

1 Digital Screener of Socio-Motor Agency Balancing Autonomy and

2 **Control**

3 Theodoros Bermperidis¹, Richa Rai², Elizabeth B Torres^{1,2,3*}

- ⁴ ¹Sensory Motor Integration Lab, Psychology Department, Rutgers University, Piscataway, NJ USA
- 5 ²Rutgers University Center for Cognitive Science, Piscataway, NJ, USA
- ³Rutgers University Center for Biomedicine Imaging and Modelling, Computer Science Department,
 Piscataway, NJ, USA
- 8 * Correspondence:
- 9 Elizabeth B Torres
- 10 ebtorres@psych.rutgers.edu

11 Keywords: autism, socio-motor agency, autonomy, control, entropy, stochastic analyses, signal-

12 to-noise ratio, wearables, ADOS

13 Abstract

- Dyadic social interactions evoke complex dynamics between two agents that while exchanging unequal levels of body autonomy and motor control, may find a fine balance to take turns and gradually build social rapport. To study the evolution of such complex interactions, we currently rely exclusively on subjective pencil and paper means. Here we complement this approach with objective biometrics of socio-motor behaviors conducive of socio-motor agency. Using a common clinical test as the backdrop
- of our study to probe social interactions between a child and a clinician, we demonstrate new ways to streamline the detection of social readiness potential in both typically developing and autistic children.
- 21 We highlight differences between males and females and uncover a new data type amenable to
- 22 generalize our results to any social settings. The new methods convert dyadic bodily biorhythmic
- 23 activity into spike trains and demonstrates that in the context of dyadic behavioral analyses, they are 24 well characterized by a continuous gamma process independent from corresponding binary spike rates.
- 24 Well characterized by a continuous gamma process independent from corresponding binary spike rates. 25 We offer a new framework that combines stochastic analyses, nonlinear dynamics, and information
- 26 theory, to facilitate scaling the screening and tracking of social interactions with applications to autism.
- 27

28 1 Introduction

29 All research involving autism is (arguably) fundamentally tied to the Autism Diagnosis Observation 30 Schedule (ADOS, currently in version 2 [1; 2; 3; 4].) Research spanning disparate fields, from genomics to complex social interactions relies on this test as the gold standard to classify humans 31 across the lifespan as autistic or autism spectrum. Although clinically validated, the ADOS-based 32 33 diagnosis misses females [5; 6; 7]. Moreover, there are not enough raters to absorb the large number of toddlers, children, and adults that according to various screening tools, are suspected as autistic 34 today. The test is long and taxing on both the child and the clinician administering it because it has an 35 average of 27 tasks aimed at engaging the child through social presses and expecting overtures from 36 37 the child.

38

39 The ADOS is a dynamic and flexible test in the sense that the clinician can choose the tasks according 40 to the flow of the child's performance. It also adapts the test on demand, choosing the module that best 41 agrees with the child's communication level. However, the interaction occurs while the clinician also 42 scores the child's performance. Though valid to probe social competence, many of the tasks artificially 43 rob the child of a chance to be naturally social, as the interaction is also taxing on the clinician and at 44 times, awkward and seemingly forced. In this sense, several of the tasks might be biased, interfering 45 with the child's agency, and robbing the clinician of the spontaneity characteristic of a natural social 46 exchange. In this sense, we need objective ways to quantify this potential bias that such a taxing effect 47 may produce on both social agents.

48

49 Prior work analyzing thousands of ADOS score records found non-obvious issues with the statistical foundations used to validate this test. While there are theoretical requirements of normality and 50 51 homogeneous variance in the signal detection theory used to validate the ADOS [8; 9], as these 52 assumptions are required for independence between bias and sensitivity [10], the empirical data across 53 thousands of records, violate these assumptions [10]. New methods have then been proposed to help 54 reduce the number of tasks [11], while also utilizing motor signatures to identify females [4; 11; 12; 55 13]. However, there are no means to define naturalistic social agency in the dyad and to identify tasks 56 that enhance it. Furthermore, no means to implement these tasks using artificial intelligence (AI) and 57 machine learning (ML) methods have been proposed. Such approaches would help us speed up, 58 automate, and scale the assessment process, particularly doing so with respect to currently 59 underdiagnosed females [5].

60

61 We reasoned in the present work that the digital ADOS [11], *i.e.*, the ADOS that is digitally recorded 62 while the child and clinician interact, could leverage the validity of this test as the gold standard for

63 clinical and research use, while providing a streamlined version of it that could help us (1) identify

64 objective biometrics of social agency and (2) automate the process of identifying socially compliant

65 tasks using methods from Artificial Intelligence (AI) and Machine Learning (ML).

66

In our approach, socially compliant tasks are those which provide social agency to the child that is being diagnosed. More precisely, we here define social agency as the balance between autonomy and control during a social exchange. Autonomy is defined as the ability of the child to lead the conversation as much as the clinician does, rather than always following the lead of the clinician. Control is defined as the ability of the child to effectively predict the consequences of impending social actions and overall behaviors, based on intact motor control. Both concepts are illustrated in Figure 1.

- 73
- 74
- 75



77 Figure 1. Digitally characterizing rapport and turn taking during an ADOS-based social 78 interaction. (A) Snapshots of the interaction between a child and clinician wearing 6 biosensors, 3 on each body, synchronously registering motion at 128Hz. Clinician-led interaction and note 79 80 taking while rating the interaction. (B) Sample standardized micromovement spikes (MMS) derived from angular speed capturing approximately 11.7seconds of social exchange. (C) 81 82 Frequency histograms of the MMS peaks, (one frame) from each sensor on the child and 83 clinician. (D) Pairwise comparison of the histograms evolution using the Earth Movers' distance similarity metric. Entries reflect the 6x6 matrix (child and clinician, 3 sensors at the torso, right 84 and left wrist) as in (A). Off diagonal entries are the shared dyadic space, while entries next to 85 86 the diagonal are the child's or clinician's activities in standalone mode. Blue to yellow color EMD

87 scale ranks from most to least similar spike patterns. Star marks the maximal similarity.

88

76

89 Socio-motor agency can be impeded if neurodevelopment undergoes a different maturation path [3; 90 14: 15]. If the child, for example, has excessive motor noise and motor randomness in its performance, 91 the predictive ability required for self-motor control will be compromised [3] and with it, the overall 92 control ability will be altered. This alteration will also in turn affect the clinician's perception of the 93 child's nuanced micro-motions underlying social behaviors, thus biasing the assessment [3; 11; 14]. 94 Under such circumstances, socio-motor agency can be impeded, as can be the rating of the child by the 95 clinician. Therefore, the tasks that manifest excess random noise of the joint dyadic motor patterns (lower control of the dyad) and / or excess lead of the clinician within the dyad (lower autonomy of the 96 97 child), are inevitably bound to bias the clinician's scoring towards a deficit model of autism. In contrast, 98 the tasks that manifest high dyadic control and autonomy of the child are bound to boast social agency, 99 according to our biometric definition. These tasks can provide a more appropriate model of readiness 100 potential for social exchange. While detecting a problem relative to normative data, this new model 101 can also do so in a fair, unbiased manner. In this sense, the child has a chance to succeed. In turn, the 102 clinician can presume competence and identify areas of strength to recommend treatments more 103 appropriately. Such treatments will rather be grounded on the non-obvious, nuanced aspects of 104 behaviors occurring at a micro-level that escapes the naked eye. Yet they will be quantifiable with biosensors that read out biorhythmic activities from the nervous systems with sub-second resolution. 105

106 We here introduce a theoretical framework grounded on empirically derived power (scaling) laws of

- 107 human ontogenetically orderly (neurodevelopmental) maturation. This framework connects stochastic
- analysis of human biorhythmic (time series) data with information theoretical metrics. We define new
- truly personalized computational indexes of dyadic control, autonomy, and socio-motor agency from
- biosensors' digital data using as guidance the digitized ADOS-2. Then, we identify socially compliant tasks *i.e.*, ADOS-2 tasks with balanced socio-motor agency, thus streamlining the digital ADOS-2.
- Lastly, we propose new ways to help automate and speed up autism screening and detection based on
- these socially appropriate tasks identified from the motor variability of the interactive dyad, rather than
- 114 from the child's or the clinician's performance alone.
- 115

116 2 Methods and Analyses

117 2.1 Participants

118 A total of 29 children including 19 males and 10 females spanning 4-15 years of age and two adult 119 clinicians participated in the study (See Table 1). Children participated in multiple sessions over the 120 span of 2 $\frac{1}{2}$ years with one clinician per session and were administered a specific module per session, 121 *i.e.*, a specific subset of ADOS tasks, in accordance with their age, level of development and spoken 122 language.

Record Number	Participant ID	Age	Sex	Visit 1 Module (Total Score)	Visit 2 Module (Total Score)	Visit 3 Module (Total Score)	Visit 4 Module (Total Score)
1	NT01	(7-9)	F	3(0)	-	-	-
2	NT02	(10-12)	F	3(0)	-	-	-
3	NT03	(7-9)	F	3(2)	-	-	-
4	NT04	(10-12)	F	3(2)	-	-	-
5	NT06	(7-9)	М	3(2)	-	-	-
6	NT08	(7-9)	F	3(0)	-	-	-
7	NT09	(7-9)	F	3(1)	-	-	-
8	NT10	(10-12)	М	3(1)	-	-	-
9	NT11	(13-15)	М	4(2)	-	-	-
10	NT12	(10-12)	F	4(1)	-	-	-

123 Table 1. Participants information.

11	NT13	(13-15)	М	4(0)	-	-	-
12	EP01	(4-6)	М	3(10)	2(7)	3(9)	2(8)*x
13	EP02	(7-9)	М	3(9)	2(8) *X	-	-
14	EP03	(10-12)	М	3(12)	2(13)	3(24) *X	2(21) *X
15	EP04	(13-15)	F	3(7)	4(8) *X	3(11) *X	4(8) *X
16	EP05	(4-6)	М	3(9)	2(7)	3(8) *X	-
17	EP07	(10-12)	М	3(12)	2(9) *X	3(18) *X	2(13) *X
18	EP09	(4-6)	М	1(18) *X	1(16)	-	-
19	EP10	(7-9)	М	1(13)	1(17)	-	-
20	EP13	(4-6)	М	3(17)	-	-	-
21	EP14	(13-15)	М	1(8)	2(10)	1(14)	2(15) *X
22	EP15	(10-12)	М	1(17)	1(15) *X	1(26)	1(21) *X
23	EP16	(4-6)	М	3(11)	2(11)	3(22) *X	2(20) *X
24	EP17	(10-12)	М	1(16)	1(18) *X	1(18)	1(19) *X
25	EP18	(7-9)	F	3(10)	2(11)	3(16)	2(10)
26	EP19	(7-9)	М	3(8) *X	2(8) *X	3(8)	-
27	EP20	(7-9)	М	3(11)	2(11) *X	3(17) *X	2(16) *X
28	EP21	(10-12)	F	3(11)	2(9) *X	3(8) *X	2(11) *11
29	EP22	(4-6)	М	1(26)	1(24) *X	-	-

124 *X Denotes that the participant came for the visit, but his/her data could not be used either because it 125 was not available/lost/corrupted or the information available was incomplete. NT stands for 126 neurotypical participant and EP stands for expected-asd participant. The latter were confirmed by the 127 clinician to have an ASD diagnosis at the end of the session. The identifiers EP/NT are not known 128 to the clinicians or parents and are used as a way to number the participants internally in our 129 lab.

131 2.2 Raw Data Acquisition

Digital data were acquired during each session using light wearable sensors (APDM Opals, Portland, OR, USA). Six sensors were used, two on the left and right wrist and one on the torso, both on the child and clinician. The sensors continuously and synchronously recorded triaxial accelerometry and gyroscopic data at a sampling frequency of 128 Hz. The recording environment followed the standardized ADOS requirements using similar table and sitting arrangements for the clinician-child dyad. The two clinicians were unaware of the goals of the study.

138 2.3 Data Type: The Micro-Movement Spikes Derivation

Scalar values of angular speed from orientation data that the gyroscopes recorded (or acceleration from the inertial measurement units) were acquired using the Euclidean norm (using Equation 2.1) of the coordinate components of motion as measured by the sensors:

142
$$V = \sqrt{V_x^2 + V_y^2 + V_z^2}$$
 Equation (2.1)

143 From here onward, all analyses refer to the scalar value V, the angular speed in deg/sec.

144 Human motion data such as angular speed and acceleration, are inherently biased by allometric effects 145 and anatomical differences across subjects. To scale out such artifacts, we normalize motion data 146 fluctuations (peaks and valleys) (using Equation 2.2) as relative deviations from the empirically 147 estimated Gamma mean. The mean is estimated by fitting the continuous Gamma family of 148 distributions to the raw peak data. The Gamma family of distributions has been consistently found to 149 be the best candidate to fit human peak activity motion data, according to Maximum Likelihood 150 Estimation (MLE). After we shift and center our data around the Gamma mean, we scale and map peak 151 motion activity to the [0,1] interval according to the local minima average:

152
$$NormPeak = \frac{Peak}{Peak+Avg_{min to min}}$$
 Equation (2.2)

153 The normalized peak series, called Micro Movements Spikes $(MMS)^{TM}$ conserve the temporal 154 structure of the original speed/acceleration time series. They represent "*quiet*" times interspersed with 155 bouts of activity away from mean activity.

156 **2.4 The Gamma Process of the MMS**

157 The normalized speed MMS are best fit (in the MLE sense) by the continuous Gamma family of 158 probability distributions [3; 16]. Furthermore, the parameters of the Gamma distribution, shape k and 159 scale θ have been found across multiple studies from our laboratory, including the present one (see 160 Results), to follow a Power Law of the form described in Equation 2.3:

162
$$k \cong a\theta^b \to log(k) = a + b log(\theta) + \varepsilon$$
 Equation (2.3)

163

164 Where ε is a small error term and b < 0. This Power Law for the standardized MMS time series reveals 165 a maturation process of the motor code for voluntary [3; 17] and involuntary [18] motions. This law is

¹⁶¹

166 very important because it provides us with a quantitative framework to interpret fluctuations in 167 biorhythmic data that range from random to predictive.

168

- 169 Importantly, the continuous Gamma family of probability distributions has the first (mean) and second 170 (variance) moments expressed in terms of the shape and scale described by Equations 2.4)
- 171

172
$$\mu = k\theta, \sigma^2 = k\theta^2$$
 Equation (2.4)

173

Then, using Equations 2.4, the Noise-to-Signal Ratio (NSR) of the MMS reduces to the Gamma scaleparameter as in Equation 2.5:

176
$$NSR = \frac{\sigma^2}{\mu} = \frac{k\theta^2}{k\theta} = \theta$$
 Equation (2.5)

177

178 The Gamma scale parameter in Equation 2.5 fully characterizes the noise of the motor patterns of the 179 interactive dyad (or of the participant), *i.e.*, in relation to the level of fluctuations of angular speed 180 during the ADOS activities.

181

Empirical estimation of these parameters in thousands of participants over a decade of work with 182 183 humans along the lifespan, and across disorders of the nervous system, has revealed an interpretation 184 for the Gamma log-log parameter plane. Distributions that fall along high NSR regimes are also close 185 to the memoryless random regime of the special exponential distribution case (when k = 1). Points in mid NSR correspond to heavy-tailed Gamma distributions. Then, low NSR (or high signal = 1/NSR) 186 187 are congruent with symmetric shapes (Gaussian-like) distributions. High-signal Gaussian regimens are highly predictable in contrast to High-noise memoryless random Exponential regimes. As such, this 188 189 parameter plane is empirically interpretable.

190

191 **2.5** Quantifying Motor Control from the Perspective of an Agent

192 Noise-to-Signal Ratio measures the degree of motion variability away from mean activity. Small NSR 193 characterizes steady and smooth motion, akin of goal-oriented behavior, as experienced from the 194 perspective of the agent/ child. On the other hand, a high NSR indicates unpredictable and random 195 motion. In that sense, the NSR is a proxy for motor control and quantifies the existence of predictable 196 motor patterns. Because the NSR is calculated on the standardized MMS, motor noise does not depend 197 on the anatomy of the individual as it is scaled by the mean amplitude of motion.

198

199 2.6 An Information Theoretic Approach to the Analyses of the MMS

The presence of MMS peaks indicates an outburst of activity away from baseline. This is informative of a motor activity. When we also consider the temporal distribution of MMS peaks, a train of such spikes can be viewed as a representation of information regarding human motion variability through time. When we consider multiple sensors sampling in synchrony, the MMS spikes carry spatiotemporal information about the bursts of distributed bodily activity in the motor system.

205 2.7 Binary Trains of MMS

If we transform the MMS data so that the presence of a peak corresponds to the binary "1" and an absence of a peak corresponds to the binary "0", we can represent normalized speed (or acceleration) as a stochastic binary sequence. An underlying mechanism stochastically generates bursts of activity, and this is equivalent to randomly generating 0s and 1s from an underlying binary alphabet as in Equation 2.6.

- 211
- 212 $B_t = 1$, if NormPeak_t > threshold Equation (2.6)
- 213

214 Let's assume that B_t is a random sample drawn from an underlying probability distribution at time t.

215

216 $B_t = 1$ with probability p_t , 0 with probability $1 - p_t$ Equation (2.7)

217

Entropy H in Equation 2.8 is an information theoretic measure that quantifies the amount of information in a random variable that follows a probability distribution P_x [19] and is equal to:

- 220
- 221 $H = -\sum_{x} p_x \log_a p_x$ Equation (2.8)
- 222

In the case of the binary process, the amount of information of the random variable of an activity outburst (MMS) is given in Equation 2.9:

- 225
- 226 $H_t = -p_t \log_a(p_t) (1 p_t) \log_a(1 p_t)$ Equation (2.9)
- 227

Which takes the maximum value of 1 when $p_t = 0.5$ and the minimum value of 0, when $p_t = 0$ or 1. Intuitively, entropy measures either the uncertainty regarding the outcome of a random realization of the random variable before that variable is measured or equivalently, the amount of information we get when we observe the variable. If we know for examples, that with a 100 % chance $B_t = 1$, the entropy is zero as we have no uncertainty about the outcome of the measurement, and no valuable information

is provided to us. However, if with a 50 % chance $B_t = 1$, the entropy is at its maximum because we are totally uncertain whether the outcome will be 0 or 1 and observing the outcome gives us maximal

information, specifically, 1 bit of information in the case of a base 2 logarithm (a = 2).

236

237 2.8 Measuring Randomness vs. Predictability Using Entropy Rate

The definition of entropy can be generalized for the case of multiple random variables $X_1, X_2, ..., X_N$, as in equation 2.10, by considering the joint probability distribution $P_{X_1, X_2, ..., X_N}$:

240

241
$$H(X_1, X_2, ..., X_N) = -\sum_{X_1, X_2, ..., X_N} P_{X_1, X_2, ..., X_N} \log_a(P_{X_1, X_2, ..., X_N})$$
 Equation (2.10)

242

In the case of a stationary stochastic process X (*i.e.*, statistical properties preserved over time) which takes values from a discrete alphabet K (in the case of the binary MMS), we can define the entropy rate of the process as in Equation 2.11:

246

247
$$H(X) = \frac{1}{T} \lim_{T \to \infty} H(X_1, X_2, ..., X_T)$$
 Equation (2.11)

248

This quantity measures how much the process changes over time, *i.e.*, the information that is carried in a new value. It measures the degree of randomness (unpredictability) of the underlying dynamical system [19; 20; 21; 22].

252

253 2.9 Randomness for Dynamical Systems

The concept of entropy rate is not limited to random processes, but it can also be defined in the case of deterministic dynamical systems. Let x_t be a continuous univariate times series. Then we can construct a state-space representation of the process as in Equation 2.12, if we choose an appropriate dimension *d* of the presumed underlying dynamical system and an embedding delay τ [23; 24].

259
$$X_t = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_d \end{pmatrix}_t = \begin{pmatrix} x(t) \\ x(t+\tau) \\ . \\ . \\ . \\ x(t+(d-1)\tau) \end{pmatrix}$$
 Equation (2.12)

The existence and calculations of the embedding dimension and delay are ensured by Taken's embedding theorem [25]. For more information on dynamical systems theory see *e.g.*, [26; 27]. Essentially, any univariate time series can be viewed as being sampled from a high dimensional dynamical system [28]. The dynamical system follows a trajectory in the d-dimensional space defined by the *d* degrees of freedom. All possible states of the dynamical system define the phase space of the system.

266

If we partition the phase space across F dimensions, with $F \le d$, we have an F-dimensional grid of cells of volume r^F . Then, we can measure the state of the system at constant time intervals equal to the embdedding delay τ . Then we can define the joint probability $p(i_1, i_2, ..., i_d)$ that X_{τ} is in cell $i_1, X_{2\tau}$ is in cell $i_2, ..., X_{d\tau}$ is in cell i_d . The degree of "randomness" of the determistic system can then be calculated using the Kolmogorov-Sinai (KS) entropy[29] using Equation 2.13:

272

273
$$KS = -\lim_{\tau \to 0} \lim_{r \to 0} \lim_{d \to \infty} \frac{1}{d\tau} \sum_{i_1, i_2, \dots, i_d} p(i_1, i_2, \dots, i_d) \log_a p(i_1, i_2, \dots, i_d) \quad \text{Equation (2.13)}$$

274

The KS entropy is almost always equal to the entropy rate of the original signal x_t and characterizes the degree of randomness of the system (and subsequently the sampled one-dimensional signal). For completely deterministic systems it is equal to zero and it is infinite for random systems.

In practice, the entropy rate is approximated using what is known as the correlation integral [30] in Equation 2.14:

280

281
$$C_d(r) = \lim_{N \to \infty} \frac{1}{N^2} [\# (n,m), \left(\sum_{i=1}^d |X_{n+i} - X_{m+i}|^2\right)^{\frac{1}{2}} \le r]$$
 Equation (2.14)

282

i.e., the (#) number of pairs of trajectory points that are close to each within a tolerance threshold r and measures the regularity (frequency) of patterns like a given template of specific length.

It can be shown that:

286

287
$$\lim_{d\to\infty,r\to0}\frac{1}{\tau}\log_a\frac{C_d(r)}{C_{d+1}(r)}\sim K_2$$

288

289 Where K_2 is the Renyi entropy of order 2. The Renyi entropy K_a in Equation 2.15 is a generalized 290 form of the usual Shannon entropy and is defined as:

291
$$K_a = \frac{1}{1-a} \log_a(\sum_x p_x^a)$$
 Equation (2.15)

This is a provisional file, not the final typeset article

We leverage these tools to calculate the entropy rate in the case of a discrete time series u(n). Consider two different blocks of length *m* sampled from the time series:

294

295
$$x(i) = \{u(i), u(i+1) \dots, u(i+m-1)\}$$

296
$$x(j) = \{u(j), u(j+1), \dots u(j+m-1)\}$$

And define the distance in Equation 2.16:

298

299
$$d[x(i), x(j)] = \max_{k=1,2,\dots,m} (|u(i+k-1) - u(j+k-1)|)$$
 Equation (2.16)

300

301 *i.e.*, the maximum distance between the two vectors (Chebyshev distance). Then, we can define a 302 quantity in Equation 2.17 like the correlation integral, for a template of length m at x(i) within a 303 tolerance threshold r:

304

305
$$C_i^m = \frac{\# j \le N - m + 1, d[x(i), x(j)] \le r}{N - m + 1}$$
 Equation (2.17)

306

307 Then, the entropy rate can be estimated as in Equation 2.18:

308

309
$$ER = \lim_{r \to 0} \lim_{m \to \infty} \lim_{N \to \infty} [\varphi^m(r) - \varphi^{m-1}(r)]$$
 Equation (2.18)

310

311 Where as in [31] Equation 2.19 gives:

312

313
$$\varphi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \log_a C_i^m(r)$$
 Equation (2.19)

314

Since $C_i^m(r)$ is essentially the probability that any sequence of length *m* is very close to the template sequence at time *i*, and $C_i^{m-1}(r)$ the probability that the same holds true for sequences of length m - 1, then $\frac{C_i^m(r)}{C_i^{m-1}(r)}$ is the conditional probability that any sequence of length *m* is very close to the template of length *m* at time *i* given that the same holds true for m - 1. Then $log_a\left(\frac{C_i^m(r)}{C_i^{m-1}(r)}\right) = log_a(C_i^m(r)) -$

319 $log_a(C_i^{m-1}(r))$ the logarithm of this conditional probability. It is easy to see that $\varphi^m(r) - \varphi^{m-1}(r)$ 320 is the average over *i* of the logarithm of this conditional probability [29].

However, due to finite sample sizes and stochasticity in time series analysis, the entropy rate can be estimated by what is known as Approximate Entropy [26] and is given by Equation 2.20:

325
$$ApEn(m, r, N)(u) = \varphi^m - \varphi^{m-1}$$
 Equation (2.20)

326

327 Where N is the length of the time series u(n), m is the choice of the length template and r is the 328 threshold tolerance choice. Approximate entropy measures the logarithmic frequency with which 329 segments of length m that very close together (according to the threshold), stay together through time.

331 An approximate formula for ApEn, which we implemented in our study is given by Equation 2.21:

332

333
$$ApEn(m,r,N) \cong \frac{1}{N-m} \sum_{i=1}^{N-m} \log_a \frac{\sum_{j=1}^{N-m} [\# j,d[|x_{m+1}(j)-x_{m+1}(i)|] < r]}{\sum_{j=1}^{N-m} [\# j,d[|x_m(j)-x_m(i)|] < r]}$$
Equation (2.21)

334

335 2.10 Entropy Rate estimation for a binary MMS speed sequence

- 336 Generally, in the case of discrete alphabet sequences with k symbols, $0 \le ApEn \le log_a k$ 337
- 338 Where ApEn = 0 for deterministic time series and $ApEn = log_a k$ for random series.
- 339 In our case (binary MM series), k = 2 and $0 \le ApEn_{MM} \le log_a 2$
- 340

For a = e (natural logarithm choice), the maximum value is ln(2) = 0.69, which the base we use in this study [29].

343

344 A good choice of m is equal to the embedding dimension, which can be estimated using the False

345 Nearest Neighbor (FNN) algorithm [32]. Usually, m is of low dimension, in our case the dimension of

346 the data was estimated to be 2. The threshold r is usually set between 0.1 to 0.25 standard deviations

of the time series [29].

348 2.11 Quantifying Information Flow Between Binarized MMS with Local Transfer Entropy

Local Shannon Entropy is defined in Equation 2.22 as the negative logarithm of the probability of an outcome x of a random variable [33]:

352
$$h(x) = -\log_2 p(x)$$
 Equation (2.22)

353

where low probability outcomes carry more information than high probability outcomes. Entropy as defined in Equation 2.23 can then be expressed as the average value of all such outcomes:

356

357
$$H(X) = E[h(x)] = -\sum_{x} p(x) \log_2 p(x)$$
 Equation (2.23)

358

359 Where E[.] is the expectation (average) operator. An estimator based on samples x_n is given by 360 Equation 2.24:

361

362
$$H(x) \cong \frac{1}{N} \sum_{n=1}^{N} h(x_n)$$
 Equation (2.24)

363

The Local Mutual Information i(x; y) and Mutual Information (MI) I(X; Y) are respectively defined in Equations 2.25 and 2.26, [34]:

- 366
- 367 $i(x; y) = \log_2 \frac{p(x|y)}{p(x)} = h(x) h(x|y)$ Equation (2.25)

368

- 369 I(X;Y) = E[i(x;y)] Equation (2.26)
- 370

371 Equation 2.26 quantifies the information that we gain when observing X after we have already observed 372 another variable Y.

373

Similarly, the Local Conditional Mutual Information and Conditional Mutual Information are given by
 Equations 2.27 and 2.28 respectively, [34]:

376
$$i(x; y|z) = \log_2 \frac{p(x|y, z)}{p(x|z)} = h(x|z) - h(x|y, z)$$
 Equation (2.27)

377

378
$$I(X;Y|Z) = E[i(x;y|z)]$$
 Equation (2.28)

379

380 It quantifies the information that we gain when we observe X after considering both Y and Z versus 381 considering only Z.

Finally, local transfer entropy quantifies the flow of information from Y to X and is defined in Equation2.29, [35; 36]:

384

385
$$t_{Y \to X}(n+1,k,l,u) = i(\mathbf{y}_{n+1-u}^{(l)}; \mathbf{x}_{n+1} | \mathbf{x}_n^{(k)})$$
 Equation (2.29)

386

Where *l* and *k* denote the length of the vectors $\mathbf{y}_{n+1-u}^{(l)} = \{y_{n+1-u-l+1}, \dots, y_{n+1-u-1}, y_{n+1-u}\}$ (storing past information of the process Y with a memory of *l* samples up to point n + 1 - u) and $\mathbf{x}_{n}^{(k)} = \{x_{n-k+1}, \dots, x_{n-1}, x_n\}$.

390

The integer u denotes the source-destination lag, *i.e.*, the causal time delay between Y and X that we are interested in when we want to calculate the transfer entropy from Y to X. For u = 1, a typical choice of source-destination lag is given by Equation 2.30:

394
$$t_{Y \to X}(n+1,k,l) = i(y_n^{(l)}; x_{n+1} | x_n^{(k)})$$
 Equation (2.30)

395

The local transfer entropy is the mutual information between Y and the future state of X, u samples ahead, conditioned on the history of X. In other words, it measures the information gained that we get about the future state of X when considering both its own past and the past states of Y versus considering only its past state. Transfer entropy is the expected information gain, averaging over all states given by Equation 2.31, [33; 35]:

401
$$T_{Y \to X}(k, l) = E[t_{Y \to X}(n+1, k, l)]$$
 Equation (2.31)

402

403 **2.12** Quantifying Autonomy of an Agent from the Perspective of the Observer

For a child/clinician dyad, we obtain the normalized MMS derived from the fluctuations in angular speed from the right- and left-wrist sensors throughout the course of the dyadic interaction. Then, we calculate the entropy rate for consecutive non-overlapping time windows, small enough to ensure stationarity but not too small, as to ensure convergence. We calculate the entropy rate both for the normalized MMS and for the corresponding binary MMS trains that we obtain by setting peak values to "1" and zero values to "zero".

410

To estimate the entropy rate we used Approximate entropy ApEn (developed by Steve M. Pincus [37]), which measures the amount of regularity or unpredictability of fluctuations over time-series data that have lengths compatible with experimental settings (unlike other measures of entropy aimed at measuring regularity but requiring very long times). There are caveats to the use of the ApEn algorithm [29]:

- 416 i. The ApEn algorithms allows self-counting when counting the number of templates that are 417 similar to a given data segment, which helps avoid the occurrence of log(0) in the calculation. 418 ii. However, when the self-similarity threshold r is very small, the template vector coincides only 419 with itself, giving ApEn low values, indicating regularity when the system may in fact, be very
- 420 irregular.
- 421 iii. ApEn is biased by a factor of $\frac{1}{N-m}$, which means that it depends on the template length and data 422 stream length.

423 ApEn generally depends on the threshold r, and the embedding delay and embedding dimension of the 424 reconstructed space (which is equal to the template length). It is generally suggested, that in order to compare the approximate entropies of different time series, all parameters must be equal. However, for 425 426 the scope of our study, we chose the threshold parameter r to be equal to 0.2σ , as recommended in the 427 literature [29]. The embedding delay was chosen according to the minimum Average Mutual 428 Information criterion, to ensure maximum novelty between consecutive samples in the reconstructed 429 space. As for the template, we chose it to be 1/R, where R is the average rate (frequency) of MMS in 430 the time window of interest. This equals to the average time-distance between two spikes and our 431 choice ensures that in the reconstructed space, the coordinates of a point in time include both zeros 432 ("quiet moments") and spikes and that the system does not bounce back and forth from a single 433 coordinate of zeros components. In this way, we can minimize any bias introduced by differences in 434 spike rates, in the computation of self-similarity by the algorithm. Sparser time windows will contain 435 the same percentage of "active" moments as denser time windows. Since it turns out that, for our 436 datasets, 0.1 < R < 0.5, we have 2 < m < 10, which according to the literature is within the optimal 437 range [29]. Moreover, since N = 1000, the bias introduced by m in the prefactor is very small.

438 ApEn is computationally efficient. One can easily see that the worst-case time complexity of ApEn is 439 $O(N^2)$. Furthermore, it has lower effect from noise in the data. If data is noisy, the ApEn measure can 440 be compared to the noise level in the data to determine what quality of true information may be present 441 in the data [29]. We here notice the difference between the criterion for randomness in the Gamma 442 parameter space, when the shape is 1, which is the special case of the memoryless exponential 443 distribution. In our empirical characterization of the MMS from the peak fluctuations, which follow a 444 scaling power law, as the shape approaches the value of 1 representing the exponential distribution 445 case, the $NSR = log\theta$ approaches its maximum levels [3]. The differential entropy for the Gamma 446 distribution has the general form in Equation 2.32, [38]:

448
$$h(X_g) = k + \log\theta + \log\Gamma(k) + (1 - k)\psi(k) \qquad \text{Equation (2.32)}$$

449 We will show later that discrete samples X_G that follow the Gamma distribution, such as the MMS,

have entropy roughly equal to $h(X_g) - log\Delta$, when Δ , is the discretization step. Because of the Power 450

Law discussed before, $log(k) = a + b log(\theta) + \varepsilon$, we have in Equation 2.33: 451

453
$$H(X_G) \cong e^{a+\varepsilon} e^{bNSR} + NSR + \log\Gamma(k) + (1-k)\psi(k) - \log\Delta \qquad \text{Equation (2.33)}$$

454

As the NSR increases, $k \to 1$ and thus, $H(X_G) \to e^{a+\varepsilon}e^{bNSR} + NSR + 1 - \log\Delta$. In Section 4, we 455 will experimentally show that, for $log\theta < 1$ in Equation 2.34: 456

457

458
$$H(X_G) = O(NSR), k \to 1$$
 Equation (2.34)

459

However, in the case of ApEn, we consider a process of the form $X_G * X_B^t$, where X_B^t is the binary spike 460 series, determining the temporal distribution of the peaks in time. In fact, we will empirically 461 demonstrate that X_B^t is almost independent from X_G , implying that ApEn measures the information 462 content of the binary spike time series, characterizing the motor code. On the other hand, randomness 463 464 in the sense of NSR (or equivalently $H(X_G)$), refers to the temporal component of the events and 465 answers the question of predictability in time, whereby predicting future events in time does not benefit from knowledge of prior or current event times. We will see later that these two elements of the Gamma 466 467 distributed MMS are indeed separable and within the current context, tend to be orthogonal.

468

469 In this sense, we propose that the entropy rate (ER) derived from ApEn is a measure to characterize 470 autonomy in the system. Since ER is a proper way to quantify regularity vs. randomness, we can safely presume that the information levels that it carries also measures the ability of an observer to predict the 471 472 motor behavior of an agent, when the two of them engage in a dyadic social interaction. For example, 473 when the clinician observes the behavior of a child that engages in repetitive and predictable motions, 474 they can easily learn their behavioral and motor patterns. This also implies that they can more easily 475 detect a problem in a child that behaves predictably and set up the context to better control the situation. 476 In this sense, the more predictable the situation is, the more control it will be afforded by the external 477 agent.

478

479 Following our argument, we redefine socio-motor agency as the balance between control and 480 autonomy. Signal-to-Noise ratio characterizes the ability of an agent to (internally) control their own 481 behavior. Entropy rate characterizes the ability of the agent to act autonomously (while minimizing 482 external control by another agent) in a social interaction.

- 484 Finally, using transfer entropy we can quantify the amount of causal influence from the clinician to the
- 485 child and vice versa, without the need to use any model or make any other assumptions.

486



487

488 Methods Figure. Digitization of the Autism Diagnostic Observation Schedule (ADOS): Angular
 489 speed samples (128 Hz) from wearable sensors on the wrists and torso of the child and clinician
 490 are normalized and binarized to obtain discrete sequences of 0's and 1's. Entropy rate estimates

491 measure the unpredictability of the underlying binary processes to characterize the agents'

- 492 autonomy in the dyadic social interaction. The analysis is performed on data from time windows
- 493 of ~7.8 secs which proved optimal to attain high confidence intervals.

494

495 2.13 The Autism Diagnostic Observation Schedule (ADOS-2) Scoring System

The ADOS-2 Modules consist of tasks that the clinician performs with the child to observe behavior related to the diagnosis of ASD and reach a conclusion. There are different Modules. Each child is administered a single Module based on their expressive language level, developmental age and their unique interests and abilities. However, they are designed in such a way that ensures that judgements about social and communicative abilities are as independent as possible from level of language ability and chronological age.

- -----

502

503 Both Modules (toddler) T and 1 are administrated to non-speaking children, Module T for ages 12-30 504 months and Module 1 for children over 31 months. Module 2 is administrated to children of all ages 505 who are using phrase speech but are not yet speaking verbally with fluency. Modules 3 and 4 are 506 administrated to individuals that are speaking with verbal fluency, with Module 3 specifically designed 507 for Children / Younger adolescents that can still play with action figure-type toys and Module 4 for 508 older adults. All Modules are administrated under the assumption that the individual can walk 509 independently and is free of visual or hearing impairments. This assumption is erroneous, but we use

- 510 the ADOS-2 test not to diagnose but to evoke social situations leading to movement patterns likely 511 present in such situations.
- 512
- 513 Our current analysis focuses on Modules 1,3 and 4. We suggest the following categorization of tasks 514 to better relate our digital biomarkers to the clinical tasks that evoke some aspect of social interactions
- and emotions present in human gestural communication, which is mediated by movements:
- and emotions present in numan gestural communication, which is mediated by movements:
- 516

517 Socio-Motor Tasks: These are tasks that engage interactive movements within the child, the clinician, 518 and jointly between the child and clinician. Construction Task, Joint Interactive Play, Demonstration Task, Cartoons, Conversation and Reporting and Break Tasks all have in common the Child's Socio-519 520 Motor behavior involvement. Construction Task consists of an interaction between the Clinician and 521 Child that involves reaching over the Clinician's arm to ask for block pieces that may form a shape. 522 Joint Interactive Play consists of a Play Sequence between the Child and the Clinician that involves 523 body movements. During Demonstration Task the Child uses their body to represent objects and mime 524 the use of each object. During Cartoon Task, the Clinician observes the Child's gestures and 525 coordination with speech. Similarly, during Conversation and Reporting body language and facial 526 expressions / gestures are observed alongside general communicative skills. During Break the Child is 527 expected to move around the room.

528

529 Emotional Tasks: These are tasks that probe the child's emotional states. Emotions, Social Difficulties, 530 Friends, Relationships, and Marriage and Loneliness all evoke strong emotional responses from the 531 Child. During the Emotions Task, the Child is asked questions about social relationships, different 532 emotions such as happiness, fear and anxiety and details about the manifestation of these emotions 533 under different circumstances. Social Difficulties and Annovance consist of questions related to social 534 interactions at school or work, such as bullying or teasing. Friends, Relationships, and Marriage are 535 designed to evaluate the Child's concepts on topics such as friendship and social relationships and the 536 questions asked can cause strong emotions in the Child. Similarly, during Loneliness task, questions 537 are asked about the concept of loneliness, which is a heavy topic, especially for Children on the Autism 538 Spectrum, that struggle with social rejection and bullying from a young age.

539

540 Abstract Tasks: These are tasks that require higher, abstract-level of cognition. Make-Believe Play, 541 Description of a Picture, Telling Story from a Book, and Creating a Story all help observe higher 542 cognitive skills. Make-Believe Play involves the use of dolls an action figures and the Child is tested 543 for their ability to perceive them as animate beings and produce imaginative sequences of actions that 544 involve these objects. Perception or the lack of it of objects as animate beings is a concept frequently 545 encountered within the context of the Theory of Mind. During Description of a Picture Task the 546 Clinician observes the Child's use of language/ communication and the level of interest in the picture presented. Telling a Story from a Book is similar but involves a story from a book instead of a picture 547 548 and humor and presumption of the feelings of the characters from the book are evaluated as well.

549 After the administration of Module 3 a scoring system is used to evaluate the levels of Social Affect 550 (SA) and Restricted and Repetitive Behavior (RRB). The scores are added up to determine a final

551 score, from 0 to 10. A score of 0 or 1 indicates Minimal to No Evidence of ASD related symptoms, scores between 2 and 4 indicate a Low Level of ASD related symptoms, 5 to 7 Moderate and 8 to 10 552 553 High. A score of 9 or more determine that the Child is Autistic whereas a score of 7 of more that the 554 Child is in the Autism Spectrum. Furthermore, Social Affect consist of Communication (Reporting of 555 Events, Conversation and Descriptive, Conventional, Instrumental, or Informational Gestures) and 556 Reciprocal Social Interaction (Unusual Eye Contact, Facial Expression Directed to Examiner, Shared 557 Enjoyment in Interaction, Quality of Social Overtures, Amount of Reciprocal Social Communication, 558 Overall Quality of Rapport) Scorings. RRB consists of scoring Stereotyped/ Idiosyncratic Use of 559 Words or Phrases, Unusual Sensory Interest in Play Material/ Person, Hand and Finger and Other 560 Complex Mannerisms and Excessive Interest or Highly Specific Topics/ Objects or Repetitive 561 Behaviors.

562 **3 Results**

3.1 Age-dependent Dyadic Motor Control Separates Neurotypical (NT) from Children on the Autism Spectrum Disorders (ASD)

565 The micro-movement spikes (MMS) derived from the biosensors' signals, offer a standardized time 566 series that scales out anatomical differences across participating children of diverse ages. This 567 standardized signal within the ADOS-2 tasks contexts, is well characterized by a continuous Gamma 568 process. Here (as in prior work involving other biosensors) the Gamma shape k and the Gamma scale 569 $\boldsymbol{\theta}$ parameters can be empirically estimated from the normalized spikes, using Maximum Likelihood 570 Estimation (MLE) with 95% confidence [3; 39]. The normalized spikes, which conserve the original 571 time latencies of the raw peaks, represent a time series of quiet times (at averaged activity) interspersed with bouts of activity evoked individually for each child and pertinent to the task at hand. We measure 572 573 these biorhythms individually for the child's and clinician's dominant hand. We also measure them 574 from the shared, synchronous activities of the social dyad composed by the child and the clinician.

575

The empirically estimated Gamma parameters localize each child-clinician's dyadic interaction for each task on the Gamma parameter plane with 95% confidence intervals (Figure 2A). These points represent the empirical probability density function (PDF) of their joint dyadic interaction. When we plot the full scatter estimated from each task in the ADOS, for all children, a tight linear relation emerges whereby the log-log plot follows a scaling power law, of the form $\mathbf{k} \cong \mathbf{a}\theta^{\mathbf{b}}$. (See methods for a more in-depth analysis of the power law and the micro-movement spikes (MMS) time series transformation).

583

This relationship, first described as a maturation law in humans' voluntary decision-making, mediated by pointing motions [3], is reproduced here for gyroscopic data reflecting joint dyadic angular speed, such that as the Gamma scale value decreases, the Gamma shape value increases. Because knowing one, we can predict the other with high certainty, we can then reduce these two parameters of interest to one parameter summarizing these motor signatures of the interacting dyad. We can also do so for

589 each individual signature, *i.e.*, those of the child and those of the clinician in standalone mode.



592 Figure 2. Summary stochastic characterization of micro-movement spikes, MMS, derived from 593 ADOS-driven dyadic interactions, using angular speed registered from the right (dominant) 594 wrist. Activity encompasses the entire ADOS session. Filled markers represent first visits to the 595 clinician, unfilled markers are subsequent visits. (A) Empirically estimated Gamma Shape and 596 Scale (NSR) for each participant using Modules 1, 3 and 4 of the ADOS test as a backdrop 597 behavioral assay. The size of the marker is proportional to the age of the participants. Empirical 598 Gamma plane of individual children and clinicians separating young from older children and 599 adults in an interpretable map of human neuromotor maturation. (B) Child's negative Gamma 600 scale parameter (log -NSR = log SNR) denotes control as a function of age. Observe (cyan line) 601 the decreasing trend of SNR with age for ASD in contrast to the opposite trend for NT (green 602 line). (C) Parameter space spanned by the empirically estimated Gamma mean (x-axis), standard deviation (y-axis) and skewness (z-axis) derived in (A). Marker size is proportional to age. (D) 603 Ouantification of Transfer Entropy for social dyads involving clinician and child obtained for 604 males and females in the NT and the ASD groups, using 6 sensors, 3 on the clinician and 3 on the 605 child, outputting time series of angular speed motion on the left and right wrists and the trunk 606 of each social agent in the dyad. Off-diagonal entries represent joint dyadic activities. 607

Importantly, the continuous Gamma family of probability distributions has the first (mean) and second
 (variance) moments expressed in terms of the shape and scale as in Equations 3.1:

610

611
$$\mu = k\theta, \sigma^2 = k\theta^2$$

612

Equation (3.1)

613 Then, unfolding Equation 3.1, the Noise-to-Signal Ratio (NSR) of the MMS reduces to:

614

615
$$NSR = \frac{\sigma^2}{\mu} = \frac{k\theta^2}{k\theta} = \theta$$
 Equation (3.2)

616

617 The Gamma scale parameter in Equation 3.2 fully characterizes the noise of the biorhythmic motor 618 patterns of the interactive dyad, *i.e.*, in relation to their joint level of fluctuations of angular speed 619 during the ADOS activities.

620

Empirical estimation of these parameters across thousands of participants over a decade of work involving humans along the lifespan, and across disorders of the nervous system, in voluntary, involuntary, spontaneous, and autonomic processes, has revealed an interpretation for the Gamma loglog parameter plane. Distributions that fall along high NSR regimes are also close to the memoryless random regime of the special exponential distribution case (when k = 1). Points in mid NSR correspond to heavy-tailed Gamma distributions. Then, low NSR (or high signal = 1/NSR) are congruent with symmetric shapes (Gaussian-like) distributions.

628

629 High-signal Gaussian regimens are highly predictable in contrast to High-noise memoryless random 630 Exponential regimes. As such, the Gamma parameter plane is empirically interpretable and the 631 locations of the distributions representing the shape and scale signatures of individual participants 632 change in an orderly ontogenetic manner whereby a decrease in the NSR is accompanied by a drecrease 633 in randomness (away from the memoryless exponential distribution at shape = 1). The lack of 634 maturation of the human nervous system is well characterized by high NSR and random fluctuations 635 previously found across ASD [3]. In this sense, we equate high SNR=1/NSR with an index of 636 controllability. As per the scaling power law, high SNR of the MMS is equated with high predictability 637 of the person's self-referenced, self-generated motor code. This motor code represents a proxy of 638 kinesthetic reafference, *i.e.*, the continuous stream of motor activity from the periphery, serving as an 639 index reflecting the quality of the motor feedback to the central controller of the nervous system.

640

641 Then, as this motor code is shared with another agent during social dyadic interactions, the distributions 642 of the joint dyadic interactions of the participant and the clinician, for the 26 participants (11 643 neurotypically developing NT and 15 ASD), can be appreciated in Figure 2A following a power law. 644 These distributions are derived from the MMS that fluctuations in angular speed produced in the 645 dominant hand (see Methods Figure).

646

Furthermore, Figure 2B shows that the log(SNR)=-log(NSR) (denoted as the index of control) of the interacting socio-motor dyad has an age-dependent pattern. In NT children, as the age increases, control tends to slightly increase *i.e.*, a slight positive trend is reflected in the slope of the line fitting the (blue

650	NT scatter), NT: interc	ept = 3.0271 p =	0.0067, slope =	0.0786, p = 0.36	52. In contrast, as	s ASD children
-----	-------------------------	------------------	-----------------	------------------	---------------------	----------------

age, control tends to decrease, *i.e.*, a strong negative trend is quantified in the slope of the line best fitting the red ASD souther intersect = $5.5217 \text{ m} = 5.65 \text{ m} 10^{-12}$ slope = 0.1074 m = 0.0462

652 fitting the red ASD scatter, intercept = $5.5317 \text{ p} = 5.65 \text{ x} 10^{-12}$, slope = -0.1074, p = 0.0463.

653

In Figure 2C we compare the two groups by localizing each participant on the Gamma moments space

- 655 spanned by the empirical mean, variance, and skewness, whereby each point represents the empirically 656 estimated moments of the Gamma PDF of joint dyadic activity for each child-clinician pair. This result
- 657 demonstrates a tendency of the joint dyad moments to separate NTs from ASD participants, as they
- 658 interact with an adult clinician, expressing marked differences between males and females.

659

660 To better appreciate the sex differences, we obtain pairwise the Transfer Entropy (TE) from child to 661 clinician and from clinician to child. See methods for an extended definition, but recall that TE is the reduction in uncertainty of predicting the future of X when we consider the process Y. In Figure 2D, 662 663 we can see for each matrix the pattern that emerges when considering the time series data from each of the 6 sensors attached to the child's and clinician's two hands and trunk. The cross terms in the off-664 665 diagonal entries of the matrix (top right-hand entries 1,4 to 1,6; 2,4 to 2,6; 3,4 to 3,6; and bottom lefthand entries 4,1 to 6,1; 4,2 to 6,2 and 4,3 to 6,3) represent the dyadic cases of child \rightarrow clinician and 666 667 clinician \rightarrow child, respectively. There, in the shared entries of the matrix we see that in ASD, males 668 show a decrease in TE values while females show an increase. In the context of the ADOS, females 669 evoke a reduction in the clinician's uncertainty predicting the impending females' motions, *i.e.*, 670 perhaps an inherent bias that partly accounts for the disparate ratio of 4-5 males per each female 671 diagnosed with ASD. We will further explore these differences to try and understand the interplay 672 between the NSR as an index of controllability (predictability) and the overall sense of socio-motor agency in each of the ADOS tasks, for males and for females. 673

674

In the diagonal sub-matrices (top left-hand entries 1,1 to 1,3; 2,1 to 2,3; 3,1 to 3,3; and bottom righthand entries 4,4 to 4,6; 5,4 to 5,6 and 6,4 to 6,6) we represent the patterns within the individual's body parts. There we appreciate higher values of TE from child \rightarrow child in both ASD males and females, with ASD females having higher TE than ASD males. As with the shared dyadic activity, here in the individual patters, the highest differences for clinician \rightarrow clinician can be appreciated in the ASD females.

681

682 **3.2** Quantifying Dyadic Social Agency Reveals Differences Between NT and ASD

High levels of NSR in the MMS fluctuations from the angular speed coincide with memoryless random regimes of motor patterns - well characterized by the exponential distribution previously found in autistic individuals [3; 40]. It has been proposed that under such random and noisy motor code, it is difficult to have high quality motor feedback contributing to a predictive code. Such predictive code would be necessary to compensate for sensory-motor and inertial time delays inherent in the nervous system [3; 14; 40; 41; 42].

690 In a dynamic dyadic social interaction such as that taking place during the ADOS, it is then difficult to 691 exert control over the interaction because presses by the clinician and overtures by the child are not 692 occurring at the expected timely rates. This temporal mismatch in autism alone can bias the rating by 693 the clinician in ways that differ between NT and ASD, but also differ between males and females. Here 694 we equate high NSR with low predictive control and posit that the type of socio-motor agency required 695 in a naturalistic social interaction will be impacted by poor controllability levels on one side of the 696 dyad. We then question whether dyadic-based control (*i.e.*, shared by the child and clinician) is 697 differentially impacted in ASD participants.

698

Another aspect of dyadic social agency is autonomy. As mentioned earlier, autonomy is the ability of the child to lead the conversation as much as the clinician does, rather than always following the lead of the clinician. An obvious way to quantify the degree to which the clinician is leading would be by using some form of causal analysis between data recorded on the clinician and data recorded on the child. As our main approach however, we choose to quantify autonomy in ways that depend on data recorded from wearable sensors on a single agent which as we will show, is intuitive and can be applied in a clinical setting, to help digitize the ADOS

in a clinical setting, to help digitize the ADOS.

706

We introduce (behavioral) spike trains from the MMS derived from the time series of angular speed. We use entropy metrics to examine the degree to which the spikes behave randomly or deterministically (*i.e.*, containing periodic, systematic patterns.) To that end, we use entropy rate, a metric well suited to interrogate the stochastic regimes of spike trains [26; 27].

711

712 From the MMS derived from time series of angular speed, recorded either from the left or the right 713 hand of the child, we derive binary sequences whereby a sequence of 1's corresponds to sudden bouts 714 of activity and 0's to "quiet" sampling periods, when no significant change above the person's average 715 activity occurs in the angular speed profile. Another way to view these binary sequences is as the 716 manifestation of an underlying "alphabet" that characterizes the predictability of the motor code. Zeros 717 and ones will appear with some probability, which we expect to change at some time scale, due to the 718 non-stationary nature of human motion. But if we restrict ourselves to small time windows (~7.8 719 seconds, determined as optimal for empirically estimated confidence intervals, upon sampling different 720 sizes), this time widow is small enough that the process can be viewed as stationary, yet large enough 721 to contain a satisfactory number of samples lending statistical power to our empirical estimation per 722 window. As such, 7.8 seconds is our unit of time for the spike trains that we derived. Using this MMS 723 per unit of time as our data type, we can then measure the degree of randomness of the child's motions, 724 by estimating the entropy rate (see Methods). Furthermore, we then compare it to transfer entropy 725 (TE) obtained from the child and clinician, a causal metric that can inform us of who leads the 726 interaction for any given task.

727

We argue that a suitable scale of autonomy is one in which, at one extreme, a high degree of randomness is a measure of a system at its highest degree of autonomy. This is the type of state where the system is uncontrollably "hidden" from the controller. There is no opportunity to control the person.

At the other end, the lowest degree of randomness leads to a systematic, deterministic pattern, highly controllable. While in the former, the child's system with excessive autonomy prevents social exchange with the clinician in that the clinician cannot control the child, in the latter, the clinician can absolutely control the child. Either extreme is detrimental to the development of rapport or turn-taking in a social exchange. A happy medium is one in which while the child preserves a degree of autonomy that enables a balanced social exchange, the clinician also partakes in a give-and-take interaction, rather than leading the child most of the time.

738

739 Figure 3. Scales of socio-motor agency according to index of autonomy. (A) Average transfer 740 entropies between child→clinician taken over windows of ~7.8 secs duration, per participant 741 (filled in markers are first visit to the clinician, unfilled markers are subsequent visits) vs. 742 autonomy (Ap Ent) reveal higher autonomy index in NT, a trend that is also quantified in (B). 743 As the child autonomy decreases, the $CL \rightarrow CH$ TE (left hand) decreases. Adding the CL past activity does not contribute more information about the CH state than looking at the CH past 744 745 activity alone. (C) This is also the case for the right hand. (D) Autonomy variability (variance 746 over the mean) throughout a session, is higher for the ASD group, both for the child and the 747 clinician involved.

748

749 We test our new hypothesis that motor autonomy relates to measures of entropy by comparing TE (a 750 measure of causality) from the child to the clinician, with entropy rate, a measure spanning a scale 751 from totally random to totally deterministic states of the spike-based code. We show in Figure 3A an age-dependent trend spanning two scatters. In older neurotypical children, the scatter aligns such that 752 753 as the child's entropy rate (denoting a scale of autonomy) increases, so does the TE denoting a causal 754 lead of the child over the clinician. In contrast a second scatter emerges for younger children whereby 755 the trend is less visible, indicating that these children's index of autonomy is not as evident during the 756 exchange and the causal lead (TE) denoting the child's lead over the clinician's lead, is less evident.

757 We can appreciate the shift in this metric of autonomy in Figure 3B where the histogram of the ASD

children is shifted to the left, indicating lower density values than NT children.

Since the left hand is not the dominant hand in these children, we plotted the histograms pertaining to the left hand as well, to see if these effects consistently emerged. We see in Figure 3C that across multiple time windows, the pdf for the neurotypical group is shifted to the right, meaning that on

average, NTs have higher values of autonomy than ASDs. Since autonomy also varies throughout sessions, plotting the autonomy variability (variance of this index over the mean of this index) for

different participants in Figure 3D shows that for ASDs, child and clinician variability is higher than

765 most NTs. This variability index tends to separate ASD from NT participants, particularly for later

visits (as the child aged, over 2 years and a half that the study spanned.)

767 **3.3** Age-Dependent Autonomy Across Children vs. Clinician's Autonomy Robustness

As we saw earlier, the SNR (1/NSR) of the control index, has trend with age that differs between the

two groups. NT children show increasing control with age, whereas ASD children show a decreasing

trend. Likewise, here we ask if the index of autonomy also changes with age. To that end, we examine

this index as a function of age across the children. We also examine it for the clinician across the

children's ages.

773

774 Figure 4. Non-equivalence of the index of autonomy and the index of control. Plots reflect the 775 average child and clinician index of autonomy for left hand motions vs. age as well as right vs. 776 left hand index of autonomy, index of autonomy variability and index of control. (A) Child index 777 of autonomy is positively and linearly correlated with age. (B) There is no trend between the clinician index of autonomy across children's ages. (C) Equivalence of index of autonomy derived 778 779 from the left hand vs. the right-hand motions. (D) Index of autonomy variability also correlates 780 between the two hands and separates NT (blue) vs. ASD (red). (E) No definite relationship 781 between index of control and index of autonomy is observed, however for small values of control 782 index there seems to be a positive trend which then becomes negative for high values. (F) Left 783 hand motions have higher variability in the index of control than do right hand motions.

785 We find that the child's index of autonomy for both NT and ASD increases with age in all cases (Figure

4A). This result reveals that the ability of the ASD child to actively participate in a dyadic interaction

is a human socio-motor developmental trait that improves with age. In contrast, Figure 4B shows that the clinician's autonomy is independent of the child's age. In this case, the adult clinician shows no

- 789 discernable trend.
- 790

791 **3.4** Indexes of Autonomy and Control are Not Equivalent

792 The child's index of autonomy and the variability of this index extracted from the sensors in both 793 hands, are linearly correlated (Figure 4C). In the case of the index of control however, there is higher 794 variability of the mean autonomy index across subjects when we use data from the left-hand sensor (as 795 shown in Figure 4D, where separation of the NT from ASD is evident). For this reason, we focused 796 our analysis on the non-dominant, left-hand motions. Furthermore, the index of autonomy derived from 797 the left-hand motions as well as its index of control, are positively correlated for small values of control index, but negatively correlated for higher values. This is shown in Figure 4E. In other words, 798 799 autonomy and control are not equivalent metrics. This can be further appreciated in Figures 4D 800 (autonomy variability) vs. Figure 4F (index of control.)

801

802 3.5 Male vs. Females Respond Different to ADOS Tasks - The case of ASD Females

Besides the quantification of *indexes of* control and autonomy as components of socio-motor agency, we rendered important to consider the heterogeneity of tasks in the ADOS' modules 1, 3 and 4 used here across children with different levels of spoken language. We grouped tasks into three main categories: Socio-Motor, requiring high motoric components (frequent movements and gestures); Abstract, tasks more "mental" in nature, requiring abstraction, theory of mind, and other cognitive components; Emotional, tasks that elicit feelings and emotional reactions, strongly *visibly* impacting the child's emotional states.

810

We calculated the average indexes of autonomy and control across all participants, derived from samples corresponding to the different ADOS tasks. Then, we assessed potential differences between ASD and NT participants, focusing on the comparison of males *vs.* females. We found that ASD males respond with lower index of autonomy than do NT males. In contrast, ASD females *vs.* NT females, manifest very modest differences, inclusive of three tasks with no significant differences (Social Difficulties and Annoyance, Loneliness (both Emotional type tasks) and Construction Task (Sociomotor type task)).

818

It is therefore clear, that ADOS tasks inherently bear a lack of differentiation between NT and ASD females, unlike their male counterparts for which the differences are large. This can be appreciated in Figure 5AB (males) and Figure 5CD (females) where we color code the task type and code it numerically according to the name of the task (Methods describe the ADOS tasks included from each module.) Modest differences were observed between females. Tasks with nonsignificant differences in females were Social Difficulties and Annoyance, Loneliness (both Emotional type tasks) and

825 Construction Task (Socio-motor type task). Notice that despite the non-significance, emotional tasks 826 have broader spread in ASD females along the index of control than do NT females. In contrast, the 827 index of autonomy has comparable spread for both. As socio-motor agency is defined as the ratio of 828 index of autonomy/ index of control, this implies that across these ADOS emotional tasks, NT females 829 have more social agency than ASD females. In contrast to females, a statistically significant difference 830 between the two male groups was observed for all tasks. ASD males shift significantly to lower values

831 of the index of autonomy across all tasks, but visibly socio-motor tasks are deeply affected.

832

Figure 5. Differences between males and females in average index of autonomy *vs.* control. Filled circles code non-significant differences at the .05 level, while non-filled circles denote significant

differences between NT and ASD participants. (A) NT females. (B) ASD females. (C) NT males.

836 (D) ASD males.

837

4 Theoretical Considerations at the Intersection of Stochastic Analyses and Information Theoretical Metrics

840 The following section of the paper aims at exploring the relationship between the temporal code of the 841 binary spikes embedded in the Gamma process, and the Gamma process itself, spanning information about the fluctuations in spike amplitude and inter-peak-timings. The latter follows a Poisson process, 842 while the former is more generally revealing of multiple overlapping processes. Part of our quest is to 843 844 try and deconvolve these overlapping processes in physiological data that contains multiple afferent 845 streams from different levels of functionality. These levels may span from autonomic (pacemaker like 846 regularities) to reflexive, to involuntary, to spontaneous and automatic, to voluntary levels of control 847 previously proposed {Torres, 2011 #241}. At the core of our proposed measure of socio-motor agency

848 lies the balance between bottom-up autonomy and top-down control, which we track through the 849 spatio-temporal code of the time series of spikes.

850

4.1 Entropy-Spike rate and NSR relationship in the case of standardized biometric data sampled from child-clinician dyadic interactions.

853 Recall that the entropy rate of a discrete process is defined as:

854

855 $H(X) = \lim_{n \to \infty} \frac{1}{n} H(X_1, X_2, ..., X_n)$ Equation (4.1)

856

- 857 The case in Equation 4.2 is when all X_i are independently identically distributed (*i.i.d.*) [43]:
- 858

859
$$H(X) = \lim_{n \to \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n) = \lim_{n \to \infty} \frac{1}{n} (H(X_1) + H(X_2) + \dots + H(X_n)) =$$

860
$$\lim_{n \to \infty} \frac{1}{n} \left(nH(X_1) \right) = H(X_1)$$
 Equation (4.2)

861

862

863 If X_i have identical entropies but are not independent, the following inequality holds in Equation 4.3 864 [43]:

865

866
$$H(X_1, X_2, \dots, X_n) \le nH(X_1) \Rightarrow \frac{1}{n}H(X_1, X_2, \dots, X_n) \le H(X_1) \Rightarrow H(X_1) \le H(X_1) \quad \text{Equation (4.3)}$$

867

The standardized MMS processes that we extract from the wearable sensors consist of non-zero values ("peaks" that are gamma distributed) and zero values ("*quiet moments*" at the person's average level of activity) and can be treated as the product between two processes. A gamma distributed process X_G^t and a binary process X_B^t , where *t* denotes discrete time. If for small time windows the process is roughly stationary, then at each point in time we have the processes X_G and X_B , respectively. Then, from Equation (4.3) we have the upper bound for the entropy rate expressed in Equation 4.4:

874

875
$$H(X^t) \le H(X_G X_B)$$
 Equation (4.4)

876

This is a provisional file, not the final typeset article

If $H(X_G, X_B)$ is the joint entropy, because f(x, y) = xy is a measurable function, we have Equation 877 878 4.5 [43]: 879 880 $H(X_G X_B) \leq H(X_G, X_B)$ Equation (4.5)881 882 The upper bound for the joint entropy is expressed in Equation 4.6, [43]: 883 884 $H(X_G, X_B) \leq H(X_G) + H(X_B)$ Equation (4.6)885 886 Equality holds true if and only if X_G and X_B are independent. Ultimately, from Equations (4.4) - (4.6): 887 888 $H(X^t) \leq H(X_c) + H(X_B)$ Equation (4.7)889 In practice, X_G is a discrete approximation of a continuous gamma variable X_g . If Δ is the size of the 890 bin used in the approximation and $h(X_g)$ the differential (continuous) entropy of X_g , it can be shown 891 892 that [43]: 893 $h(X_g) = \lim_{\Delta \to 0} (H(X_G, \Delta) + \log \Delta)$ 894 Equation (4.8) 895 896 Therefore, for small Δ we can write: 897 $h(X_{q}) \cong H(X_{G}, \Delta) + \log \Delta$ 898 Equation (4.9) 899 900 Equivalently: 901 $H(X_G) \cong h(X_a) - \log \Delta$ 902 Equation (4.10)903

904 905	We mentioned earlier that the differential entropy of the gamma distribution (with θ) has the closed form:	th shape k and scale			
906					
907	$h(X_g) = k + ln\theta + \ln\Gamma(k) + (1 - k)\psi(k)$	Equation (4.11)			
908					
909	where $\Gamma(k)$ is the gamma function and $\psi(k)$ is the digamma function.				
910					
911 912	Because of the power law between shape and scale that we discovered from anal data:	yzing human motion			
913					
914	$logk \cong alog\theta + b$	Equation (4.12)			
915					
916 917 918	Where $a \le 0$ and b are parameters that are determined by fitting a regression model on the population of interest. Because of this, differential entropy ends up being a univariate function of θ , for a specific set of parameters.				
919					
920 921 922 923 924	For the ranges of values typically found in the MMS, we see that Differential H positive linear relationship with the natural logarithm of the scale, which is equivaled Signal Ratio of the gamma process. For greater values, that are not usually encount the Differential entropy diverges. (Here the calculation of differential entropy for due to numerical instability). Figure 6A shows these results.	Entropy has almost a alent to the Noise-to- ntered in human data, smaller values failed			
925					
926 927 928 929	Because of Equation (4.10), we expect the empirical discrete entropy to also respect to the log-NSR, which turned out to be the case. For the range of phys found a slope of 0.4694 nats ($p<0.01$) for the differential entropy and a slope of for the empirical entropy.	behave linearly with siological values, we 0.2606 nats (p<0.01)			
930					
931 932	Moving on to the process $H(X_B)$, the theoretical entropy (if probabilities do not window, <i>i.e.</i> , spike process is Bernoulli) is equal to [44]:	change within a time			

935 Figure 6. Theoretical considerations of relationships between NSR and Entropy metrics. (A) 936 Near positive linear relationship between Differential Entropy and the natural logarithm of the 937 Gamma scale, the NSR of the gamma process. (B) The empirically estimated entropy is the same 938 as the equivalent Poisson process. Within the range of physiological data from the ADOS dyadic 939 interaction, the entropy of the gamma component increases with the NSR, and the entropy of the binary component increases with the Spike Rate. (C) Sampling the empirical ranges across the 940 941 entropy of the gamma component and the entropy of the binary component of the MMS yields 942 the theoretical maximal entropy defining the upper bound (green). Their ratio (blue) indicates 943 that the maximum entropy is greater than the empirical entropy by a quantity that is increasing 944 as the binary entropy drops (i.e., as the spike rate decreases) and it increases as the gamma entropy increases. (D) The binary spike process has a relatively small dependence from the 945 946 gamma process, suggesting that in human motion timing and amplitude (spatial) aspects of these 947 motions are independent.

948
$$H(X_B) = -p \log p - (1-p)\log(1-p)$$
Equation (4.13)950951Where p is the probability of a spike occurring. We approximate it in Equation 4.14 as:952953 $H(X_B) = -R \log R - (1-R)\log(1-R)$ Equation (4.14)954954954

955 Where *R* is the spike rate measured as "number of spikes" per "number of samples".

956

As we see, the empirically estimated entropy is the same as the equivalent Poisson process. As a conclusion, for the range of physiological data from the ADOS dyadic interaction, the entropy of the gamma component is increasing with the NSR, and the entropy of the binary component is increasing with the Spike Rate.

961

962 4.2 The Separability of Peak Activity from Standardized Angular Speed.

If we plot the entropy of the process $H(X_G X_B)$ and the maximum theoretical entropy defining the upper bound, $H(X_G) + H(X_B)$, we see that the maximum entropy is greater than the empirical entropy by a quantity that is increasing as the binary entropy drops, or equivalently (by the previous finding) as the spike rate decreases. It increases as the gamma entropy increases. Fitting a surface function, we find that:

968

969
$$H_{max}(X_G, X_B) - H(X_G X_B) = -2.368H(X_B) + 0.8141H(X_G) + 1.114, RMSE = 0.04857$$

970

971 Then, we get an approximate relation:

972

973
$$H(X_G X_B) = 3.368H(X_B) + 0.1859H(X_G) - 1.114$$

974

975 Recall that the following inequality holds:

976

977 $H(X_G X_B) \le H(X_G, X_B) \le H(X_G) + H(X_B)$

978

979 Which yields:

980

981
$$3.368H(X_B) + 0.1859H(X_G) - 1.114 \le H(X_G, X_B) \le H(X_G) + H(X_B)$$

982

984

983 Or:

985 $3.368[H(X_B) + H(X_G)] - 3.1821H(X_G) - 1.114 \le H(X_G, X_B) \le H(X_G) + H(X_B)$

986987Finally:988989
$$3.368H_{max}(X_G, X_B) - 3.1821(X_G) - 1.114 \le H(X_G, X_B) \le H_{max}(X_G, X_B)$$
990991If X_G and X_B were to be independent, we can see that:992993 $3.368H_{max}(X_G, X_B) - 3.1821H(X_G) - 1.114 = H_{max}(X_G, X_B)$ 994995Or:996997This independence criterion, is obviously data dependent. In the general case where we can fit a surface998 $H_{max}(X_G, X_B) = H(Y, Y_A) = cH(Y_A) + bH(Y_A) + c$

 $H_{max}(X_G, X_B) - H(X_G X_B) = aH(X_B) + bH(X_G) + c$

1000 The condition for independence is:

1001
$$H_{max}(X_G, X_B) = \left(1 - \frac{b}{a}\right) H(X_B) - \frac{c}{a}$$

1002 If the actual data can be well fit by a linear model with slope *A* and we ignore small differences between 1003 the intercepts of the two linear models, an easy way to quantify departure from independence is by 1004 computing the angle θ between the two lines:

1005
$$\theta = \tan^{-1} \frac{A - \frac{a-b}{a}}{1 + \frac{a-b}{a}A}$$

1006 Then, we define the degree of departure from independence as the ratio between the θ and $\frac{\pi}{2}$. If the 1007 data cannot be well fit by a linear model, it's best to perform a standard goodness-of-fit test and/or 1008 measure the mean error between the model and the data. If $\theta > 0$, the rate of maximum joint entropy 1009 increase over the gamma entropy is bigger than in the case of independence, if $\theta < 0$ it's smaller.

From our data, we see that this condition holds pretty well, with 9.93 % departure from independencefor the ASD group and 11.55 % for the NT group. This implies that:

1013
$$H(X_G, X_B) \cong H(X_G) + H(X_B)$$
 Equation (4.15)

1014

We just showed that the binary spike process is not very dependent from the gamma process. This supports the independence of time and space in human motion, previously proposed for intentional, goal-directed movements at a behavioral level of kinematic analyses [16; 45; 46; 47; 48] under a geometric modeling approach to address the brain control and coordination of the bodily degrees of freedom problem [49; 50].

1020

10214.3Controllability of an Agent in a Dyadic Social Interaction is Inversely Proportional to1022Autonomy: Leveraging Sociomotor Agency to Protect the Agent

1023

1024 In the methods section we defined what Transfer Entropy $T_{Y \to X}(k, l)$ between two processes X and Y 1025 is:

1026

1027 $T_{Y \to X}(k, l) = E[t_{Y \to X}(n+1, k, l)]$ Equation (4.16)

1028

1029
$$t_{Y \to X}(n+1,k,l) = i(\mathbf{y}_n^{(l)}; x_{n+1} | \mathbf{x}_n^{(k)})$$

1030

1031 Equivalently, TE can be seen as the difference between the conditional entropy rate (which is equal to 1032 entropy rate for stationary processes) h_X of process X and the generalized entropy rate $h_{X,Y}$ of X 1033 conditioning on the source Y [51]:

1034 $T_{Y \to X}(k, l) = h_X - h_{X,Y}$ Equation (4.17)

1035

1036 With:

1037
$$h_X = -\sum p(x_{n+1}, x_n^{(k)}) \log p(x_{n+1}|x_n^k)$$

- 1038
- 1039

1040
$$h_{X,Y} = -\sum p(x_{n+1}, x_n^{(k)}, y_n^{(l)}) \log p(x_{n+1}|x_n, y_n^{(l)})$$

1041

1042 The generalized entropy rate measures the uncertainty in predicting the future values of X, given its 1043 history and the past values of Y. Transfer Entropy is the reduction in uncertainty of predicting the

This is a provisional file, not the final typeset article

- 1044 future of X when we consider the process Y. If we call $h_{Y,X}$ uncertainty, then h_Y is what we already
- 1045 defined as autonomy and $T_{X \to Y}$ is the transfer entropy.
- 1046

We chose embedded history of length 20 for TE and for the entropy rate of our processes we used a template (embedding) length equal to the average distance between two spikes, to ensure that in the reconstructed space, the coordinates of a point in time include both zeros ("quiet moments") and spikes

- and that the system doesn't bounce back and forth from a single coordinate of zeros components. The
- 1051 embedding delay was chosen using Average Mutual Information.
- 1052
- 1053 If we plot the Child or Clinician Autonomy with respect to the log(NSR) and the Spike Rate, we see 1054 in Figure 7 that the relationship between entropy rate, noise and spike rate is rather complex. It also
- 1055 differs between NT and ASD, more data are needed to get a clear picture but we can definitely see that
- 1056 there is a small positive trend with respect to noise and spike rate. Nonetheless, this shows that the
- 1057 processes cannot be treated as *i.i.d.*
- 1058

1060Figure 7. Non i.i.d. process revealed by the relationship between autonomy, NSR and spike rate1061for clinician (A) and child (B) for the gamma and binary components of the MMS, relative to the

- 1062 process entropy.
- 1063
- 1064 Now that we have established the speed/peak activity independence and the positive correlation 1065 between entropy rate and NSR or Spike Rate, we are ready to study how TE behaves in the shared 1066 space of the child-clinician dyad.

1067

1068 Figure 8. Linear relationships between transfer entropy and the entropy rates (autonomies) for 1069 child and clinician differentiating between NT and ASD participants.

1070

1071 We find that $TE_{CL \to CH}$ decreases when the child exhibits high autonomy and increases when the 1072 clinician has higher autonomy and vice versa for $TE_{CH \to CL}$. In fact, this relationship is well 1073 characterized by linear relationships between transfer entropy and the entropy rates (autonomies), as 1074 the fitted linear surfaces indicate in Figure 8.

1075

1076 In this sense, we can safely conclude that by manipulating standardized human biorhythmic time series 1077 either by increasing the NSR or by increasing peak activity, we can increase autonomy and reduce the 1078 controllability of human agents by other human or by artificial agents, including those potentially 1079 created by AI.

1080

10814.4Validation of the Digitization of the ADOS: Automated, Streamlined and Scalable Screener1082of Socio-Motor Agency

To make our basic scientific results actionable, we need to validate our digital data with the clinical criteria, a paradigm that we have coined *clinically interpretable digital biomarkers*. In this model, the objective digital indexes that we used to define socio-motor agency as the autonomy-to-control ratio, are examined in relation to the ADOS clinical scores that a trained human rated during the session. We employ a machine learning technique, Support Vector Machine (SVM) to classify the digital data as a function of the clinical score. Then we apply tools from signal detection theory, specifically the receiving operating characteristic curve, ROC, to assess the validity of our classifier.

1090

1091 Each of 26 participants with the full ADOS session (digital and clinical) produces on average between 1092 50-60 minutes of time series digital data from biosensors registering motion at 128Hz. We used the 1093 left-hand wrist sensor in these analyses, as we showed that it is highly correlated with the right wrist,

1094 yet more variable, thus expanding our sampling space. Upon exploration of several time windows to 1095 segment the data, sweeping across the time series and tasks, while maximizing statistical power in each

1096 locally stationary segment, we arrived at 7.8 second windows as optimal.

1097

1098 The data were validated using the Leave-one-person-out cross-validation (LOOCV) method. As 1099 features for our classifier, we used autonomy (entropy rate), NSR and the embedding delay of the data, 1100 which is the time scale at which deterministic properties arise and characterize the dynamical behavior 1101 of motion (for more information, see Methods). Two classifiers were used, one trained on female 1102 subjects and the second one trained exclusively on male participants. When trying to digitally diagnose 1103 autism in one participant, we trained our classifier on the data from the remaining male or female 1104 participants and then tested how accurately the trained model predicts the participant class (NT vs. 1105 ASD). This method avoids overfitting and trains models that can diagnose autism in novel participants, 1106 thus automating the screening process. Digitizing the ADOS in this way makes the diagnosis of autism 1107 more inclusive of females, historically underdiagnosed by a test that we objectively showed has biases 1108 towards males across all tasks [10]. A larger sample size and a longitudinal study are required to 1109 validate our model at scale. Yet, as shown in Figure 9A, there is no confusion of our biometrics about 1110 the clinician ADOS scores, which classify ASD males with 100% accuracy and performs remarkably 1111 well for ASD vs. NT females. Indeed, Figure 9B confirms the validity of these biometrics for clinical 1112 use with an area under the ROC curve of 95.76%.

1113

1114 Figure 9. (A) Support Vector Machines (SVM) classifiers were trained on all subjects except one 1115 and tested on the remaining subjects of the same sex (Leave-one-person-out cross-validation 1116 (LOOCV)). Therefore, each of the 26 subjects was digitally diagnosed with a classifier trained 1117 on a different dataset, which ensured zero overfitting and bias. Training and testing features 1118 were the entropy rate, the signal-to-noise ratio, and the embedding delay (the time scale at which 1119 a dynamical system behaves in the most deterministic way) calculated on normalized speed 1120 samples of \sim 7.8 secs duration windows. Here, we report the percentage of time windows per subject that gave a positive diagnostic label and plot them versus the ADOS scores, as determined 1121 1122 by the clinicians. (B) We use the positive rate scores as a metric used to diagnose ASD and report 1123 the Receiver Operating Characteristic curve (ROC curve), which shows the true positive and

false positive rates of the digital diagnostic tool we developed for different thresholds. The Area Under the Curve (AUC) is 0.9576, which indicates great performance.

1126

1127 **5 Discussion**

1128 In this work, we use the ADOS test as a backdrop to study social interactions between children and 1129 adult clinicians with the purpose of defining new ways to automate and speed up the autism screening 1130 process, while leveraging the clinical validity of this test. To that end, we explored anew the concept 1131 of socio-motor agency by defining a ratio of two indexes of autonomy and control. Autonomy was 1132 defined as the non-parametric entropy rate spanning from totally random to totally deterministic 1133 behavior of standardized micro-movements spike trains. These were derived from nuanced fluctuations 1134 in motion data that contains goal-directed segments of behavior interspersed with spontaneously 1135 occurring, more ambiguous, transient segments that are known to interconnect the goal-directed ones 1136 [16; 52]. Control was defined in terms of the NSR, empirically estimated from such spike trains as 1137 well, such that high regimes of NSR correspond to the memoryless regimes denoting high uncertainty 1138 (poor predictability and randomness) in the motor code. We took a step further to examine the 1139 parameterization of the MMS as a binary-spike and a Gamma process and demonstrated the 1140 independence between them.

1141

1142 We reasoned that these binarized sequences of spikes bear a motor code whereby the observer may or 1143 may not be able to predict and therefore control the observed agent. At high randomness, the observed 1144 agent affords more autonomy than at deterministic ranges. At deterministic ranges, with high 1145 regularity, the observer can predict and control the actions of the observed agent. At higher NSR, the 1146 agent has lower self-control. This is so because the kinesthetic reafferent feedback from the motions is 1147 noisy and with such poor signal quality it is difficult to predict a desired outcome and plan the action 1148 consequences to compensate for sensory transduction, transmission, and motor integration delays 1149 inherent in the person's system. As predicting his/her/their motor actions consequences can then be 1150 compromised by noise in the motor code, the child is more controllable by the clinician. The observer 1151 clinician can exert higher control over the observed agent. In this sense, the child's socio-motor agency 1152 may also be compromised. This is the case whether the child / adult is autistic.

1153

Underlying both indexes and the ratio of autonomy to control are then discrete pockets of information making up a continuous stream of dyadic motor code, contributed by both social agents. Thus, we can infer the existence of an underlying shared alphabet in the motor code that manifests during dyadic social interactions of the type studied here. Agents with discrete motor signatures that appear more random are thus harder to control and behave more autonomously and independently than agents with systematically predictable motions sharing their codes.

1160

Besides describing new biometrics of shared socio-motor agency in dyadic social interactions, our analyses showed ways to streamline the ADOS test, thus making it less taxing on the child and the clinician. A handful of tasks affording more socio-motor agency to the child can indeed uncover the social readiness potential of the child rather than biasing the diagnosis by the clinician towards a deficit

1165 model. Along those lines, using these newly defined indexes of dyadic autonomy and control, we 1166 demonstrated fundamental differences across the tasks for males and females, thus confirming that 1167 despite previously quantified differences in motor control separating males and females at the 1168 voluntary [4] and involuntary [12; 13] levels, the ADOS remains biased towards males. These digital 1169 indexes of shared socio-motor agency, nevertheless, used within the context of an unbiased ML 1170 classifier, could detect the differences between males and females for both the NT and ASD randomly 1171 chosen participants. This digitized automated version of the test resembles the type of scenario that a 1172 clinician faces at the clinic, any given day. Namely, a random arrival of a case that the clinician may 1173 see for the first time. In that sense, the leave-one-person-out classifier provides robust digital screening 1174 of autism and may be a way to scale our pilot study to encompass larger numbers of NT, ASD 1175 participants across ages, sexes, and do so longitudinally as well.

1176

1177 Future longitudinal studies of autism with an eye for the evolution of the neuromotor code and its 1178 impact on social perception and cognition, will require the type of normalization that we introduced 1179 earlier with the MMS [3] and further used here, namely, scaling out allometric effects due to anatomical 1180 differences across participants (see also [17; 18; 53; 54]). This step is crucial in any study that involves 1181 biorhythmic motions whereby kinematic analyses will be impacted by such anatomical differences. 1182 This is so because kinematic parameters such as speed, acceleration, distance, etc. are impacted by the 1183 limb sizes and masses in ways that confound results and interpretation of such studies [55]. It will be 1184 particularly important to consider these caveats present in all current studies that do not account for 1185 allometric differences during the very early neurodevelopment when the rate of bodily growth is highly 1186 non-linear and accelerated [54]. These rates of changes in anatomical growth produce different ranges 1187 of values in such kinematic parameters and impact the empirical distributions of the values associated 1188 with natural behaviors such as those examined here.

1189 5.1 Implications of Socio-Motor Agency Metrics for AI and Privacy Protection

1190 The theoretical considerations at the intersection of stochastic analyses and information theoretic 1191 approaches with non-linear dynamics offers the MMS and analyses as a viable way to obtain the 1192 personalized signatures of autonomy and control and tweak the NSR to mask the spike trains derived 1193 from the person's physiological biorhythmic activity. This ability to separate the binary spike rate code 1194 from the gamma process denoting levels of randomness vs. predictability, offers the possibility of 1195 creating a device that alerts the persons involved in the dyadic exchange to balance their autonomy and 1196 control, to attain socio-motor agency. By enhancing autonomy and avoiding excessive external control 1197 by the other agent, be that agent another human or an AI-driven one, the person can be protected from 1198 excess control. This approach will be critical to revamp autism therapies with an emphasis to respect 1199 the child's autonomy and support the bottom-up development of autonomous motor control. The 1200 maturation of bottom-up autonomous motor control (building blocks of autonomy) is a necessary pre-1201 requisite for the further neurodevelopment of top-down control. Without considering and balancing the 1202 orderly maturation rates of these two building blocks of socio-motor behavior, therapies in autism will 1203 cause trauma to the nervous system.

We propose that this methodology can also be used to protect our privacy more generally from surveillance systems, as ultimately these systems rely on biometric data, which we can now, using the present personalized methods, manipulate to hide our fingerprint-like signatures from an external agent trying to control us. This solution to the controllability issue can then be extended from individuals to dyads, from dyads to social groups and from social groups to society. In this sense, socio-motor agency

1209 can serve as a foundation for societal agency, now quantifiable using the methods that we offer in this1210 work.

1211

1212 6 Conclusions

1213 In summary, we found that variability in the dyadic index of autonomy is more pronounced in ASD 1214 than in NTs, across a broad range of ages from 4-15 years old. Furthermore, we found that the dyadic NSR, indicative of socio-motor control, increases with age. This result is consistent with prior work on 1215 1216 individuals across ages and sex [3; 4]. In contrast, both ASD and NT showed increases of the autonomy 1217 index with age, an indicator that regardless of the human condition, whether developing along a 1218 neurotypical trajectory, or along the trajectory of autism spectrum disorders, respecting the child's 1219 autonomy will be necessarily our best ally when designing future treatments that unveil the child social 1220 readiness potential. We would not have known this had we treated the ADOS as the criterion test that 1221 it is (*i.e.*, based off children with neurodevelopmental issues only), rather than treating it as a normative 1222 test (*i.e.*, including NT controls as well, to define normative ranges and quantify similarities and 1223 departures from it.)

1224 We have uncovered new indexes of shared, dyadic autonomy and control, objectively defined socio-

motor agency and provided new means to automate its digital screening with already routinely used

1226 clinical tools. This works offers novel ways to scale our clinical science and make it actionable, diverse,

1227 and inclusive at more than one level.

1228 7 Conflict of Interest

1229 The authors declare that the research was conducted in the absence of any commercial or financial 1230 relationships that could be construed as a potential conflict of interest.

1231 8 Author Contributions

TB and EBT contributed to conception and design of the analyses. TB analysed the data while EBT
designed the clinical study. RR organized and curated the database. TB wrote methods and derivations.
EBT wrote report. TB and EBR wrote and edited full manuscript. All authors contributed to manuscript
revision, read, and approved the submitted version.

1236 **9** Funding

This work was supported by the Nancy Lurie Marks Family Foundation Career Development Award
to EBT and by the New Jersey Governor's Council for Autism to EBT. TB and RR were funded by the
NJ GCA grant.

1240 10 Acknowledgments and Disclaimer

We thank the children and families who kindly participated in the study. This study was approved by the Rutgers University IRB and signed consent was obtained from the legal guardian/parent of the child.

- 1243
- 1244

1245 11 References

- [1] C. Lord, S. Risi, L. Lambrecht, E.H. Cook, Jr., B.L. Leventhal, P.C. DiLavore, A. Pickles, and M.
 Rutter, The autism diagnostic observation schedule-generic: a standard measure of social and
 communication deficits associated with the spectrum of autism. J Autism Dev Disord 30 (2000)
 205-23.
- [2] K. Gotham, S. Risi, A. Pickles, and C. Lord, The Autism Diagnostic Observation Schedule: revised
 algorithms for improved diagnostic validity. J Autism Dev Disord 37 (2007) 613-27.
- [3] E.B. Torres, M. Brincker, R.W. Isenhower, P. Yanovich, K.A. Stigler, J.I. Nurnberger, D.N.
 Metaxas, and J.V. Jose, Autism: the micro-movement perspective. Front Integr Neurosci 7 (2013) 32.
- [4] E.B. Torres, R.W. Isenhower, P. Yanovich, G. Rehrig, K. Stigler, J. Nurnberger, and J.V. Jose,
 Strategies to develop putative biomarkers to characterize the female phenotype with autism
 spectrum disorders. J Neurophysiol 110 (2013) 1646-62.
- [5] A.M. D'Mello, I.R. Frosch, C.E. Li, A.L. Cardinaux, and J.D.E. Gabrieli, Exclusion of females in autism research: Empirical evidence for a "leaky" recruitment-to-research pipeline. Autism Res (2022).
- [6] S. Lundstrom, C. Marland, R. Kuja-Halkola, H. Anckarsater, P. Lichtenstein, C. Gillberg, and T.
 Nilsson, Assessing autism in females: The importance of a sex-specific comparison. Psychiatry
 Res 282 (2019) 112566.
- [7] R. Loomes, L. Hull, and W.P.L. Mandy, What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. J Am Acad Child Adolesc Psychiatry 56 (2017) 466-474.
- [8] E. Somoza, and D. Mossman, ROC curves and the binormal assumption. J Neuropsychiatry Clin
 Neurosci 3 (1991) 436-9.
- 1269 [9] M. Hollander, and E.A. Pena, Nonparametric Methods in Reliability. Stat Sci 19 (2004) 644-651.
- [10] E.B. Torres, R. Rai, S. Mistry, and B. Gupta, Hidden Aspects of the Research ADOS Are Bound to Affect Autism Science. Neural Comput 32 (2020) 515-561.
- [11] H. Bokadia, R. Rai, and E.B. Torres, Digitized Autism Observation Diagnostic Schedule: Social
 Interactions beyond the Limits of the Naked Eye. J Pers Med 10 (2020).
- [12] E.B. Torres, and K. Denisova, Motor noise is rich signal in autism research and pharmacological
 treatments. Sci Rep 6 (2016) 37422.
- [13] E.B. Torres, S. Mistry, C. Caballero, and C.P. Whyatt, Stochastic Signatures of Involuntary Head
 Micro-movements Can Be Used to Classify Females of ABIDE into Different Subtypes of
 Neurodevelopmental Disorders. Front Integr Neurosci 11 (2017) 10.
- [14] M. Brincker, and E.B. Torres, Noise from the periphery in autism. Front Integr Neurosci 7 (2013)
 34.
- [15] E.B. Torres, H. Varkey, J. Vero, E. London, H. Phan, P. Kittler, A. Gordon, R.E. Delgado, C.F.
 Delgado, and E.A. Simpson, Sensing echoes: temporal misalignment in auditory brainstem
 responses as the earliest marker of neurodevelopmental derailment. PNAS nexus 2 (2023)
 pgac315.
- 1285 [16] E.B. Torres, Two classes of movements in motor control. Exp Brain Res 215 (2011) 269-83.
- [17] E.B. Torres, R.W. Isenhower, J. Nguyen, C. Whyatt, J.I. Nurnberger, J.V. Jose, S.M. Silverstein,
 T.V. Papathomas, J. Sage, and J. Cole, Toward Precision Psychiatry: Statistical Platform for
 the Personalized Characterization of Natural Behaviors. Front Neurol 7 (2016) 8.
- [18] E.B. Torres, C. Caballero, and S. Mistry, Aging with Autism Departs Greatly from Typical Aging.
 Sensors (Basel) 20 (2020).

- [19] C.E. Shannon, and W. Weaver, The mathematical theory of communication, University of Illinois
 Press, Urbana, 1998.
- 1293 [20] C.E. Shannon, Claude Elwood Shannon papers, 1932-1995 (bulk 1938-1989), pp. 7,000.
- [21] C.E. Shannon, N.J.A. Sloane, and A.D. Wyner, Claude Elwood Shannon : miscellaneous writings,
 Mathematical Sciences Research Center, AT&T Bell Laboratories, Murray Hill, N.J., 1993.
- [22] C.E. Shannon, N.J.A. Sloane, A.D. Wyner, and IEEE Information Theory Society., Claude
 Elwood Shannon : collected papers, IEEE Press, New York, 1993.
- [23] J. Shi, L. Chen, and K. Aihara, Embedding entropy: a nonlinear measure of dynamical causality.
 J R Soc Interface 19 (2022) 20210766.
- [24] E. Tan, S. Algar, D. Correa, M. Small, T. Stemler, and D. Walker, Selecting embedding delays:
 An overview of embedding techniques and a new method using persistent homology. Chaos 33
 (2023) 032101.
- [25] F. Takens, Detecting strange attractors in turbulence. in: D.A.R.a.L.-S. Young, (Ed.), Dynamical
 Systems and Turbulence, Lecture Notes in Mathematics, Springer-Verlag, 1981, pp. 366–381.
- 1305 [26] A. Delgado-Bonal, and A. Marshak, Approximate Entropy and Sample Entropy: A
 1306 Comprehensive Tutorial. Entropy (Basel) 21 (2019).
- [27] J.T. Lizier, JIDT: an information-theoretic toolkit for studying the dynamics of complex systems.
 Frontiers in Robotics and AI 1 (2014) 1-20.
- [28] N.H. Packard, J.P. Crutchfield, J.D. Farmer, and R.S. Shaw, Geometry from a time series. Physical
 Review Letters 45 (1980) 712-716.
- [29] A. Delgado-Bonal, and A. Marshak, Approximate entropy and sample entropy: A comprehensive tutorial. Entropy 21 (2019) 541.
- [30] P. Grassberger, and I. Procaccia, Estimation of the Kolmogorov entropy from a chaotic signal.
 Physical review A 28 (1983) 2591.
- [31] J.-P. Eckmann, and D. Ruelle, Ergodic theory of chaos and strange attractors. Reviews of modern
 physics 57 (1985) 617.
- [32] C. Rhodes, and M. Morari, The false nearest neighbors algorithm: An overview. Computers &
 Chemical Engineering 21 (1997) S1149-S1154.
- [33] J.T. Lizier, JIDT: An information-theoretic toolkit for studying the dynamics of complex systems.
 Frontiers in Robotics and AI 1 (2014) 11.
- [34] R.M. Fano, and D. Hawkins, Transmission of information: A statistical theory of communications.
 American Journal of Physics 29 (1961) 793-794.
- [35] J.T. Lizier, M. Prokopenko, and A.Y. Zomaya, Local information transfer as a spatiotemporal
 filter for complex systems. Physical Review E 77 (2008) 026110.
- [36] M. Wibral, R. Vicente, and J.T. Lizier, Directed information measures in neuroscience, Springer,
 2014.
- [37] S.M. Pincus, and A.L. Goldberger, Physiological time-series analysis: what does regularity
 quantify? American Journal of Physiology-Heart and Circulatory Physiology 266 (1994)
 H1643-H1656.
- [38] A.V. Lazo, and P. Rathie, On the entropy of continuous probability distributions (corresp.). IEEE
 Transactions on Information Theory 24 (1978) 120-122.
- [39] E.B. Torres, Atypical signatures of motor variability found in an individual with ASD. Neurocase
 19 (2013) 150-65.
- [40] E.B. Torres, Atypical signatures of motor variability found in an individual with ASD. Neurocase
 1335 19 (2011) 150-65.
- [41] M. Brincker, and E.B. Torres, Chapter 1- Why Study Movement Variability in Autism_. in: E.B.
 Torres, and C. Whyatt, (Eds.), Autism : the movement sensing perspective, CRC Press/Taylor
 & Francis Group, Boca Raton, 2018, pp. xviii, 386 pages.

- [42] S.N. Mohamed Thangal, and J.M. Donelan, Scaling of inertial delays in terrestrial mammals. PLoS
 One 15 (2020) e0217188.
- 1341 [43] T.M. Cover, Elements of information theory, John Wiley & Sons, 1999.
- [44] A. Klenke, and A. Klenke, The Poisson point process. Probability Theory: A Comprehensive
 Course (2014) 543-561.
- 1344 [45] E.B. Torres, New symmetry of intended curved reaches. Behav Brain Funct 6 (2010) 21.
- [46] E. Torres, and R. Andersen, Space-time separation during obstacle-avoidance learning in monkeys. J Neurophysiol 96 (2006) 2613-32.
- [47] E.B. Torres, Zipser D., Simultaneous control of hand displacements and rotations in orientation matching experiments. Journal of Applied Physiology 96 (2004) 1978-1987.
- [48] E.B. Torres, R. Quian Quiroga, H. Cui, and C.A. Buneo, Neural correlates of learning and trajectory planning in the posterior parietal cortex. Front Integr Neurosci 7 (2013) 39.
- [49] E.B. Torres, Zipser, D., Reaching to Grasp with a Multi-jointed Arm (I): A Computational Model.
 Journal of Neurophysiology 88 (2002) 1-13.
- [50] E.B. Torres, Theoretical Framework for the Study of Sensori-motor Integration., Cognitive
 Science, University of California, San Diego, La Jolla, 2001, pp. 115.
- [51] M. Prokopenko, J.T. Lizier, and D.C. Price, On thermodynamic interpretation of transfer entropy.
 Entropy 15 (2013) 524-543.
- [52] E.B. Torres, Signatures of movement variability anticipate hand speed according to levels of
 intent. Behav Brain Funct 9 (2013) 10.
- [53] C. Caballero, S. Mistry, and E.B. Torres, Age-Dependent Statistical Changes of Involuntary Head
 Motion Signatures Across Autism and Controls of the ABIDE Repository. Front. Integr.
 Neurosci. 14 (2020) 1-14.
- [54] E.B. Torres, B. Smith, S. Mistry, M. Brincker, and C. Whyatt, Neonatal Diagnostics: Toward
 Dynamic Growth Charts of Neuromotor Control. Front Pediatr 4 (2016) 121.
- [55] E.B. Torres, J. Vero, and R. Rai, Statistical Platform for Individualized Behavioral Analyses Using
 Biophysical Micro-Movement Spikes. Sensors (Basel) 18 (2018).
- 1366

1367 12 Data Availability Statement

- The datasets [GENERATED/ANALYZED] for this study can be found in the repository
 <u>https://zenodo.org/records/10032169</u>. Please see the "Availability of data" section of <u>Materials and</u>
 data policies in the Author guidelines for more details.
-
- 1371
- 1372
- 1373