the development of cognitive behavior, as we conceive the motor output as reafferent sensory input and provide a sense of the quality of this type of signal that the brain receives as a continuously self-generated sensory stream [9].

Recent theories of motor organization and control adopting a nonlinear dynamical systems approach [10] focus on how, from the start of life, endogenously self-generated spontaneous movements emerge as a key ingredient, scaffolding cognitive control [11]. These motions, which occur largely beneath
awareness, can be harnessed with wearables (Figure 1) and longitudinally tracked as neonates develop. They support the transition from writhing to fidgeting to intentional motions (Figure 2). Intentional motions emerge aimed at sensory goals, supported by spontaneous motions, which help the nascent nervous system of the neonate self-discover the purpose and consequences of self-generated motions.

Figure 1. Neurodevelopmental Derail Tracked by Wearable Sensors. Motivation for this work is from previous findings, concluding that we can detect neurodevelopmental derailment by 3 months of age. (A) Wearable sensors affixed to leg warmers collected data for 9-12 hours once a month, over the course of 5 months in 60 neonates [12]. By three months of age (third visit) it was evident which infant stunted in growth and/or neuromotor control development captured by the wearable biosensors. (B) Derivative data from the rate of change, based on body weight, was also collected for preterm at risk (AR) infants and from full-term, typically developing infants (TD). The inset plot is from the absolute values of weights using growth chart data. Rate data clearly shows differences between infant cohorts. The plots in the bottom panel show (left) different groups by rate of change in weight, body length and head circumference (proportional to marker size). It is a map of growth and traditional AIMS data from Pediatrics, and revealed 3 clusters of infants (red at risk and blue controls) with representative insets of each class showing empirically estimated statistics across the first three visits (head circumference proportional to...
marker size). Rank 1 are TD infants, Rank 4 are AR infants, and Ranks 2 and 3 are mixed (red AR vs. blue TD). The right plot shows the relationship between the nonlinear rate of change of body weight vs. the rate of change of NSR.

Among relatively recent research developments complementing those from the clinical fields, Esther Thelen’s theory concerning the dynamic motor perspective [13, 14] proposed a bridge between motor and cognitive development, and brought new methods to further develop her innovative research program. While studying the infant’s kicking patterns, she proposed that changes in biodynamic properties of bodily segments, and motor synergies, were key to establishing neurological maturation as a basic foundational component of communication [15]. While these ideas certainly revolutionized the field of Developmental Psychology, new quantitative methods are lacking to scale objective research beyond subjective observation.

Part of the challenge is that movements evolve quite rapidly in asynchronous patterns across infants, with overlapping epochs. Transitioning reflexes while self-organizing and coordinating the body and limbs in motion, evolves rapidly across the first 30 weeks of life (Figure 2.) The general Theory of Movements that clinicians use situates the transitions from writhing to fidgeting at about the 6-9th week of life by observing the evolution between proximal and distal joints. However, during this transitional period, we know very little about
the articulated joints of the limbs. These are key to solving the degrees of freedom problem that Nikolai Bernstein posed in the 1930s.

Figure 2 General Movements (GM) Theory and New Joint Decomposition. (A) Proposed schema by Prechtl and others of general movements maturation. Writhing movements are small, with a moderate amplitude, and a slow to moderate speed often ellipsoidal in shape. They appear starting at birth. Fidgeting movements are of variable acceleration and present a moderate speed of the neck and trunk, with limbs moving continuously in all directions. They appear at around 6-9 weeks and coexist with writhing movements. They last till 15-20 weeks, when antigravity motions with intent begin. (B) Proximal and distal joints are commonly tracked at the clinic to ascertain the intactness of the nervous systems using these GMs maturation. (B) We add articulated joints to proximal and distal, to address the Bernstein DOF problem connecting motor and cognitive learning. Tracked joints here are 8 of the 25 estimated by OpenPose. Based on confidence threshold, we track Proximal (head, neck, both shoulders), Articulated, (both elbows) and Distal (both wrists). Color schema will be used throughout the thesis to represent these joints.

Indeed, the complexity and nonlinearities of movements are well appreciated in the early work of Nikolai Bernstein. Upon building his own motion caption system in the early 1930s, Bernstein discovered the variability of the limbs’ kinematic trajectories, the synergies that self-emerge during complex movements,
and more generally the hard problem of determining, given an external goal, the appropriate set of internal synergies and postural configurations that would attain such an external goal along an energy efficient path. He noted that such synergies help simplify the DOF problem by reducing the complexity of trajectory formation and joints’ coordination [16, 17]. Yet, how to build motor programs that automate these synergies and recruit/release them on demand as a function of the task demands and environmental context remains a computational challenge across more than one discipline.

One approach to the DOF problem is to understand the motor learning process from very early life. To that end, Dr. Torres’s Sensory Motor Integration Lab has created a phylogenetically orderly taxonomy that offers a multi-layered approach to model and empirically assess neuromotor function. This taxonomy spans from endogenously self-generated motions (from central patterns generators at various levels of the spinal cord) to exogenously driven motions (with intent) derived from exogenous sensory goals in the environment. Motor variability is seen in this model as a form of sensory feedback, that contributes to bodily self-awareness, and the formation of a coherent (motor) percept when development is on the right path. We use these principles here and aim to understand the emergence of this percept, and its role on the acquisition by the
infant’s brain of mental intent in connection to physical bodily autonomy and volitional control.

Thelen and Bernstein both define the development of neuromotor control as an emergent phenomenon that integrates many elements, spanning from higher level cognitive goals to postural configurations and motor synergies gracefully attaining such goals [13, 18]. Keeping these inter-relations in mind, we study the organization of bodily motions through three different classes: Proximal joints, Articulated joints, and Distal joints, discussed in Section 2.2.1 and shown in Figure 2, as they evolve in the first 17 weeks of life. Our main goal is to provide new data acquisition and analytical methods that permit scaling successful neurodevelopmental research broadly across the population, and expanding it to the comfort of the home. We further seek to empower the new parents by including them in the research loop, thus creating a bridge between this community and the research in neurodevelopment. As Thelen, who built a van to travel across rural Indiana in the 1980s and study infants in the comfort of the home, here we provide new means to do this with off-the-shelf, commercially available technology. In the times of COVID, this is our virtual van to advance early neurodevelopmental research.
1.1. Clinical Significance

The tracking of neuromotor development during formative years remains a challenge, given its accelerated rate of change. Wellness visits are not as consistent as daily interactions between the infant and the parents. Waiting until 2-3 years of age for a diagnosis wastes precious time to leverage the plasticity of the nascent nervous system, as the rates of physical growth are intrinsically related to the rates of neuromotor control [11]. In this sense, combining digital means to register natural motions at home, with verified clinical milestones in motor development from traditional observational inventories, increases the likelihood of detecting neurodevelopmental derail earlier. This type of research involving parents in the loop, can better inform early intervention programs of important milestones in the development of autonomous volitional control.

For example, the average age of diagnosis for ASD is around 4-5 years of age, when challenges in social communication and repetitive movements become unambiguously evident [19]. Social interactions critically depend on the scaffolding of proper motor sensation, which is disrupted in Pediatric Movement Disorders (PMD) from an early age [11]. When the motor and perceptual systems have reached adequate levels of maturation, social cues and behavioral differences are typically easily detectable by humans. However, in cases where maturation has stunted or deviated from normalcy, the expected social norms will not be
present. Such an outcome can be prevented or helped if an abnormal motor signature is detected earlier in life, prior to the development of the types of atypical social interaction patterns that take place as the infant transitions from 3 to 4 years of age [20].

In over two decades, the average age of diagnosis for ASD has not improved significantly [21] and it is unlikely that it will, given that it relies on core social/communication criteria, unambiguously detectable after three years of age. Prior to this pre-cognitive, pre-language period, clinical inventories assessing social interactions used to flag early signs of autistic traits, are too coarse. They cannot detect critical and subtle changes in key milestones (like reflexes gone astray in autism [22]) or the spontaneous acquisition of bodily autonomy and volitional control [23].

In low-income families it may take an additional two years to diagnose autism, even after inventories flag early signs. These communities are not guaranteed appropriate access to medical care [24]. Such lack of care prevents early intervention informed by proper diagnostic criteria within appropriate time windows. It is ideal to detect the neuromotor problem earlier, when the somatic sensory motor systems are still developing, and the basic building blocks of social interactions are forming [25]. In the absence of such foundational blocks, coping nervous systems may develop differently, e.g., leading to frequent falls or head
injury. Therefore, early detection of PMDs for early intervention is critical to aid in the development of systems that are essential, to support socio-motor behaviors during infancy [26, 27]. Traditional diagnostic criteria focusing on social cognition during pre-cognitive times will inevitably propagate delayed diagnosis until subclinical conditions develop into clinically apparent symptoms. Video-based means paired with state-of-the-art algorithms for pose estimation, which can detect motion trajectories, allow us to reach underserved communities in rural areas or in lower socio-economic neighborhoods.

Here we propose to develop new techniques for longitudinal assessments of video-based tracking (e.g., a few minutes per week) and cross-sectional data from open access repositories, with the purpose of eventually combining this information with traditional methods. Among these are those to measure the rates of physical growth (e.g., using the growth charts from the World Health Organization (WHO) and the Centers for Disease Control (CDC)). This combination may help us develop a personalized assessment tailored to the true rates of development of each infant, subject to the environmental constraints where the family lives.
2. Methods

2.1 Participants

There are 9 infants in the present study. Their data was obtained from two main sources explained below: (i) videos selected from YouTube; (ii) home videos collected from the parents of two infants, starting as neonates at 10 days of age to 17 weeks; this data collection is ongoing. The videos collected from part (i) are used for cross-sectional analyses. The infants studied in part (ii) were tracked longitudinally over the course of 17 weeks, a video from each infant was submitted for 12-13 weeks of the total of 17. All video data was obtained in the supinated position, with the infant on their back, and the camera posed above the infant. In future extensions of the present pilot study, additional positions will be requested to track the development of antigravity muscles and bodily strength and dexterity more effectively, as other positions, such as the prone position with the infant lying on their abdomen, provide information to the strength of the infant’s body as they grow. Figure 3 displays the pipeline of data acquisition.

2.1.1. Data Acquired from YouTube Videos

Videos were selected from YouTube based on the following criteria: (i) the infant was in the supine position, with the full body in view, (ii) the age of the infant at the time of the video was stated in the description or title of the video. Videos were found by searching key words such as “My Baby Moving”, “My Baby
Kicking”. A total of 74 videos were initially found, however after filtering to ensure that the age and full body were present, seven were selected for this work. The selected videos are included in the cross-sectional dataset. The disadvantage of selecting videos from YouTube is that personal information of the infant is not readily available. Gender and underlying genetic or medical conditions cannot be easily collected. In addition, camera angles and movement can change rapidly, as most of the acquired videos are home videos. This motivated the approach of recruiting parents with newborn infants to submit weekly videos of their child, to ensure accurate information could be collected, and that the infant could be tracked longitudinally through video. In this way, the tracking of development can be scaled-up and done in the home. Figure 1A displays previous work where neurodevelopmental derail could be detected by 3 months of age, based on the data acquired from wearable sensors affixed to the ankles of neonates over the course of the first three months of life [12]. Extending the acquisition to video-based motion caption will expand the research more easily but will also require validating video and wearables’ data.

2.1.2. Data Acquired from a Home Setting

Two families with neonates were recruited to collect longitudinal data of their infants over time. The parents of each infant submitted a 1–2-minute video of the infant every week. Parents were instructed to take the videos with the infant
in the supine position, with the full body in view. The videos were recorded with the personal phones of the parents, each with a video sampling rate of 29 frames per second (fps).

Through the submission of these videos, it was found that pants or loose clothing worn on the lower body obstructs accurate pose estimation from the infant, specifically in early weeks. The problem is resolved in later weeks as parents are more comfortable with their infant not being fully clothed, since by that point the infants’ systems are able to better regulate body temperature. Participants recruited in the future will be requested to dress their infant appropriately for the video collection. The parents in the current study were provided with guidelines, (i) the full body of the infant should be in the video and (ii) to hold the camera directly above the infant in one spot. In future work, we will also add specific instructions in regards to appropriate clothing.

2.2. Data Generation from Pose Estimation

Recent advances in computer vision offer new algorithms that permit the use of video data to accurately estimate full bodily poses in detail [28]. These include the face, hands, body, and feet (Figure 3D-E). The primary advantage of these methods is that they provide time series positional data, which can be used to estimate kinematic trajectories of the various parts of the body.
Here we use OpenPose, a pose estimation software developed by the Perceptual Computing Lab at Carnegie Mellon University [29], to estimate bodily poses of 25 positional joints extracted from the supinated pose of the infants. OpenPose outputs confidence matrices which reflect the quality of the estimation procedure (Figure 4B). All frames below a .75 confidence were rejected, to ensure good continuous data across all joints tracked. Figure 9 displays the joints tracked in this study, all of which are on the upper body. Further, we convert pixels per frame to spacial and temporal units (mm/s). This is done solely for the longitudinal dataset, as we do not know the recording devices used for the YouTube videos. Based on the pixel density (pixels per inch/PPI) of the phone used to record the video, the computation of distance is shown below:

\[ \text{mm} = \text{pixels} \times \frac{25.4}{\text{PPI}} \]  

(1 in = 25.4 mm)

The smart phones used in this work both had pixel densities of 458 PPI.
2.2.1. Motivation for Joint Decomposition

The theory of General Movements advanced by Prechtl and others [30] had previously focused on proximal vs. distal joints, to track the evolution of bodily synergies, and the transitions across different stages of motor control. These are elaborated in Figure 2, following their large body of work primarily used in clinical
settings [31]. Our work extends the proximal and distal parts of the body, to include the articulated joints. As shown in Figure 2B, the Proximal Class (blue) is defined as the joints and body parts closest to the head and core. Including the neck, shoulders, eyes, ears, and hips. The Articulated Class includes the larger hinge joints, the elbows, and knees. The Distal Class includes the minor hinge joints, the ankles, and wrists, as well as the joints within the hands and feet.

Based on the Bernstein Degrees of Freedom problem, we hypothesize here that there is an order of maturation from proximal, to articulated, to distal positional joints. The articulated joints are introduced since they reflect the transition from a flexed position, as fetuses have in the womb and continue to present during the first three months of life, to extended poses. Extended poses heavily recruit articulated joints, as the infant develops motor control of its distal joints and explores the peripersonal space. As antigravity muscles mature in the shoulders, the infant will expand the arms and legs, which lead to well-known motor milestones such as rolling over and crawling. This study tracked only the supine position, and due to limitations discussed in section 2.1.2, only the upper body is evaluated in this work, specifically the head, neck, shoulders, elbows, and wrists.
2.3. Data Generation Through Digitization of Continuous Analogue Streams into Spike Trains

In recent years, Dr. Torres’s lab has converted analog signals of biorhythmic data from continuous streams of motion to digital signals. This is made possible by obtaining a new data type created by the lab, coined “micro-movement spikes” (MMS). These are spikes reflecting the fluctuations away from the empirically estimated mean, obtained from the entire sequence of points in the time series. In traditional approaches of continuous streams e.g., motion data (analog streams), the approach is to assume a theoretical mean (e.g., from the Gaussian distribution), and to take epochs of data from repetitions of the movement. These epochs are then averaged across those repetitions to obtain the standard deviation from the theoretically assumed mean. During this procedure, fluctuations away from the standard error are deemed noise, incurring inevitable data loss. This data (gross data) is precisely the information needed to examine the body at an individual level, as it informs us of the variations that are unique to an individual.

We follow a new procedure that obtains the peaks of the kinematic parameters (e.g., the speed parameter derived from the time series of the positional data), and takes overlapping windows of data, continuously sweeping through the time series. The peaks conforming the entire time-period of the data (in this case 1.5 minutes, sampled at 29Hz) are then gathered into a frequency histogram,
and using maximum likelihood estimation (MLE), we estimate the best continuous family of probability distributions that best fits the data. We then use the empirically estimated moments of the distribution to subtract at each frame the deviations away from the empirically estimated mean.

These fluctuations are scaled into standardized, unitless values (ranging from 0 to 1) to scale out possible allometric effects in subsequent analyses. Allometric effects are due to anatomical differences across infants, due to asynchronous cross-sectional growth patterns in the population, and even within the longitudinal profile of an infant. Scaling out such anatomical factors enables us to plot all infants on the same parameter space, and to appropriately track their differences.

We discovered the need to scale out anatomical differences in neonates and young infants, due to non-linear patterns of growth present in derivative data. Such data can be obtained from the growth chart, involving longitudinal (from birth to 2 ½ years old) and cross-sectional (from 2 ½ to 4 years old) data from head circumference, body length and weight [32].

Figure 4 displays the pipeline used in this study to obtain the kinematic parameters of interest. The output from the OpenPose estimation (Figure 4A) offers the confidence matrix for a sample video (Figure 4B). The top panel shows the matrix for one frame. Here we see 25 columns by 3500 rows. The columns are
the joints detected by OpenPose. The rows are the frames at 29 Hz. The bottom panel of Figure 4B displays the reduced version upon filtering the mode above 0.75 confidence, to ensure high quality of the trajectories. This means that most frames are above this threshold across the entire set of infants. The color bar reflects the confidence scale ranging from 0-1. Figure 4C depicts the positional trajectories plotted as X-Y coordinates (top plot) as frames of X and frames of Y below. This type of plot lets us know e.g., which dimension is greatest in value. In this case, we can see that the head moves mostly along the Y dimension. The data is filtered to eliminate noise, due to the estimation process, and to preserve fluctuations mostly due to actual motion. This is ensured by spline interpolation methods present in MATLAB. Figure 4D displays the output of the speed computation, upon taking the derivative of the position (velocity plot not shown). From this we obtain a time series of scalar values in pixels/FPS or Hz, (later converted to mm/s to obtain distance traveled by each joint). To obtain the speed, we use the Euclidean distance.

We later report the distance traveled as the physical excursions spanned by each body part under consideration. This is done by cumulatively summing the speed per timepoint across the time series. These excursions indicate how much physical motion the infant generates, which in turn can provide estimates of how
much energy the infant is consuming as it grows and develops neuromotor control.

**Figure 4. Pipeline to Transform Data from Continuous Analog Data to Digital MMS Derived from Pixel Trajectories.** (A) OpenPose provides a kinematic skeleton, detecting all the major joints and body parts in motion throughout the input video, recording at 29FPS. (B) The software provides a confidence score for each keypoint detected. A threshold of 0.75 confidence is used for this work. After thresholding, eight body parts were selected for the analysis of this study. These parts were decided based on the availability of useable data after thresholding and based on the Proximal, Distal, and Articulated Classes. For the purposes of this figure, the data taken from the detected head keypoint, during one video taken 10 days after birth, is shown. (C) The Head Positional Trajectory displays the XY coordinates of the pixels detected as the head, throughout the duration of the video. (D) The speed of the pixel-based positional coordinates is then computed, based on the change in their position over time. (E) The speed profile is then windowed into 5-second intervals with a 50% overlap. The windowing is performed to account for the difference in video lengths across the dataset. Each window is then individually treated as a data occurrence for that video. (F) For the speed profile of each window, the fluctuations in the amplitudes of each peak from its corresponding speed profile are obtained and normalized. These are the micro-movements (MMS) of the speed profile. Using the Statistical Platform for the Individualized Analyses of Behavior (SPIBA), the MMS waveform is analyzed. (G) The peak amplitude of the MMS from this window are collected into a frequency histogram, from which Maximum Likelihood Estimation is used with a 95% confidence interval to identify the distribution of best fit. (H) The Gamma distribution is identified as the best fit. (I) The fitted probability distribution provides the shape and scale Gamma parameters. Steps E – G are
Importantly, we sweep through the time series with overlapping time windows of 1 second, providing enough peaks to build a frequency histogram. We then use MLE to empirically estimate the parameters of the continuous family of probability distribution functions (PDFs), characterizing the infant’s behavioral states. Figure 4D depicts the windowing of the speed peaks (marked as red dots), while Figure 4E depicts the window size for one occurrence. Note that the size of the window is obtained after the confidence intervals are examined, to settle on an appropriate estimation of at least 100 micro-movement points, for later use in the PDF trajectory estimation for our stochastic analyses. Figure 4F shows a window of MMS, which convert analog video data to digital spikes. Prior work has revealed important results in autism, with diagnostic power across cognitive levels (from non-speaking to speaking some phrases, to fully verbal, highly correlated with clinical measures of Intelligent Quotient, IQ). For this reason, we chose to use the MMS in this work [33].

To normalize the fluctuations away from the empirically estimated mean, and obtain the MMS, we use equation 2 below [33, 34].

\[
\text{Normalized Peak} = \frac{\text{Local Peak}}{\text{Local Peak} + \text{Local Average}_{min-to-min}}
\]  

(2)
2.4. Stochastic Analyses and Parameter Spaces

The MMS data represents a point process which we can study using stochastic analyses to derive stochastic trajectories, as well as their non-stationary behavior, as the infant moves during the video recording time. Using methods developed in Dr. Torres’s lab, we design several parameter spaces with interpretable outcome, to perform inferences about the data, leading to hypotheses testing. To that end, we start out with exploratory analyses, and let the data lead us to surprising results later used to elaborate testable hypotheses.

Figure 5. Stochastic Analyses Interpretability and Inferential Statistics. (A) Above is a depiction of all the PDF curves for every window in one video. Each of these curves (windows) is treated as a separate marker of time throughout the video. (B) Based on the parameter curves, the shape and scale of each window is represented as a marker. The lines through these markers represent a confidence interval of 95%. (C) The shape parameter identifies the shape of the distribution, the lower values of which illustrates a more exponential distribution, the higher values illustrate a more gaussian distribution. The scale parameter represents the noise-to-signal ratio (NSR). The lower scale parameter corresponds to a lower NSR, and as it increases indicates a higher NSR. For each video, the median shape and scale were calculated, seen in C as solid lines across the parameter plane. These median values separate the plane into four quadrants.
The Left Upper Quadrant (LUQ) represents a more random and exponential space with high noise. The Right Lower Quadrant (RLQ) represents a more symmetric parameter space with a low noise to signal ratio. (D) The parameter space was defined empirically per body part for each day. (E) Depicts the critical point, the maximum EMD value described in Figure 7.

2.4.1. Gamma Parameter Space

The normalized peaks from the MMS are used to plot a frequency histogram (e.g., Figure 4G). We then fit a PDF using MLE (e.g., Figure 4H). This is done through the estimation of the Gamma (a) shape and (b) scale parameters of the continuous Gamma family of probability distributions. The Gamma family comes as a result from MLE, whereby it has been found to be the optimal means of representing MMS derived from human biorhythmic data [32, 35]. This has been the case in voluntary motions, in spontaneous motions, in involuntary motions, and in autonomic motions [35].

The plane spanned by the shape and the scale of each Gamma PDF derived from the MMS in each window, are then plotted with 95% confidence intervals as points along a trajectory, on the Gamma parameter plane. For example, Figure 4I displays 8 points of such a trajectory for one video that had 8 overlapping windows. Importantly, the reader can appreciate different values ranging from less than a 100 to over 150 along the shape axis. Likewise, the reader can appreciate different values ranging from 0.007 to 0.0035 along the scale axis. These values tend to fluctuate in no order, from low to high, yet the trend when neurodevelopment is on the right path, is tending towards lower values of the
scale and higher values of the shape. These values are explained in Figure 5 as they have been mapped out for humans across the lifespan, using a multiplicity of biorhythmic data (EEG, EMG, ECG, kinematics, etc.). They have been collected with various biosensors, such as inertial measurement units (IMUs), electromagnetic sensors, motion capture cameras, and obtained through a variety of tasks (pointing, resting, walking, dancing, performing clinical inventories, etc.) [34, 36].

The general formula for the PDF of the gamma distribution is shown below, where $a$ is the shape parameter and $b$ is the scale parameter.

$$f(x) = \frac{1}{\Gamma(a)b^a}x^{a-1}e^{-\frac{x}{b}} \quad (3)$$

The moments (mean and variance) are $a^*b$ and $a^*b^2$ respectively. For this reason, the scale is equivalent to the noise to signal ratio (NSR), which we track as part of the evolution of the stochastic signatures. The significance of this parameter is provided by the fact that movements and their consequences are sensed through afferent nerves, reflecting spatio-temporal information for the prediction of sensory consequences from actions [37]. This proposition extended from von Holst and Mittelstaedt, the principle of reafference in voluntary actions to include endogenously generated spontaneous motions. Such motions are present in neonates, and we track them here in the form of writhing and fidgeting movements, as they tend to lower the NSR and become more systematic. These
motions then transition into intentional movements. Intentional movements can then become goal-directed, as exploratory motions undergo self-discovery of goals evoked by surprising events. Such surprising events lead to curiosity and further exploration of the infant’s surrounding environment. Since we can continuously track the quality of the motor signal that feeds back to the brain from the periphery, we can infer when these transitions occur [12]. The present video-based methods enable this investigation at scale.

Importantly, Dr. Torres’s lab has discovered that the log-log parameter of the Gamma estimation yields a tight linear fit, and that when the NSR reduces, the shape increases toward a more Gaussian (symmetric) shape (increasing values of the shape parameter, a) [12]. This tight linear relation reduces the parameter to track to one dimension (e.g., the NSR, b), which we can then follow along the time of neurodevelopment, as an indicator of the quality of the feedback that the brain is receiving from the periphery [11]. We will show later the results of tracking 1/NSR, which gives an estimate of the signal quality (low referring to noisy feedback, high referring to a clearer signal).

The stochastic trajectory on the Gamma plane has two components, the direction of the change, as one PDF transitions to the next PDF, and the magnitude of this change reflecting the speed of the trajectory. The latter can be obtained using the Earth Mover’s Distance [38]. This is a metric derived from transport
problems, measuring the amount of effort (work) that it would take to convert one frequency histogram (‘a pile of dirt’) to another frequency histogram (‘another pile of dirt’), when transitioning from one MMS window (behavioral state) to another, sweeping across the continuous video data. Figure 5 depicts the pipeline to go from empirically estimated Gamma PDFs (Figure 5A) to a stochastic trajectory on the Gamma parameter plane (Figure 5B), to interpretable trajectories (Figure 5C).

Importantly, we obtain the median value of the trajectory (median shape and median scale) and divide the Gamma plane into, quadrants whereby the Left Upper Quadrant (LUQ) contains the instances of the stochastic trajectory that have high NSR and low shape value. The latter are closer to the exponential distribution, which is a limiting case of the Gamma family, whereby the shape is $a=1$, and reflects random, memoryless states of the renewal process under consideration. These states are ubiquitous in autism and neurodevelopmental disorders of the like [33, 37]. The opposing quadrant in the log-log Gamma plane is the Right Lower Quadrant (RLQ), reflecting Gamma PDFs towards the Gaussian like range (symmetric distributions with values >$100$ of the shape $a$). The range between $a=1$ and $a=100$ are skewed distributions with heavy tails to the left or right. These are present in learning processes and neurotypical developmental states. Transitions between these quadrants reflect healthy learning and development. When only
one regime prevails, we see pathologies of the system owing to excess randomness and noise, and/or a very narrow bandwidth of values.

2.4.2. Critical Points of the Gamma Trajectory

The direction and magnitude of the shifts in the points of the Gamma trajectories are quantifiable as stochastic kinematics features. We track the jumps from the LUQ to the RLQ as the system evolves through time, on the scale of seconds (our time window resolution) to minutes and days to weeks. To that end, we use the trajectories of the video data collected in two modes previously explained in section 2.1. We track two infants longitudinally over the span of 16 weeks (data is still being collected). We also track infants cross-sectionally using a combination of the YouTube videos, with infants in the supinated position. Figure 5D displays the trajectories of the Gamma Parameter plane for day 10 (1.5 weeks) and Week 16, showing the evolution of the median lines and scatter over time. Figure 5E shows the trace of critical points (transitions, explained in section 2.4.4.) as they fluctuate along the stochastic trajectory. These are equivalent to the magnitude of the shift (obtained with the EMD), and the colors and shape reflect whether the point lies in the LUQ or RLQ.
Figure 6. Gamma Moments Parameter Space. Pipeline to Represent Change. (A) The gamma estimated PDFs for each window through one video can be summarized in a 3-dimensional Gamma parameter space (B). The empirically estimated summary statements, the mean (µ), variance (σ), and skewness of each window is shown. The size of each marker represents the kurtosis of the window, a larger size marker indicating a higher kurtosis. The shape of the marker indicates whether the window is in the LUQ or RLQ. (C) The parameters for each day are calculated and allow us to track the general change in this space over time, here showing a shift from Day 10 to Week 16. (D) Shows all the scatter of all windows. (E) The face color of each marker represents the MMS speed of the given window. The edge color depicts the point in time at which the video was taken, a lighter blue depicting closer to birth. The LUQ points (blue circle) vs. the RLQ points (red triangles) track the levels of NSR and predictability of the empirically estimated Gamma distributions.

2.4.3. Gamma Moments Space

The stochastic trajectories that we compute on the Gamma plane using the shape and scale parameters to represent each window of the video, can also be visualized on the Gamma moments space. This parameter space is spanned by the mean (µ) along the x-axis, the variance (σ²) along the y-axis, the skewness along the z-axis, and the kurtosis, which we make proportional to the size of the marker
(higher kurtosis equates to larger marker size). We add to this visualization the color of the marker edge to represent the time along a gradient of blue, whereby lighter color represents earlier weeks. Furthermore, we represent the MMS range by a gradient color that we use for the marker face. These can be appreciated in Figure 6, where we show the pipeline to get this representation. Figure 6A is the family of Gamma PDFs empirically estimated drawn using the shape and scale parameters and colored according to the week. Furthermore, because of the normalization, larger MMS values represent lower speeds (the ratio in equation 2 has a term in the denominator which penalizes the numerator). Larger speeds on average, result in lower scaled MMS values. Lower speed values result in higher scaled MMS. Figure 6B represents the trajectory’s critical points, with triangles representing points which lie in the RLQ of the shape and scale parameter space, and circles in the LUQ. Panel 6C contrasts this parameter space between day 10 and week 16, taken from the head. We appreciate a shift of the scatter separating LUQ from RLQ points, with lower speeds in the RLQ signifying more controlled motions. Panel 6D mixes all values from Week 1 to 16 and shows more kurtotic distributions in the RLQ with overall lower speed on average.
Figure 7. Pipeline to Obtain the Critical Points Profile. (A) For each video, multiple windows were obtained for which a PDF was empirically estimated. (B) The Earth Mover’s Distance was calculated between each window. The matrix depicts the EMD value between the MMS of each window in the given video. A lower (darker blue) value represents windows of more similar distributions. A higher (yellow) EMD value depicts windows that are highly differentiable. (C) For each day a video was submitted, the EMD can be tracked longitudinally. The change in EMD over time throughout the video and through the longitudinal dataset represents the stochasticity of an infant’s movement over time. This can provide insight to the change in head speed over the course of video submissions. (D) The median EMD value for each day is plotted over time, referred to as the Critical Point. The marker of each critical point identifies whether it is obtained from a window located in the LUQ or RLQ parameter space shown in Figure 5. This can depict the level of change from week-to-week. Blue circles are from the LUQ, and red triangles are from the RLQ, displaying interpretable stochastic shifts whereby LUQ are high noise and random regimes vs. RLQ of high signal and high predictability.

2.4.4. The Critical Points

As explained above, we use the EMD to obtain the magnitude of the shift between quadrants, defined by the median shape and median scale values. As the scatter on the log-log Gamma parameter plane fluctuates between these quadrants
along the tightly fit straight line, the points trend toward the RLQ. They do so with a relation inversely proportional between the scale (NSR) and the shape. As the NSR decreases, the shape value increases. The PDFs become more symmetric, and the dispersion decreases. The distributions tend to be more kurtotic and Gaussian like. When all is well, the signal to noise ratio increases as the system matures [12].

To measure this evolution across windows, we first obtain the frequency histograms of the MMS (Figure 7A). We can then visualize the pairwise EMD values across windows (Figure 7B) and obtain the curve of the EMD value fluctuating across weeks for each video-based window. Each day of measurement is then summarized in Figure 7D, whereby the median EMD value is obtained for each day. Using the median shape and median scale of that day, the quadrant where this median value falls is represented by a blue circle (LUQ) or a red triangle (RLQ.) This profile of the critical point is obtained for each body part and examined in terms of Proximal, Articulated and Distal positional joints. This is done for each infant longitudinally assessed, and across all infants to build a cross-sectional set.
Figure 8. Cross-sectional Analyses. For the cross-sectional data set, (A) the shape and scale parameters were calculated for each video (as a whole, not windowed) within each cohort (as seen in Figure 5). (B) The EMD for each cohort was computed as seen in Figure 7. (C) The mean EMD point (critical point) for each cohort was plotted. The marker face indicates whether the video was in the LUQ (blue circle) or RLQ (red triangle). (D) The average critical point value, based on all the critical points in each cohort, were computed and plotted. This was done for all body parts; this figure depicts only the head.

2.4.5. Cross-sectional Analyses

To ascertain group behavior, we pool the data from all infants. The videos from the longitudinal infants were placed with the YouTube videos, depending on the age of the infant at the time of the video, to create 6 cohorts: 0-2 weeks, 3-5 weeks, 6-8 weeks, 9-11 weeks, 12-14 weeks, and 15-17 weeks. From this we build a trajectory of the EMD in a pairwise fashion, building a matrix that represents how similar the infants are to one another (as in Figure 7A-B). Figure 8A depicts the sequence of the Gamma parameter plane shifting stochastic trajectories, relative to the median shape (a) and scale (b). Figure 8B depicts the evolution of all
the infants in each age cohort. There were fewer infants included in the 0-2 weeks sample, however as time progresses, these head stochastic trajectories yield more similarity across infants in weeks 15-17. This flags baby 32 in this week as different from the cohort of babies 27-31. We use these patterns to track how individual body parts evolve as a group, i.e., which body parts converge to a similarity pattern, and what correspondence they have with mean critical points of the stochastic trajectory. This is depicted in Figure 8C-D for the head.
3. Results

The results divide the selected body parts into the three joint classes, the Proximal division (head, neck, left shoulder, right shoulder) shown in blue, Articulated (left elbow, right elbow) shown in green, and Distal (left wrist, right wrist) shown in orange.

3.1. Video Data from Smart Phones Yield High Confidence Data – Can Track Development of Neuromotor Control

The video data, filtered to include only that with confidence levels above 0.75, enabled the use of enough positional data across the infants to ascertain 8 body parts inclusive of the head, neck, shoulders, elbows, and wrists. This enabled us to build a model of the stochastic progression of the maturation process starting from 10 days after birth, until week 16 (data collection ongoing). Figure 9 shows sample speed profiles of one day’s worth of data from one infant.

Given that the videos are recorded with smart phones, during early days of infancy, it was suspected that the quality of the videos would not be suitable for accurately tracking changes over time. However, the videos submitted were deemed adequate for the data analyses done in this work. This is relevant, as it opens a more accessible and novel way to assess the development of neonates longitudinally, in a cost-effective way from the comfort of one’s home.
The American Academy of Pediatrics recommends a well-child visit the first week, 1st, 2nd, 4th, 6th, 9th, and 12th months, and every year thereafter. However, changes during the early stages of life change day-to-day and can vary significantly on a weekly basis. It is unrealistic to expect clinicians to see infants every week, and it is difficult for parents to bring infants into offices. During this rapidly changing and critical period of neurodevelopment, a method of monitoring which can be done efficiently in the comfort of the home is ideal. These results provide proof of concept, that we can easily deploy this study at a larger scale.
scale, since most households in the US have a recording device capable of rendering video.

Figure 10 depicts the conversion of pixel/frame movement to the total physical movement of the infant per day. The stochasticity of movements on a week-by-week basis in the Distal joints show a much greater variability of movement, in the wrists, whereby the Proximal joints maintain a more consistent degree of movement over time. Figure 11 depicts the cumulative sum of movements over time i.e., Week 1 displays the total mm/s of Week 1, Week 2 the total in Week 2 and Week 1, and Week 16 the total of all weeks prior. The rate of change in the Proximal class is stabilized early on, however the Distal class depicts a slightly more exponential increase.

Figure 10. Sum of Movement Excursions (mm) per Day. The total sum of movement excursions (mm/s) per day. The Proximal class displays less variability, and a consistent level of movement over time, specifically the head and neck. The shoulders in both infants stabilize by Week 6. The Distal class shows the greatest level of variability from week to week, displaying that the wrists are the most exploratory of the body parts studied.
3.2. Longitudinal Data Reveals Order of Maturation

The cohort of longitudinal data consisting of two newborn infants reveals individual patterns, but also a common trend to both infants. Below we describe these individual features along with the common findings.

3.2.1. Proximal → Articulated → Distal – Commonality of Quadrant Progression

Both infants (B1 and B2) in this work were tracked longitudinally over the course of 16-17 weeks, through video submissions provided by their parents. The submitted videos revealed an orderly pattern of convergence to the RLQ of
symmetric distributions (the quadrant with higher signal quality, and more controlled movement with reaferent feedback to the brain). This pattern is seen in both infants first in the proximal joints (head, neck, and shoulders), where the shifts in the RLQ from the LUQ are more frequent. This is quantified by tallying the frequency of window occurrences in each quadrant, represented by the bar plots in Figure 12 (B1) and Figure 13 (B2). The proximal pattern reflects head-neck-shoulder motions that begin as faster and less controlled, with more random distributions and a lower signal to noise ratio. Over the weeks, these motions evolve to slower and more controlled motions, with a stochastically more predictable distribution and greater signal content. This suggests a maturation of the proximal joints, supporting the head, and the building of a foundation for the strengthening of antigravity muscle control within the shoulders.

The evolution of the proximal joints’ maturation is followed in both infants by the articulated joints. These are still shifting heavily between the two quadrants, however by week 9, the right side of B1 presents more RLQ visits, signaling a higher signal and more predictive code than in earlier weeks. Lastly, the distal joints are still lagging in both infants (particularly B2), but we see that week 9 in both infants produces larger changes in one side of the body. According to Prechtl’s method (see Figure 2A), this week is generally seen as the transition from writhing movements to fidgeting movements. We anticipate that as we collect
more data over subsequent weeks, we will see the distal joints trending to slower, more controlled motions with a higher signal and more predictable distribution.

Figure 12. Longitudinal Bar Plots of Quadrant Occurrences – Baby B1. The pattern of maturation for Baby 1 over the first 16 weeks. The quantification of the shifts to the RLQ vs. those to the LUQ reveal earlier convergence to a lower NSR, and increased shape values (tending to the Gaussian distribution) for the proximal joints, followed by the articulated joints, and finally the distal joints which reveal more shifts in the LUQ. This signifies that the distal joints are still noisier and more random than the rest of the upper body.
3.2.2. Critical Points Reveal Differential Maturation in Each Infant

While the tallying of frequencies of shifts to the RLQ reveals a common pattern of maturation from proximal to articulated to distal, the tracking of the critical points is informative toward both the magnitude of the shift (EMD value) and the directionality of the shift (from the LUQ or RLQ). These are shown in Figure 14 for B1 and B2, superimposed for comparison, to appreciate the individual differences in stochastic trajectory. We see here how by weeks 13-16, B1 has consistent right wrist maturation with high signal to noise ratio and high predictability (higher Gaussian like shape). This is also appreciated in B2 for the proximal joints (neck and right shoulder). However, the distal joints are still
lagging in random and noisier locations of the Gamma parameter plane. Importantly, weeks 8-9 can be seen as critical for B2, suggesting that the important transition from writhing to fidgeting movements took place and maturation is ensuing. The critical points show that B2 seems to take longer than B1 to attain predictable movements with high signal to noise ratio. Yet, both infants seem to be on steady stochastic trajectories.

3.2.3. RLQ – LUQ Ratio - Useful to Assess Longitudinal Progression

The Quadrant ratio proposed here to help assess maturation provides a tally of the occurrences of shifts in a normalized way, which can now be used to compare both infants, despite the asynchronous patterns across proximal, articulated, and distal joint discussed above. The Quadrant Ratio is as follows:

\[
\text{Quadrant Ratio} = \frac{\text{# of RLQ Occurrences}}{\text{# of RLQ Occurrences} + \text{# of LUQ Occurrences}} \tag{4}
\]

This ratio spans values between 0-1, whereby as the number of occurrences in the LUQ decreases, the value of the ratio increases toward 1. We can then visualize how each block of joints progresses toward a more mature state (approaching 1), as this characterizes when motions are slower, more controlled, with a higher signal content and more predictable toward the Gaussian regimes of the Gamma family (away from the memoryless, random, exponential regime at \(a=1\)). Figure 15 depicts the quadrant ratio calculations of B1 and B2 over the course of the first 16-17 weeks. It is apparent that the proximal joints in both infants most
rapidly approaches 1, followed by the articulated joints. In the first 16-17 weeks, the distal joints do not approach 1, however future work with further longitudinal data will elaborate on this pattern.

Figure 14. Superimposed Critical Point Profiles of B1 and B2. Critical point profiles for baby 1 and baby 2. The blue circles indicate that the critical points lie in the LUQ of the shape and scale gamma parameter plane, whereas the red triangles indicate that the points lie in the RLQ. By weeks 13-16, B1 displays higher SNR and higher predictability, seen by the prevalence of RLQ occurrences in later weeks. B2 displays this pattern in later week in the proximal joints. Both distal joints in B2, and the left wrist specifically in B1, are most consistently in the LUQ. However, a change occurs for B1 in the left wrist during weeks 8-10, indicating a potential transition from writhing to fidgeting movements.
Figure 15. Quadrant Ratio to Track Maturation in Babies 1 and 2. The ratio is calculated by Equation 4. As the value approaches 1, it signifies a more controlled and mature state, with a higher signal content and increased predictability. The proximal joints fluctuate to and near 1 in both infants within the first 5 weeks. The distal joints follow the proximal class, however not as closely and frequently. The distal class does not reach a value of 1 and reveals a state with increased noise and decreased predictability.

3.2.4. Signal to Noise Ratio

The signal to noise ratio (SNR) is calculated by $1/b$ (scale). A greater SNR is ideal, as it signifies that more signal is present than noise. Figure 16 plots the average SNR over time for B1 and B2. Based on this figure, the RLQ has a consistently greater SNR than the LUQ in the proximal and articulated joints in both infants. However, there is greater fluctuation in the RLQ in the distal class.
Figure 16. **Average Signal to Noise Ratio.** The average Signal to Noise Ratio (SNR) in the LUQ and the RLQ for both infants over time. The SNR is calculated as 1/b (scale) from the gamma family probability density function. The RLQ displays a consistently higher average SNR than the LUQ. The variability in the RLQ is greater than that of the LUQ, verifying that there is higher signal quality in the more controlled motions which converge to the RLQ.

3.2.5. **Summary Gamma Statistics Reveal Shifts in Maturation**

The plots in Figure 17 reflect the evolution of the Gamma moments. They provide visual information regarding the shifts of the stochastic values, and the variability of patterns between the two infants. By weeks 16-17 both infants display that their proximal joints provide substantial foundational support to the rest of the upper body, with a slower, more controlled MMS range. B1 displays a prevalence in the RLQ in all joints, whereas B2 is still evolving from the LUQ for articulated and distal joints. The proximal and articulated joints in B1, and proximal joints in B2, display a slower more controlled speed in later weeks.
However, the distal joints in B1, and the articulated and distal joints in B2, show slightly faster movements in later weeks.

3.3. Cross-sectional Data Reveal Patterned Behaviors of the Cohort

The longitudinal data from the two infants was combined with the YouTube video data to form a set of six cohorts spanning from 0 to 17 weeks of age. Using the pairwise similarity metric (the EMD), we obtained colormap matrices to visualize patterns across the cross-sectional cohorts. These are shown in Figure 18, where we can clearly see blocks of blue gradients signaling similarity across those infants as time progresses. In the initial weeks there is high variation in the marginal distributions of the infants conforming each cohort. Yet, as time progresses, a grouping of oscillatory patterns across several infants for each of the joints is shown. The following patterns emerge: we see the head begin to group in weeks 15-17 for babies 27-31; the neck and left shoulder patterns group at weeks 12-14 for babies 20-25 and 21-24 respectively; right shoulder patterns for babies 28-32 at 9-11 weeks and again at 15-17 weeks. We also see similarity patterns for the articulated elbow joints for babies 7-10 during weeks 3-5 in the right elbow, and again for weeks 15-17 for babies 27-31 at the left elbow. The distal joints at both wrists show weeks 3-5 as common to babies 6-9 and as time progresses, we anticipate other similarity patterns emerging from the cohort.
Figure 17. Tracking B1 and B2 Longitudinally Through the Empirically Estimated Gamma Moments Space. Each marker here represents the window with the maximum EMD for each day. The right body joints from each of the three classes are shown here. The speed of the MMS (shown by the face color of the markers) indicates the average MMS speed, the darker red depicting slower and more controlled movements, and the yellow depicting faster less controlled movements. The shape identifies whether the marker resides in the LUQ (circle) or RLQ (triangle) of the shape-scale gamma parameter plane. The edge color indicates the time period of the windows. In later weeks, both infants display slower and more controlled movements in their proximal joints, with higher signal and predictability (RLQ). B1 displays a transition toward the RLQ in the articulated and distal joints, and B2 remains in the LUQ with faster movements throughout later weeks.

This can be interpreted as individual variations that when collated at the population level, converge to congruent patterns of neuromotor control development. These patterns surprisingly coincide with prior clinical criteria, however we non-trivially extend them to articulated joints, and provide more detailed trajectories of motor variability.
The maximum and minimum critical points (maximum and minimum EMD value) for each cohort are shown in Figure 19. An interesting pattern presented in the right side shows all right joints (shoulder, elbow, and wrist) have decreased variability after 6-8 weeks in the maximum critical point, whereas the left side better maintains this variability. Figure 20 panel A highlights that the left body parts display a greater range than their right counterparts.
Six cohorts were formed spanning 0-17 weeks of age: 0-2 weeks, 3-5 weeks, 6-8 weeks, 9-11 weeks, 12-14 weeks, 15-17 weeks. Above are the colormap matrices of the EMD values for each cohort, within each body part studied. The blue depicts a similarity between infants within each cohort. The blocks of blue gradient forming in later cohorts indicates a signaling in similarity across those infants as time progresses. The following patterns emerge: the head begins to group in weeks 15-17; the neck and left shoulder patterns group at weeks 12-14; right shoulder patterns emerge at 9-11 weeks and again at 15-17 weeks. We also see similarity patterns for the articulated elbow joints during weeks 3-5 in the right elbow, and again during weeks 15-17 at the left elbow. The distal joints at both wrists show weeks 3-5 as common to infants 6-9, indicating symmetry between the wrists in this metric.
Figure 20. Minimum and Maximum Critical Points Summary. (A) Shows the range for minimum (top) and maximum (bottom) critical points for the cross-sectional cohorts. The left joints (shoulder, elbow, and wrist) have greater range than their right counterparts, depicting increased variability overtime in the left side. (B) Displays the overall change from end to beginning (difference between the last and first critical points). The head is the only body part to have a noticeable increase from 0-2 weeks to 15-17 weeks. (C) Displays the standard deviation of the minimum and maximum points.

Figure 19. Maximum and Minimum Cross-Sectional Critical Points. The minimum (dotted line) and maximum (solid line) EMD values are plotted. From 6-8 weeks onward, the right sides display more consistent patterns and a decrease in variability of the critical points. Early weeks display more changes in critical points, indicating more exploration and fluctuations in movements in earlier cohorts.
4. Discussion

The guiding principle for this work is to provide reliable, accessible, and convenient forms of early detection during initial weeks of infancy. There are a myriad of assessments and measurements evolving to detect neurodevelopmental derail. However, as shown in this work, changes occur stochastically and rapidly during early weeks of infancy. Extant research in measurements of movement analysis rely heavily on physical sensors, which are inconvenient, inaccessible, and can alter the natural movements of the infant when in use. Scaling up this approach will allow for a real-world impact, as it will be more accessible to parents within their own homes, and clinics where research grade wearable sensors are not realistic or accessible means of measurement. The most effective way to accurately track development is to adapt research-grade methods to a paradigm which can be done frequently, and without hassle within the home. This work looks to add a level of precision to current clinical assessments, while making them more accessible for families, so that they can utilize this technology with off-the-shelf and commercially available technology. The technology used in this thesis hopes to contribute to the development of measures that can be readily accessible to parents, as all it requires is a smart phone with video recording abilities. The prevalence of smart phones in our current society can allow families to utilize
methods that we introduce here, which serve to analyze such continuous data streams by digitizing them as spike trains. As such, we enable the assessment of movement as sensory feedback, producing a qualitative measurement of the type of peripheral feedback that the brain of the infant is relying on.

We can also assess the pattern of synergies that naturally emerge through the spontaneous motions that take place during development. These motions begin as unintentional motions with random noise. Yet, eventually they transition to well-structured signals – giving rise to systematic, intentional movements. As the movements become intentional and goal-directed, they are utilized in a forward prediction model which can then provide the infant’s system with the necessary information to estimate the sensory consequences of their own movements. Estimating the consequences of impending actions and compensating for internal sensory-motor transduction and transmission delays, serves as a learning paradigm interfacing the infant’s body and their environment. This ability to mentally plan, control and coordinate complex bodily movements while translating external sensory goals into physical acts helps the child connect mental intent with physical volition. As the infant matures and myelination takes place, the brain will continue to develop and recalibrate accordingly. The motor signatures assessed here at the micro level permit the tracking of the quality of the peripheral kinesthetic reafference, informing the central controllers of the brain...
and the spinal cord how to coordinate actions at will. Our new methods thus provide new tools to improve the theory of general movements, interfacing it with embodied cognitive theories and extending their use from the lab and the clinics, to the comfort of the home.

In this study, we were able to empirically detect the maturation of movement of the proximal, articulated, and distal joints over time, through longitudinal and cross-sectional means. Using a personalized method of statistical analysis, we were able to detect independent maturation between infants, as well as self-emerging patterns amongst cohorts of the same age group.

### 4.1 Order of Maturation Revealed Through Quadrant Division: Proximal-Articulated-Distal

The empirically estimated shape and scale gamma parameters for each infant find themselves in either the Left Upper Quadrant (LUQ), which contains instances with higher NSR and more random regimes, or the Right Lower Quadrant (RLQ), which contains instances with high signal and higher predictability. Based on the occurrence of points in the LUQ vs. the RLQ, we can quantify the direction and the amplitude of the stochastic shift from one probability distribution to another. More importantly, based on our prior empirical research and sampling thousands of individuals from birth to 79 years of age, we have mapped the significance of the RLQ vs. LUQ locations and can
provide interpretability to these stochastic trajectories. As such, we can determine the degree of randomness and high-quality signal present throughout different weeks of development for each infant. By tracking the occurrences and their changes weekly, we can see that the proximal class enters and remains in the RLQ first, followed by the articulated, and then the distal joints (Figure 21). This pattern of convergence toward the RLQ allows us to interpret which parts of the body become more regulated with predictable and stable movement patterns, and in what order these transitions take place. The discovery aligns with suspected conclusions, as the proximal joints are core stabilizers of the body and are required to have a strengthened foundation, for the rest of the body to develop with appropriate supports. In the period assessed in this study, the distal joints do not show a clear transition to the RLQ, implying that by 16-17 weeks they still produce more random distributions and lower signal to noise ratio. Understanding this pattern of maturation can allow us to flag an infant who does not have proper stability in one group before the other. If we see, for example, that the proximal joints are not approaching the RLQ before the articulated, we can alert pediatric movement specialists that the infant may not have developed adequate core control, which is necessary for the proper development of the rest of the body. The unique contribution of these methods is that they can flag which parts of the body may need specific attention, instead of reducing the problem to a general
complication using the more traditional one-size fits all model. The benefits of this are discussed in 4.2.

In addition to identifying which parts of the body may be maturing faster or slower than others, the quadrant identification and evaluation process allows us to assess the longitudinal progression of the infants. We can evaluate maturation patterns over time and assess whether the trajectory of maturation is appropriate or stunted. In addition to maturation patterns, we can address whether an infant is producing enough general movement, as required per clinicians and general movement specialists. If an infant is not progressing at the desired rate, the parents can be notified to be more aware of their infant’s developmental process, and if persisting, take the child to a physician.

The empirically estimated Gamma Moment space helped further visualized other aspects of the LUQ and RLQ locations on the Gamma plane. These parameter spaces can help with the identification of parameters with the maximal information to further the research.

Transitions between the RLQ and the LUQ across the 16-17 weeks when both infants were studied gave us a sense of the typical pattern of maturation that the joints underwent. The data showed that although the RLQ is a location corresponding to higher signal and more controlled movements, during these early stages of neurodevelopment, there are frequent excursions to the LUQ
region, hosting distributions closer to the memoryless random exponential and with higher NSR. The results indicate bouts of exploratory activity sampling the here and now in a memoryless fashion, interspersed with systematic trial and error-correction, converging toward self-discovered goals. We can ascertain that as infants’ micro-motions progress to the RLQ, they are more engaged with their environment. They are becoming more aware of their bodies in motion, and are beginning to acknowledge more cognitively the consequences of the brain’s self-generated peripheral motions. As this progression unfolded, in both babies, we witnessed the orderly maturation of the proximal to the articulated, to the distal linkages of the body in motion (Figure 21).

**Order of Maturation**

- **Proximal**
  - Slower Movements
  - Lower Variability
- **Articulated**
- **Distal**
  - Faster Movements
  - Higher Variability

*Figure 21: Order of Maturation.* Orderly maturation of the proximal, to articulated, to distal linkages of the body. Based on the transitions between the RLQ and the LUQ across the 16-17 weeks. The movement patterns from the proximal class were the first to progress toward increasingly controlled motions with higher signal quality and more predictable distributions, later followed by the articulated, and then the distal.

### 4.2 Identification of Differentiating Signatures Between Infants.

Besides identifying patterns common to both infants and to the cross-sectional cohort, we were also able to identify individual signatures unique to each
infant. Through the tracking of the critical points, we captured nuances of each infant’s developmental trajectory. For example, Figure 14 captures differential transitions across the joints of the two babies. The transitions of each trajectory (signaling stochastic shifts between distributions) uniquely define the infant’s maturation pattern. In the context of e.g., a treatment, these patterns would inform the clinician on improvements (towards the RLQ regimes) or setbacks (stuck in the LUQ regimes) that one would have to avoid. Probing different stimuli will give rise to different responses that we can track using these parameter spaces to optimize performance. This is so, because we know (from our prior empirical work) which region of these spaces are conducive of good maturation and which regions are conducive of pathological trajectories.

The identification of target for treatments specifically tailored to the child’s evolution of the stochastic patterns offers new alternatives to the longitudinal tracking of treatments. The traditional approach of bucketing infants into general groups and offering a blanket, one-size-fits-all model for treatment, prevents us from leveraging the plasticity of the early neurodevelopment, and from helping the child flourish. Infants who clinically present similarly often have very different underlying mechanisms at play, causing certain movement issues. As such, movement output variability can be our best ally in the treatment of
neurodevelopmental disorders. Movements and their sensation help us track the progression of the child, and isolate the gains induced by the specific treatment.

Exclusive reliance on subjective observation misses an opportunity to identify causal roots of a disorder and treat the person based on needs for support. For example, autism (ASD) is considered a spectrum disorder with a wide array of symptoms. These symptoms can be quite general, such as impairments in social interaction and communication. Once an ASD diagnosis is given, the main target becomes changing the social nature of the patient. However, this overlooks the potential developmental, medical, and/or genetic underpinnings of this highly heterogeneous condition, and the adaptations of a person’s individual nervous system are therefore not properly supported and nurtured. To that end, targets for treatment based on a one-size-fits-all model of social norms can be unproductive at best, and detrimental to the persons development at worst. However, a personalized signature can bring the clinical world a step closer to helping the individual. A one-size-fits-all model is not ideal given the coping nature of the nascent, and very plastic nervous system. Understanding individual differences in development and maturation, from birth and being able to characterize such differences broadly across the population, will allow us to provide more individualized interventions and therapies for infants that may be flagged for some type of neuromotor delay or abnormality.
4.3 Self-Emergence of Development Patterns Throughout the Body

In this work, when the probability distributions of the different age ranges were compared, we were able to identify distinct self-emerging patterns in different parts of the body. The videos used in this consisted of the longitudinal participants, B1 and B2, as well as the videos acquired from YouTube. Even with this amalgamation of a dataset, it was clear, for example, that the head micro-motions registered in 5/6 infants in the age range of 15-17 weeks, converged to similar distributions. Such patterns persisted, at different age ranges, throughout all the body parts. Though individual infants had unique stochastic trajectories (from personalized fluctuations) they all converged around similar time frames, to a common pattern. This self-emerging pattern depicts the universality of infant development, and lends support to the use of these methods. Indeed, we now know what to expect when using a smart phone to collect weekly videos below a 2-minute duration.

This shared pattern of development underlies a new dynamic and stochastic metric to contribute to a more individualized digital growth chart, which looks at the components of development independently, and yet relies on the universality of the human neurodevelopmental trajectory. This proposition for a dynamic neurodevelopmental chart adds non-trivially to the current growth chart to check developmental patterns. An infant may grow, but if no neuromotor
development towards autonomous control accompanies the growth, the infant will not develop agency towards independence.

We can also use this information to track variability throughout different ages. The lack of consistency amongst body parts, especially articulated joints, emphasizes the need to study the body as a whole. Current general movement practices, when looking at writhing or fidgeting movements, tend to focus on the outer most parts, such as the wrists and ankles, as jerking and general movement is most evident at these parts (the distal joints). However, as we saw through the diverse set of self-emerging patterns present in the cross-sectional data, there is important information at the joints. These joints contribute to synergies which simplify the brain’s effort to control the DOF and as such, tracking their evolution can inform us of the progress that the brain may be making (or lacking) in resolving this major feat.

4.4 Limitations

The current study has several limitations. First, the longitudinal results are based on only two participants. The cross-sectional results are based on only an additional seven. To make more robust conclusions about estimates of the stages of maturation, and the emergent patterns within different age ranges, we should include many more participants. Specifically, it would be helpful to include infants presumed to be healthy, and infants who may be deemed at risk, e.g., having
siblings in the spectrum of autism. Adding large cohorts of infants in various families of neurotypical and atypical development is not only warranted, but also now feasible, thanks to the proof of concept that this study provides.

Besides the low number of participants, we lacked clinical or personal information for the cohort extracted from YouTube. Besides their ages, we did not know much more. For B1 and B2, only age and gender are known, though we continue to monitor their neurodevelopment. It would be helpful to know height, weight, head circumference, medical history of the infant and of the parents and/or siblings. Having the physical data of the infants would serve as another metric of growth and could be combined with the digital data. Expanding on this, it would be helpful to have clinical scores of the infants. Prechtl’s assessment can be done from a video recording, and therefore it is possible to do this in the future with the aid of clinicians. Lacking this information limits us in connecting what is currently the clinical standard to a new digital paradigm. Having these scores would allow us to better characterize the movements which have been well studied and characterized through traditional clinical methods. Along these lines, we are now collaborating with a large clinic in the state of New York which is providing us with the medical history of infants to deploy our study at scale.

An important component left out of this work was the lower half of the body, as we only studied the upper half of the body. The lower half, specifically
kicking movements, have been well characterized in developmental literature and are vital in studying neuromotor development. Due to clothing and obstructions in the videos, we were not able to retrieve data with good enough confidence to use in this work. This will be remediated with specific instructions to the parents now that we are more aware of the limitations of pose estimation of infants.

Furthermore, the latest videos have very good quality in the legs.

Lastly, the statistical platform in this work focused on a personalized assessment. While the empirical estimation in our statistical analyses allowed us to take advantage of metrics and parameter spaces for individualized analyses, we are also interested in dyadic interactions between the infant and the caregiver. In future studies we plan to include motions of the mother in tandem with the infant’s motions.
Figure 22. Informative Poses for the Evaluation of Motor Skills Acquisition. These five (5) poses are selected for further evaluation in future work. They have been classified as positions which convey information toward the potential for skill acquisition and motor development during the first year of life. The poses have been studied monthly in infants, and the transition in skill an infant has when performing such poses is informative of the development of the kinesiology underlying varying motor skills and/or motor problems.

4.5 Future Work

In the future, we plan on expanding this study to include more participants. We plan on recruiting participants who are comfortable providing the necessary information required to participate in this study. We also plan on expanding the participants pool to include infants who are deemed healthy, as well as those born pre-maturely, infants with family members who have or who themselves have known genetic issues which lead to neurodevelopmental disorders, and infants who had birth complications. This would provide a comparative model to infants who are on what is deemed a developmentally good trajectory, and allow us to better characterize, detect, and train for infants at-risk. Participants will also be asked to avoid clothing around the legs of the infant, so that the lower body can be included in the pose estimation analysis.
This extension of the study will request more information from participants. This will include any relevant health information for the infants and their parents. Additionally, more physical measurements will be requested, including body weight, body height, and head circumference, at the time of each video submitted. We will also collaborate with clinicians to retrieve clinical scores for the infants in the videos as a frame of reference for the current clinical metrics and milestones of development.

In measuring these milestones, we will also incorporate additional poses which have been clinically characterized, and the performance of which convey information toward potential for skill acquisition and motor development during the first year of life [39]. These are shown in Figure 22 and include: (i) the supine position as studied in this thesis, (ii) the prone position, which consists of the infant lying on their stomach, (iii) pulled to sit, which involves the infant laying on their back and being gently supported and pulled by the arms, to bring them to a sitting position, (iv) supported sitting, which consists of supporting the infant to remain in a sitting posture by providing support to the back, and (v) standing/supported standing, which involves holding and/or supporting the infant to be in a standing position, until they are capable of doing so themselves.

These positions have been well characterized by clinicians systematically using them month by month throughout the first year. As such, we know what to
expect visually, when observing the infant. However, using our video-based data acquisition and biometrics throughout development in the first year, will provide precise benchmarks for motor development. We have mentioned the importance of reafferent feedback to the brain, which is necessary to ensure development. These postures are indicative of another extremely important mechanism in development, “feedforward control” based on intent and error correction [39]. This mechanism is critical to develop a predictive code and compensate for internal delays in sensory transduction and transmission across the motor systems. Understanding how these mechanisms mature in tandem with the emergence of theory of mind and other cognitive abilities will help us further bridge the principles of kinesthetic reaference, and the internal models of neuromotor control with principles of embodied cognition. We will do so while characterizing the emergence of anticipatory postural adjustments, many of which have been evaluated in the positions (i)-(v) [40].

Lastly, we will extend this work from 2-dimentional poses which are acquired by video, to 3-dimensional means by utilizing new models and validating such models with non-invasive physical wearable sensors. To advance data acquisition toward more advanced video-based methods, they must be first validated and developed with physical measurements as ground truth. The
methods presented here offer ways to validate such data and move the field toward new computational approaches to neurodevelopmental research.
References


