



Department of Psychology

Theoretical Neuroscience Group

Jun. Prof. Dr. Dominik Endres

“Dance Like Nobody Is Watching You”

The production of movement primitives in unconstrained dancing

In Partial Fulfilment of the Requirements for the Degree
Bachelor of science, Psychology

Submitted by

Anna-Lena Eckert

First supervisor: Jun. Prof. Dr. Dominik Endres

Second supervisor: Prof. Dr. Anna Schubö

Marburg, September 2017

I would like to thank Dr. Dominik Endres for sharing his knowledge and his passion for science as well as for providing many hours of theoretical and technical support.

Financial funding was received from SGB-TER135 ‘Kardinal Mechanisms of perception’ project C6, for the compensation of participants.

Special thanks to Benjamin Knopp for sharing his knowledge and expertise. Also, thank you to all members of the Theoretical Neuroscience workgroup for providing the lab and the great working atmosphere, and special thanks to my 17 participants who rocked the lab.

Thank you to Dr. Anna Schubö and Dr. Dominik Endres for supervision and review of this thesis.

Special thanks to my parents Heike and Philipp Eckert for their support and to Dominik Deffner, who had to put up with a lot but has always been a huge help.

Table of Contents

LIST OF FIGURES.....	II
LIST OF TABLES.....	II
ABSTRACT.....	IV
1. INTRODUCTION AND THEORETICAL BACKGROUND	
1.1. Motor Behavior.....	1
1.2. Challenges in motor control research.....	1
1.3. Movement Primitives.....	2
1.4. Movement Primitives and the motor hierarchy.....	3
1.5. The extraction of MP and different MP models.....	4
1.6. Experimental settings for the investigation of MP.....	8
1.7. Dance.....	9
1.8. Rationale & Hypotheses.....	9
2. METHODS	
2.1. Stimuli.....	11
2.2. Participants.....	11
2.3. Recording of movement data.....	11
2.4. Analysis	
2.4.1. Pre-processing of movement data.....	12
2.4.2. Data Analysis: model selection	13
3. RESULTS	
3.1. Questionnaire results.....	19
3.2. Results Model 1 ('Universal set of MP').....	19
3.3. Results Model 2 ('Individual repertoire').....	20
3.4. Results Model 3 ('Stimulus-determined behavior').....	20
3.5. Results Model 4 ('Interaction individual * stimulus').....	21
3.6. Results Model comparison.....	23
4. DISCUSSION	
4.1. Summary and interpretation of results.....	25
4.2. Limitations.....	26
4.3. Future research and applications	28
5. CONCLUSION.....	29
REFERENCES.....	30
APPENDIX.....	35

APPENDIX

APP 1: Participant study information	35
APP 2: Informed consent form.....	37
APP 3: Post-experimental questionnaire	39

LIST OF FIGURES

Figure 1. Model Evidence and Variance accounted for (VAF): ‘universal set of MP’	19
Figure 2. Model evidence and VAF ‘participant-determined model’	21
Figure 3. Model evidence and VAF ‘stimulus-determined model’	20
Figure 4. Model evidence and VAF ‘participant and stimulus determined model’	22
Figure 5. Comparison of Lay vs. expert dancer.....	24
Figure 6. Illustration of the beat-tracking system.....	44
Figure 7. Data arrays underlying different MP conceptualizations	45
Figure 8. AIC and BIC ‘universal set of MP’	46
Figure 9. AIC and BIC ‘individual repertoire model’	47
Figure 10. More examples for LAP in the stimulus-determined model	48
Figure 11. AIC and BIC scores ‘stimulus-determined model’: examples.....	49
Figure 12. Model comparison: Maximum LAP scores per model.....	50

LIST OF TABLES

Table 1. Model comparison: computation and results.....	23
Table 2. Results post-questionnaire.....	41
Table 3. Full list of stimuli.....	42
Table 4. Overview: current MP definitions.....	43

ABSTRACT

According to a long-established hypothesis in motor control, the brain might simplify the tremendous computational demands of motor coordination by relying on a discrete number of simpler elements, or movement primitives (MP). However, the extraction of MP from unconstrained movement data is only rarely encountered in the literature. For the present work, the movements of 17 human participants engaged in unconstrained, spontaneous dance were recorded. The fit of the Synchronous Model for MP and its respective model order was evaluated in a model selection process. MP were learned from pre-processed data and the probability of a given model was approximated with the Laplace Criterion for model selection, the Bayesian Information Criterion and the Akaike Information Criterion. Analysis revealed that generally, complex models with at least 34 synchronous components form a good basis for the decomposition of the movements. Further, model comparisons imply that dancing is a rather stimulus driven behavior which emerges from an individual x stimulus interaction. Further research regarding other models for MP proposed in the literature is necessary. Opportunities for applications in perception research, computer graphics and robotics are discussed.

1. INTRODUCTION AND THEORETICAL BACKGROUND

“Movements are complex objects in space and time that adapt their structure to solve an infinite number of motor problems”. Bernstein (1967)

1.1. Motor Behavior

Animals and, in particular, humans, are capable of flexibly adjusting their overt behavior to a multitude of contexts, conditions, tasks and behavioral goals. Understanding human motor behavior, defined as any behavior that contributes to the position or change of position of body parts by exerting muscular forces (Tresilian, 2012), is a fascinating, yet challenging endeavor as biological movements tend to be of complex and continuous nature and unfold dynamically in space and time (Clever et al., 2016; Danion & Latash, 2011; Rosenbaum, 2009). Nonetheless, humans usually seem to employ little effort in order to control, adapt and execute the movements of their bodies; to the contrary, they are capable of skillful, exquisite full-body movements, for example when dancing, doing gymnastics or martial arts.

1.2 Challenges in motor control research

Motor coordination is the process by which the many degrees of freedom (df) of the body are assembled together to produce purposeful, integrated movements (Bernstein, 1967). Df are defined as “the number of dimensions in which the physical system can independently vary” (Rosenbaum, 2009). A prevalent problem in the motor control literature is referred to as the “Degrees-of-Freedom-Problem” (Tresch et al., 2002; Bernstein, 1967; Rosenbaum, 2009). Considering the process of selecting an appropriate motor behavior while being confronted with redundancy on the musculoskeletal and neural levels illustrate this concern. Importantly, the excess df have many advantages as they allow the central nervous system (CNS) to flexibly adapt movements to a rapidly changing environment (Danion & Latash, 2011). Yet, it remains puzzling how the brain exploits this over-equipped physical system and efficiently accomplishes the fast computation and execution of specific motor options.

In addition to the computational burden the CNS is faced with when selecting an adequate motor behavior for a specific situation, its performance relies on relatively slow circuitry (Ballard, 2015). Compared to a silicon computer, axonal activation transmission rates with a maximum of 10 Hz reach only a fraction of a computer’s transmission rates (~40 kHz). From research with robotic systems like the SARCOS humanoid robot, it becomes

evident that centrally and explicitly controlling every df of a biological body would be more than unrealistic due to its enormous complexity (Ballard, 2015). The noisy environments humans move in are generally prone to disturbances from a large number of sources, which further expand the computational demands for an agent.

To this date, these theoretical problems are prevalent in the motor control literature (Ting & Macpherson, 2005; Rosenbaum, 2009). One of the core components of the theoretical considerations described above is the observation that the CNS has the capacity to quickly compute accurate and efficient motor actions while faced with tremendous computational burdens. It seems intuitive that the CNS relies on some sort of simplification approach in order to master the control of the many df of the physical body (Danion & Latash, 2011, Ballard, 2015). Suggestions on how motor coordination can be simplified are various and include the idea of ‘intelligent mechanics’, i.e., characteristic biomechanical and muscular constraints of a biological system that reduce the amount of computation necessary for movement and locomotion (van Ingen Schenau, 1989); elimination strategy, i.e. the ‘freezing’ of a surplus of df and the focusing of computational capacities on task-relevant df (Vereijken et al., 1992); or the ‘dynamical systems theory’ approach in which computation is replaced by the spontaneous evolution of spatiotemporal movement patterns over space-time (Schöner & Kelso, 1988). Finally, the concept of Movement Primitives has gained wide consideration (Flash & Hochner, 2005; Bizzi et al., 2008; Danion & Latash, 2011).

1.3. Movement Primitives

Movement Primitives (MP) constitute another facilitating strategy the CNS could rely on when planning and executing movements. The idea of MP is based on the assumption that complex movements are generally comprised of individual sequences or elementary ‘building blocks’ (Bizzi et al., 2008, Flash & Hochner, 2005). They are assumed to consist of simpler movement variables, such as joint trajectories (e.g., Kaminski, 2007; Flash & Hogan, 1985) or electrophysiological muscle activation (d’Avella, Saltiel & Bizzi, 2003). Hence, the combinations of a limited number of stored, simpler movement segments in space and time and the operations and transformations performed on them can produce the flexible and elaborate motor behavior casually performed by humans and other animals. MP therefore potentially constitute the link between an implicit internal representation of a movement and the observable motor output (Danion & Latash, 2011). In the literature, MP might also be referred to as ‘motor schemas’ (Arbib, 1992), ‘prototypes’ (Jeannerod et al., 1995), or ‘control modules’ (Schaal, Ijspeert & Billard, 2003). The configuration and

recombination of simpler elements in order to produce more complex behavioral patterns seems to be crucial not only for movement control, but also for other cognitive abilities, like language production and speech (Pinker, 2015; Moro, 2014). This notion has led DelVecchio, Murray and Perona (2003) to coin the term ‘movemes’ which accentuates the functional analogy of ‘movemes’ for movement-, and phonemes for language-production. Similar to phonemes, movemes are assumed to be reconfigured and combined obeying a pre-defined syntax (Flash & Hochner, 2005).

1.4.Movement Primitives on different levels of the motor hierarchy

A multitude of empirical findings support the existence of MP, and thus modularity, at different levels of the motor hierarchy (Ballard, 2015; Flash & Hochner, 2005).

At the *behavioral* level, the statistical analysis of motion data indeed suggests that complex action streams might be composed of discrete sub-movements. Reaching movements, for example, might consist of the combination of several straight trajectories, where each trajectory is characterized by specific velocity profiles (Flash & Hogan, 1985; Harris & Wolpert, 1998). A more commonplace example for this assumption is a dance step scheme commonly used for teaching choreographed dances. Here, individual steps that eventually are concatenated by the dance student form the building blocks of a more complex series of actions (Endres, Chiovetto & Giese, 2016).

On the *muscular* level, the assumption of muscle synergies is a recurring theme in motor control and -coordination research (Lee, 1984, Macpherson, 1991, d’Avella, Saltiel & Bizzi, 2003; Cheung, d’Avella, Tresch & Bizzi, 2005, Overduin et al., 2015). Historically, the idea of muscle synergies dates back to the early 20th century, where research on reflexes led to the concept of synergistically coordinated activation patterns (Sherrington, 1910). The term ‘muscle synergy’ refers to a functional unit that coordinates the activity of several muscles (d’Avella & Tresch, 2002). The activation of a muscle synergy produces muscular activation, and therefore force in a particular direction (Flash & Hochner, 2005). D’Avella, Saltiel and Bizzi (2003) examined muscular activation during a kicking motion in the frog’s 13 hindlimb muscles using EMG recordings. According to the authors, three discrete elements (or synergies) with varying amplitudes account for the complex behavioral pattern; suggesting an efficient means of simplifying motor control for the observable behavior.

On the *neural* level, research using microstimulation techniques has come to fascinating conclusions. For instance, Graziano, Taylor and Moore (2002) observed complex and

purposeful movements in monkeys after electrically stimulating primary motor and premotor cortical areas that resembled behaviors from the monkeys' natural motor repertoire. These findings suggest that on a cortical level, modularity is achieved by relying on a cortical map of joint angles. This would equate a MP to a 'motor program' that unfolds more or less independently of external or internal feedback signals (d'Avella, Saltiel & Bizzi, 2003). Importantly, evidence suggests that neurons in the spinal cord might be central for the organization of MP. Hart and Giszter (2010) used multielectrode neural recordings during a frog's kick movements and found that the activity of sets of interneurons is associated with specific simple movement segments; emphasizing a central role of spinal interneurons in the construction of MP. In the invertebrate literature, these populations of interneurons are also referred to as 'modules' and it is suggested that both behavior-independent and behavior-specific modules are necessary for the construction of complex motor behaviors (Jing et al., 2004). Furthermore, the notion of 'central pattern generators (CPGs)' has received considerable attention. A CPG is a neural network in the spinal cord that autonomously, i.e. without inputs from supraspinal or peripheral sources, produces several rhythmic movements such as breathing, locomotion or scratching (Marder & Bucher, 2001, Frigon, 2012). Frigon (2012) showed that even spinalized animals are capable of locomotor activity without relying on descending signals from cortical areas. Also, the combination of the activity of different CPG is thought to produce different complex movement patterns. The author concluded that the same muscles might be involved in different complex movements, but the differential modulation by a CPG is necessary to produce the specific behavioral pattern.

In sum, these findings support the concept of motor primitives on different organizational levels of the motor hierarchy, suggesting compositionality in combination with a hierarchical organization as central simplification strategies the motor system possibly relies on when planning and executing coordinated motor behavior.

1.5.The extraction of MP and models for MP

Bernstein's (1967) early ideas on motor planning and –coordination already contained aspects of movement compositionality. Yet, due to more elaborate technological opportunities for movement recording and its statistical analysis as well as appealing application areas, the research interest in complex behaviors and motor modularity has been growing within the past years (Mussa-Ivaldi, 1999). The extraction of MP is commonly performed using blind source separation (BSS) methods (i.e., the extraction of unknown 'sources', or components, from a set of given signals, Omlor & Giese, 2007) that exploit

features of the temporal, spatial or spatio-temporal structure underlying the kinematic or electrophysiological data. Common examples are principal component analysis (PCA), independent component analysis (ICA) or non-negative matrix factorization (NMF; d'Avella, Saltiel & Bizzi, 2003; see Ivanenko et al., 2005, for a comparison of PCA, ICA and NMF). BSS aim to reduce the dimensionality of kinematic or electrophysiological data by superposing a minimal, yet sufficiently representative number of sources, or MP, on the data stream (Omlor & Giese, 2007). More elaborate techniques also include the estimation of temporal delays (Omlor & Giese, 2007; 2007a; see Jenkins & Mataric, 2003, for a review).

The wide variety of methodological procedures and theoretical assumptions prevalent in the literature led to a plurality of conceptualizations and mathematical definitions of MP (Kaminski, 2007; d'Avella & Tresch, 2002; Omlor & Giese, 2007; Hart & Giszter, 2013; Chiovetto, d'Avella & Giese, 2016), each assuming different generative models, i.e., assumptions on how a specific value of a data stream is generated, underlying the respective definition. The following paragraph will summarize central mathematical properties of four common conceptualizations: the synchronous MP model (Ivanenko et al., 2005), the spatial MP model (Tresch, Saltiel & Bizzi, 1999; Ting & Macpherson, 2005), the time-varying synergies model (as introduced in d'Avella & Tresch, 2002) and the anechoic mixture model (as introduced in Omlor & Giese, 2007).

i. Synchronous MP model

Central to synchronous MP is the assumption that the movements resulting from the activation of one MP are executed in a timely coordinated manner (Ivanenko et al., 2005). This implies a timely coupling, or covariation, between the movements or muscular activations that belong to one MP. A synchronous MP hence is equivalent to a movement with a stereotypical time-course that can be executed by various effectors. For example, both a grasping motion with the arm and a kicking motion with the leg that show equal temporal characteristics would be considered as belonging to one MP in the synchronous model. Synchronous MP can be derived from EMG- or kinematic data (Ivanenko et al., 2005). The assumed generative model complies with the equation

$$x_{i,j}(t) = \sum_{s=1}^S w_{(i,j),s} \text{MP}_s(t) \quad (1)$$

The signal x in the current trial i for the joint j at a specific point in time t can be reconstructed by summing all number of weighted MP. A weight w in a specific trial i

affects which joint the MP is executed with and is used to scale the features of the MP which unfolds stereotypically over time (Ivanenko et al., 2005).

ii. *Spatial MP*

Spatial MP are based on the assumption that movements belonging to one MP are defined by the spatially stereotypical co-activation of groups of joints or muscles (Tresch, Saltiel & Bizzi, 1999; Chiovetto, d'Avella & Giese, 2016). Applied to, for example, EMG data, this implies that synergistically acting muscles are instantaneously co-activated. The following generative model is assumed for spatial MP:

$$x_{i,j}(t) = \sum_{s=1}^S w_{i,s}(t) MP_{s,j} \quad (2)$$

The signal x in trial i and for joint j at time point t is generated by s number of MP, weighted with w which scales the execution of the MP over time t and is characteristic for one MP s . The MP, unlike in the synchronous model in equation (1), now describes a spatially stereotypical primitive with a time course that strongly depends on the weight $w_{i,s}(t)$; the weight w is scaling the timely execution of the MP which in this model cannot operate with various effectors. An example: muscular efforts needed for arm wrestling, including tension in the arm and in the back for balance control, could be controlled by the same spatial primitive as is pushing a lever on a table. In the example, the spatial muscle patterns are equal; nevertheless, the time-course of the execution of the spatial MP differs from one situation to the other (depending on the opponent's strength). Also, postural control in varying contexts or with different expectations on impedances can be considered one MP in this model (Ting & Macpherson, 2005, see also *Figure 7* in the Appendix for a graphical illustration of the differences between synchronous and spatial MP).

iii. *The time-varying synergies model (d'Avella & Tresch, 2002)*

One innovation of the time-varying (TV) synergies model is that it alleviates the constraint of synchronicity, allowing for temporal delays between, e.g., synergy-activated muscles to be non-zero. The observation that muscles that are assumed to belong to one synergy often display a coordinated, but asynchronous activation pattern (with a temporal delay but a tight dispersion around the delay) led to the introduction of the TV model. The algorithm introduced in d'Avella & Tresch (2002) allows for the simultaneous analysis of more than

one synergy and is based on previously introduced NMF-techniques (e.g., see Tresch, Saltiel & Bizzi, 1999). The pattern of muscular activation is given by:

$$x_{i,j}(t) = \sum_{s=1}^S w_{i,s} MP_{s,j}(t - \tau_{i,j}) \quad (3)$$

With $x_{i,j}(t)$ representing the signal at time point t . $w_{i,s}$ is an amplitude coefficient that scales the activity pattern s in trial i , represented in vector space. The MP is a consistent movement pattern. Also, the model considers the delay t and the actual delay between activated joints τ . Consequently, a given episode is not only characterized by the relative muscle activation amplitude, but also by coordinated relative activation timings; adding an additional parameter to describe a synergy when compared to the synchronous MP model above ($M \times T$ vs. M , with M muscles and T time steps in a synergy). In this model, the same synergy can characterize patterns that are shifted in time and affect various body parts. A small number of the time-varying synergies can give rise to more complex behaviors such as walking, offering an efficient means of representing the data with few components.

iv. The anechoic mixture model (Omlor & Giese, 2007)

Similar to the TV synergies model above, the anechoic mixture model is based on the observation that the modeling of specific data streams, such as EMG signals or joint trajectories, requires the compliance with the assumption of time shifts. Omlor and Giese's (2007, 2007a) approach methodologically uses a different mixing model. Instantaneous mixtures, i.e. the linear superposition of sources without timely delays in order to reconstruct a signal, are suspected to be inadequate for the modeling of movement data. Anechoic mixtures for the linear combination of sources *with* time delays allow for a certain level of asynchronicity in the signal while still identifying the respective MP. In the anechoic mixture approach, the following generative model is assumed:

$$x_{i,j}(t) = \sum_{s=1}^S w_{(i,j),s} MP_s(t - \tau_{i,j}) \quad (4)$$

Signal x at time point t is approximated by the combination of a weight w , which scales the effector j in trial i , and the S number of MP which signifies a characteristic pattern that can be executed with timely delays ($t - \tau_{i,j}$). The anechoic model differs from the TV-synergy model, as the characteristic, time-shifted component MP is no longer applicable to just one specific joint j but can impact various effectors. The effector is determined through the

weight w in this model. Further, no initial assumptions on the size of $\tau_{i,j}$ are required in this approach. A two-step extraction algorithm has been introduced for the identification of MP with an anechoic mixture model which has outperformed conventional BSS-methods for synthetically generated and human kinematic data (Omlor & Giese, 2007a). This superiority might be observable because of the integration of temporal delays in the model and the flexible adjustment according to data characteristics. For an overview and further details about the differences between the concepts, see *Table 4* and *Figure 7* in the Appendix.

Recent developments towards a unifying framework for MP identification

The plurality of different theoretical and methodological approaches has made it difficult to interpret and generalize findings across different study contexts. This motivated attempts to identify a framework that unifies all perspectives. Recently, Chiovetto, d’Avella and Giese (2016) showed that the many prevalent definitions of MP relate to each other as they can be considered special instantiations of the anechoic mixture generative model (eq. 4). By adding constraints, the anechoic mixture model can be made equivalent to other theoretical MP definitions; for example, for the special case when $\tau_{i,j} = 0$, the anechoic mixture model is equivalent to the generative models that underlies the synchronous model in equation (1). Furthermore, the authors suggested one algorithm, referred to as the Fourier-based Anechoic Demixing Algorithm (FADA, Chiovetto & Giese, 2013), for the extraction of various types of MP. The FADA has shown better or similar performance when compared to other unsupervised learning techniques (Chiovetto, d’Avella & Giese, 2016).

1.6. Experimental settings for the investigation of MP

Research on MP is most commonly based on EMG or kinematic data in either animal (e.g., frog, octopus, cat, monkey) or human movement or locomotion (Sumbre et al., 2001; d’Avella & Tresch, 2002; Omlor & Giese, 2006; Hart & Giszter, 2013; Endres, Chiovetto & Giese, 2013). In recent years, a shift in attention from simple to complex full-body movements led to many interesting experimental approaches (Mussa-Ivaldi, 1999). For example, Kaminski (2007) examined full-body reaching movements; and Omlor and Giese (2006, 2007a) extracted emotion-specific primitive patterns from walks (actors walking angrily, happily or sadly). However, most evidence concerning MP relies on movement data recorded in highly repetitive, controlled laboratory settings that do not generally encourage flexible adjustment and improvisation that everyday human motor behavior is characterized by. For this reason, various researchers have demanded the analysis of spontaneous,

naturally occurring full-body movements, such as, for example, preparing a meal, shopping or, in particular, dancing (Endres, Chiovetto & Giese, 2016; Kaminski, 2007).

1.7. Dance

Defined as the rhythmical movement to music in the Oxford Dictionary (Stevenson (Ed.), 2010), dancing is a naturally occurring and complex sensorimotor activity. Moving to music requires not only the synchronization with complex auditory stimuli, but also the integration of coordinated whole-body movements, rhythm and visual information about, e.g., surroundings, a partner, the body etc. (Brown, Martinez & Parsons, 2005). Despite its complexity, dance is ubiquitous in all human societies and might have emerged at an early stage of the human evolution as a part of group rituals (Laland, Wilkins & Clayton, 2016; Farnell, 1999). Findings from comparative psychology suggest that when dancing, the human brain exploits neural circuitry involved in vocal or motor imitation as only species capable of imitation show sophisticated ‘dance’ or rather the integration of rhythm and movement (Laland, Wilkins & Clayton, 2016). Yet, elaborate dance seems to be a universal human characteristic that is only rarely observed in other species. Importantly, dances are frequently described as being of a modular organization, as individual steps, or rather discrete movement segments, are concatenated with one another and repeated cyclically (Endres, Chiovetto & Giese, 2016; Brown, Martinez & Parsons, 2005). These combinatoric features prompt a grammatical analysis and description of dance and thereby renders dance an interesting opportunity for the identification of MP.

1.8. Rationale and Hypotheses

Although the research interest in MP has been growing, the nature of MP during fully spontaneous and unconstrained human movements has, to our knowledge, not been researched in the past. The work by Williams, Toussaint and Storkey (2006), where MP were extracted from unconstrained movement data recorded during handwriting, constitutes an exemplary exception. However, MP derived from spontaneous and flexible full-body movement data are rarely encountered in the literature. For this reason, we record the movements of individuals dancing to music. Spontaneous dancing, alongside its potentially interesting modular features mentioned above, constitutes a natural and entirely unconstrained full-body motor behavior. We will evaluate the fit of the Synchronous MP model (eq. 1) with the movement data through statistical modelling and contrast it with the

TV-synergies model (eq. 3). The fit of the data with the respective model will be examined in an exploratory manner, guided by the following research questions and hypotheses:

a) *What MP conceptualization offers the best decomposition of the data?*

We will investigate the synchronous MP model (eq.1) and evaluate its fit with the data in a model selection process. Contrasting the synchronous model with the time-varying synergies model is of special interest (eq.3). The exploratory nature of this project prompts the model selection process, as only little is known about the question whether individuals tend to spontaneously couple their movements (suggesting the Synchronous model underlying movement generation) or rather tend to move more decoupled and asynchronously (prompting the TV-synergies model). Through this, the MP model for which the highest model evidence score is observed will be selected.

b) *How ‘universal’ are the MP extracted from dance moves?*

The extracted MP will lead to interesting conclusions regarding the universality of a specific set of MP. By basing the evaluation of the model fit on different ways of subdividing the data, we will gain insights into the following guiding questions and hypotheses (hyp):

Hyp. I) *Universal set of MP:* all participants use the same, ‘universal’ set of MP when dancing spontaneously; the movements are not influenced by any individual features or characteristics of the stimulus.

Hyp. II) *Individual repertoire:* A person relies on an individual set of MP, or his or her specific and ‘individual repertoire’, when dancing, irrespective of the music. This would be represented by a model that also considers the participant as a component of the data structure but does not assume any stimulus-specific adaptations of the movements.

Hyp. III) *Stimulus-determination:* All individuals show movement patterns that are rather determined by external stimuli, i.e., the songs of a specific genre, than by the person itself; and

Hyp. IV) *Individual \times stimulus interaction:* Movements are determined by the individual *and* the music; this would imply that the movements differ for every song on an individual level. This would be the case when a model that takes into account both the individual and the song leads to the highest model evidence.

2. METHODS

2.1. Stimuli

The stimuli used for the encouragement of spontaneous dancing in subjects were seven popular songs, including pop-, rock-, reggae-, and hip-hop-songs (see *Table 1* in the Appendix for a full list). The songs were selected so that they had a clear, regular beat while representing a wide range of genres and beats-per-minute (bpm, range= 96.0-138.3 bpm; mean = 111.96 bpm; SD=15.79). The presentation of the songs was organized on a Windows laptop using a standard media player (Microsoft: Windows Media Player, 2017) and speakers (Philips DS3205/12). The volume was adjusted to be around 70-80dB. The seven songs were presented in a random order. The total length of the playlist was 24 minutes and 54 seconds.

2.2. Participants

N= 17 physically healthy student volunteers (9 female, 19-33 years old, mean age=23,2 years, SD = 3,89) were recruited via university mailing lists and participated in the study for course credit or financial compensation (8€/hour, SGB-TER135'Kardinal Mechanisms of perception' project C6). All participants obtained extensive written and oral study information and gave their written consent to participate in accordance with the local ethics committee (*Appendix 1 and 2*, AZ2017-03k). Participants were instructed to dance spontaneously to the music 'as if nobody was watching'. The dancers were invited to take breaks and drink something whenever they felt it was necessary; additionally, they were explicitly offered to rest after two songs. The total study time of approximately 1.5 hours included the adjustment of the motion capture system, recording of the data and a brief post-experimental questionnaire with 12 items regarding demographic characteristics; mood, dance proficiency and dance-related self-esteem (*Appendix 3*).

2.3. Recording of movement data

The movements of all participants were recorded using an Xsens MVN Awinda inertial motion capture system (MTw User Manual, 2017) with 17 wireless motion trackers (MTw or 'units') that were attached to the participants' head, limbs, hands and feet using Velcro straps. MTw for sternum and shoulders were secured on a tight-fitting shirt that was worn above the participants' clothing. The system used allows for the unobtrusive and accurate assessment of 66 joints, movement velocity and acceleration of body segments by incorporating three-axial rate gyroscopes, accelerometers, magnetometers and barometers in

the individual units. Participants were instructed to hold a neutral pose (N-Pose) during the initial calibration process, defined as an upright, straight posture with the arms alongside the body and a parallel position of the feet. After successful calibration, raw motion data of the dancing participants was recorded with the update rate set to 60Hz. The term update rate (in contrast to ‘sampling frequency’) reflects that the recorded data undergoes a mathematical method referred to as ‘Strap Down Integration’ after it was recorded; this allows recording at a very high sampling frequency (1000Hz) while still enabling wireless transmission of the recorded signals to the station and a less severe computational load on the host PC. During recording, the data was filtered using the Xsens Kalman Filter for Human Movement (XKF-HM) which uses information gained during the calibration process for initialization of the filtering system (for details see MTw User Manual, 2017). The data was transferred to an Xsens Awinda Station equipped with a rotatable 2.4GHz antenna. The station was connected to a Windows PC via USB, where data assessment was managed using MVN Studio. Here, the recording process was monitored as the data is visualized in real time onto a skeleton model using algorithms provided by the manufacturers. Re-calibration was performed regularly after two songs and whenever necessary, e.g., when a participant indicated one of the MTw might have loosened. The rotations between body segments was represented using Euler angles, with the three dimensions Z (flexion/extension), X (abduction, adduction) and Y (internal/ external rotation).

2.4. Analyses

2.4.1. Pre-processing of movement data.

One necessary preparatory step before the derivation of MP from a continuous data stream using PCA is the pre-segmentation of the data into single motion segments (Jenkins & Mataric, 2003). In the present experimental setting, the pre-segmentation on the basis of the beats in the music is plausible. For this purpose, all songs used in this study were analyzed for their beat patterns using the ‘librosa’ library for python (McFee et al., 2015). This library offers functions for the derivation of beat sequences from music signals. The beat tracking system is based on the assumptions that the beats in the music can be perceived as ‘onsets’ with a specific strength in the audio signal. Also, the beat pattern of a given song is assumed to be regular and rhythmic. Both assumptions are reflected in equation 5 (Ellis, 2007):

$$C(\{t_i\}) = \sum_{i=1}^N O(t_i) + \alpha \sum_{i=2}^N F(t_i - t_{i-1}, \tau_p) \quad (5)$$

Where the beat pattern $C(\{t_i\})$ consists of N beat instants identified by the beat tracking system, $O(t)$ is the ‘onset strength envelope’ at time point t , i.e., an indicator derived from local acoustic properties, which becomes large when a beat has very likely occurred in the music analyzed, and a weighing term α which balances the importance of the two terms. The term $F(t_i - t_{i-1}, \tau_p)$ in equation (5) measures the deviation between a detected interbeat-interval $\Delta t = t_i - t_{i-1}$ and the ideal beat spacing τ_p . By maximizing the consistency between the detected interbeat-interval and the ideal beat pattern so that $\Delta t = \tau_p$ through a squared-error function, a time-sequence with a good score is identified in an iterative process and imposed onto the music signal (see *Figure 6* in the Appendix). The performance of the beat tracking system has been evaluated for synthetic training data and naturalistic songs (Ellis, 2007; McFee et al., 2015). The continuous stream of motion data was then aligned and pre-segmented with the derived beat sequences. Every 4th beat (equals one trial I) was used as a point of reference for the segmentation process. This yielded segments of varying length (average segment length = ~ 1.8736 s) which had to be resampled to a standardized length. The segments were resampled using a Gaussian process with squared exponential functions (Bishop, 2006) leading to segments of a consistent length ($T=100$ samples). The data therefor had the shape: $(I \times J), T$ after pre-processing with T number of time steps; I number of trials and $J = 66$ joints.

2.4.2. Data Analysis: Model selection

The generative models introduced above have shown to effectively decompose the data under investigation, which is why we chose a model selection approach to determine which one of the generative models is most likely to underlie the data (see Endres, Chiovetto & Giese, 2013; 2016). To illustrate this approach, suppose the data was generated by one model out of a given set of models M, \dots, M_k . ‘Model selection’ is the process of selecting one M_i of the models that has most likely led to the observed data (Wassermann, 2000). The selected model or models form a good basis for further interpretation or prediction. Several ‘information criteria’ (IC) for the evaluation of the model fit with the data have been proposed in the literature (Schwarz, 1978; Akaike, 1987; Endres, Chiovetto & Giese, 2013). These criteria help selecting a parsimonious model that still describes the data accurately enough, most commonly through the computation of the likelihood function (Endres, Chiovetto & Giese, 2013). In the framework of Bayesian statistics, the likelihood is also referred to as the ‘probability density’ of the data D given the model M under investigation. In the present analysis, one model M_i consists of the parameters $M_i = (\theta, \Phi, K, M)$ with θ

model parameters, such as the weight w and the spatial, temporal or spatio-temporal pattern MP as well as the noise in the data. Specifically, we assume that the noise will not exceed 2% in the present analysis. This noise level is required for the production of stable walking sequences in humanoid robotic systems (Clever et al., 2017). Φ are hyperparameters of the model, including the kernel function (see below; Endres, Chiovetto & Giese, 2013) and prior assumptions about the nature of the noise; K in the present case is the number of MP of a given model and M the model type (e.g. synchronous or TV-synergies). The model posterior $p(M|D)$ is related to the marginal likelihood, or model evidence $p(D|M)$, assuming a constant prior on $p(M)$ as no model is a priori favored over the other, via

$$p(M|D) = \frac{p(D|M)p(M)}{p(D)} \propto p(D|M) = \sum_{k=1}^K \int d\theta p(D|\theta, \Phi, K, M) p(\theta|\Phi, K, M) p(\Phi) p(K) \quad (6)$$

where $p(\Phi)$ and $p(K)$ are constant terms in the present case, as the respective model's hyperparameters and model order K are constant or selected a priori. The model evidence $p(D|M)$ is computed from the likelihood function $p(D|\theta, \Phi, K, M)p(\theta|\Phi, K, M)$ by marginalization over the model parameters θ ; consequently, the model evidence can be expressed as

$$p(D|M) = p(\Phi)p(K) \sum_{k=1}^K \int d\theta p(D|\theta, \Phi, K, M) p(\theta|\Phi, K, M) \quad (7)$$

This model evidence is approximated by the ‘LAP’ criterion for model selection, so that

$$p(D|M) = p(\Phi)p(K) \sum_{k=1}^K e^{\text{LAP}(D,K,M)} \quad (8)$$

The LAP (see Endres, Chiovetto & Giese, 2013) is given by

$$\log(p(D|\Phi, M)) \approx \underbrace{\log(p(D|\theta, \Phi, M))}_{\text{log-likelihood}} + \underbrace{\log(p(\theta|\Phi, M))}_{\text{log-prior}} + \underbrace{\frac{\dim(\theta)}{2} \log(2\pi) - \frac{1}{2} \log(|\mathbf{H}|)}_{\text{log-posterior volume}}$$

where θ is (θ, K) . The three above terms have different implications. Firstly, the log-likelihood can be interpreted as measuring the goodness of fit of M_i with the data D , similar to an ‘explained variance’- logic. Secondly, the log-prior constitutes a regularization term for small datasets; and thirdly, the log-posterior volume measures the volume of the posterior distribution of the parameters as \mathbf{H} indicates how well the parameters θ are constrained by

the data (a small \mathbf{H} therefore indicates that the data constrain θ well). In this analysis, the software underlying Endres, Chiovetto and Giese (2013) has been used for the computation of the LAP. All calculations were executed at the Marburg Cluster (marc2, see <https://www.uni-marburg.de/hrz/infrastruktur/zserv/cluster> (last accessed 09/28/17)) From equation (8), it becomes evident that $\sum_{k=1}^K e^{\text{LAP}(D,K,M)}$ (with K number of MP) can be approximated through the one k which maximizes LAP as the maximal LAP score in this case dominates the model evidence term. From the model evidence, approximated by the K with the max-LAP score, we can compute the Bayes factor (Lavine & Schervish, 1999; Good, 1985) which indicates when to favor one model over the other in the framework of a Bayesian Model comparison. In the present analysis, suppose we test the evidence for the model that assumes one universal set of MP, M_{ALL} (hyp. I) versus the evidence for the stimulus-determined model, M_{SONG} (hyp III). The model posteriors are given by:

$$p(M_{ALL}|D) = \frac{p(D|M_{ALL})p(M_{ALL})}{p(D)} \quad (10)$$

and

$$p(M_{SONG}|D) = \frac{\prod p(D_{SONG}|M_{SONG})p(M_{SONG})}{p(D)} \quad (11)$$

As $p(D)$ and $p(M)$ are constant terms, it follows that:

$$\frac{p(M_{ALL}|D)}{p(M_{SONG}|D)} = \frac{p(D|M_{ALL})}{\prod p(D_{SONG}|M_{SONG})} \quad (12)$$

The logarithm of eq. (12) leads to the expression

$$\log p(M_{ALL}|D) - \log p(M_{SONG}|D) = \log p(D|M_{ALL}) - \sum_{song1}^{SongN} \log p(D_{SONG}|M_{SONG}) \quad (13)$$

where again, the LAP can be employed for the approximation of $\log p(D|M_{ALL})$ and $[\sum_{song1}^{SongN} \log p(D_{SONG}|M_{SONG})]$ so that through marginalization we receive

$$\log(p(M_{ALL}|D) - p(M_{SONG}|D)) = \tag{14}$$

$$LAP(D, M_{ALL}, \max(k)) - \sum_{song1}^{SongN} LAP(D_{SONG}, M_{SONG}, \max(k))$$

$\log p(M_{ALL}|D) - p(M_{SONG}|D)$ is also referred to as the log-Bayes factor. The Bayes factor indicates when the data prompt the selection of one model over the other. Following the logic of equation (14), a positive log-Bayes factor prompts M_{ALL} over M_{SONG} , whereas a negative log-Bayes factor would lead to the contrary conclusion.

Alternative model selection approaches that are considered in the present analysis are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

I) Akaike Information Criterion (AIC, Akaike, 1987)

The AIC is based on information-theoretical considerations insofar as it labels a model as good when the Kullback-Leibler discrepancy (KLD) is minimal (Zucchini, 2000). The KLD quantifies the difference between two probability distributions, typically the empirical observation vs. an approximation, or model. In his work, Akaike has shown that the negative log-likelihood function can be considered an estimator for the KLD (Akaike, 1987). The AIC is then given by

$$AIC = -2(\log(p(D|\Theta, \Phi, M)) - \dim(\Theta)) \tag{15}$$

The goodness of fit of a given model is rewarded through the log-likelihood function which grows when the probability of the data D given the model M_i increases. This term alone would inadequately favor large, complex models. For this reason, the AIC also contains a penalty term ($-\dim(\Theta)$) that disadvantages more complex models over simpler models. Consequently, the AIC as a penalized IC favors the less complex model unless the more complex model offers a significantly better fit with the data. A low AIC score indicates good model fit. The AIC is widely used for model selection purposes in many areas of cognitive psychology (Akaike, 1987; Wagenmakers & Farrell, 2004).

II) Bayesian Information Criterion (BIC, Schwarz, 1978)

The BIC is another common tool for model selection purposes, e.g., in sociology. It aims at indicating when the evidence derived from data prompts one particular model over

the other (Weakliem, 1999). The BIC is an asymptotic approximation to the computationally demanding Bayesian model selection (BMS) where the computation of the likelihood function $P(D|M_i) = \int (D|\theta_i, M_i) \pi(\theta_i|M_i)$ is required, with $\pi(\theta_i|M_i)$ representing the ‘prior density’; or, in other words, prior beliefs or information a researcher might have before the observation of the data. The BIC does not require determining prior densities in advance and is, furthermore, computationally less costly. It is given by

$$\text{BIC} = -2 \left(\log(p(D|\Theta, \Phi, M)) - \frac{1}{2} \dim(\Theta) \log(N) \right) \quad (16)$$

N represents the number of observations; all other parameters are equal to the AIC described above. Other than the AIC, the BIC acknowledges the fact the model selection process can be distorted by large samples in an undesirable manner (Wagenmakers & Farrell, 2004). Consequently, the penalty term of the BIC grows logarithmically with the number of observations N and therefore is not distorted by large N when $n \rightarrow \infty$. When comparing the central assumptions of BIC vs. AIC, it has been noted that the BIC assumes the model selected in the model selection process is the ‘true’ model; it aims at indicating the degree of certainty that the investigated model is the true generative model. In contrast, the AIC calculates the information-theoretic KLD for each model without presuming that one of the M_k models is necessarily true (Kuha, 2004; Wagenmakers & Farrell, 2004).

The LAP introduced in equation (9) has been shown to perform as good as AIC and BIC on synthesized as well as actual kinematic data (Endres, Chiovetto & Giese, 2013). In fact, LAP’s performance is assumed to become similar to BICs’ as $n \rightarrow \infty$. An important advantage of LAP over BIC and AIC is that it efficiently deals with temporal correlations in the data. Data-related temporal dependences have shown to adversely affect the performance of conventional IC (Endres, Chiovetto & Giese, 2013). The LAP effectively handles temporal dependences in the data structure and is therefore particularly suitable when selecting a model for data that is assumed to show temporal dependences in their structure; such as movement data (Endres, Chiovetto & Giese, 2013; Bishop, 2006). From the temporal dependences in the data, or the underlying covariance matrix, it is possible to determine the ‘smoothness’ of the movements. The higher the temporal dependencies, the more the data points are interrelated and the smoother the trajectory. Kernel functions of Gaussian Processes for MP generation can determine how much the data points co-vary over the course of the movement and hence make assumptions regarding the temporal smoothness the dataset is characterized by (Bishop, 2006; Rasmussen & Williams, 2006). They can add

‘soft’ constraints to the model comparison process through hyperparameters Φ . Nevertheless, the regulations following the kernel function can be overridden in case the evidence from the data strongly suggests so. In particular, three types of kernel functions are assumed to have appropriate features for the analysis of motion capture data: the wave kernel, the radial-basis function (RBF) kernel (Endres, Chiovetto & Giese, 2013), and the polynomial kernel (Bishop, 2006; Endres, Chiovetto & Giese, 2013; 2016). In the analysis of the data, we will employ an RBF kernel which is given by

$$k(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (17)$$

$\|x - x'\|^2$ is the squared Euclidean distance between the two feature vectors. σ is a free parameter which corresponds to the width of the RBF-kernel. It will be set to 30 in the present analysis. This corresponds with the a priori assumption that the temporal dependences in the data will last for the duration of approximately 1/3 of one trial. Consequently it is assumed that an initial data point would not be correlated with another point that is observed at 2/3 of the trial duration.

During initial data analyses, we were unpleasantly surprised by the heavy computational load of the extraction and evaluation of time-varying synergies. The computation of the least complex models was extremely time consuming, which is why we had to abandon the testing of hypothesis A. For similar reasons, the model selection process was performed on data of 5 participants only.

In sum, the Synchronous MP concept with a range of corresponding model orders (3-34 MP) was evaluated with the movement data from volunteers using three information criteria (IC): the AIC, the BIC; and the LAP. In line with the hypotheses, the data was sorted so that 4 different models are tested: M_{ALL} , a ‘universal model of MP’ (hyp. I), $M_{PARTICIPANT(P)}$, where each individual is thought to rely on an individual repertoire when dancing (hyp. II); M_{SONG} , a model that assumes dancing to be stimulus-determined (hyp. III); and $M_{P\&SONG}$, where dancing is assumed to be determined by the interaction of individual and stimulus (hyp. IV). The models were represented by different ways of concatenating the data; for the M_{ALL} , MP were learned from all movement data; for M_P all data recorded for one participant was concatenated before learning the MP; for M_{Song} , all data recorded during one song was concatenated; and for $M_{P\&Song}$, each recording of one participant dancing to one song was considered individually (code examples in the supplementary material (CD)).

3. RESULTS

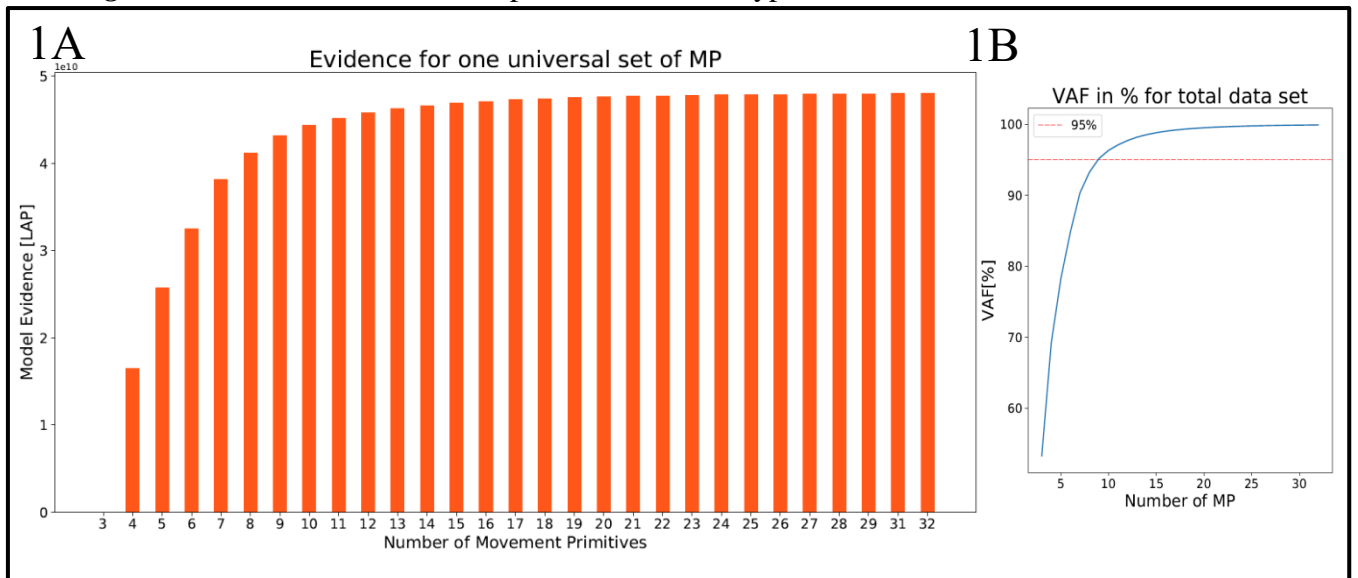
3.1. Questionnaire

Full descriptive results from the post-experimental questionnaire are illustrated in *Table 2* (Appendix). Participants indicated that they felt comfortable in the study situation ($m=6.0$ ($SD=.866$) from 1=very uncomfortable to 7=very comfortable). 5 out of 17 volunteers reported advanced experiences as dancers.

3.2. Hypothesis 1): ‘Universal set of MP’

For the testing of the model of one ‘universal’ human set of MP, the data fit of the Synchronous MP model type with three to 32 components was evaluated. Results indicate that 9 components account for 95.19% of the total variance in the data (VAF, *Figure 1B*). LAP score was maximized for the most complex model with 32 MP (*Figure 1A*). The AIC score for M_{All} were maximized for a 32 MP model, whereas BIC score was highest for 19 MP (see *Figure 8* in the Appendix).

Figure 1. Model Evidence and explained variance: hypothesis of one universal set of MP



Annotations. Model Evidence (LAP) and variance accounted for (VAF) of the ‘universal set of MP’ hypothesis. LAP scores reported in *Figure 1A* are relativized with the minimum score observed for this model; in this case, a synchronous model that considers only 3MP. Positive LAP scores here indicate good model fit. Relative LAP scores were highest for a model that considers 32MP. In *Figure 1B*, the proportion of variance that a ‘universal set of MP’ model explains (in relation to the total data variance) is illustrated. The 95% level is reached with 9 components.

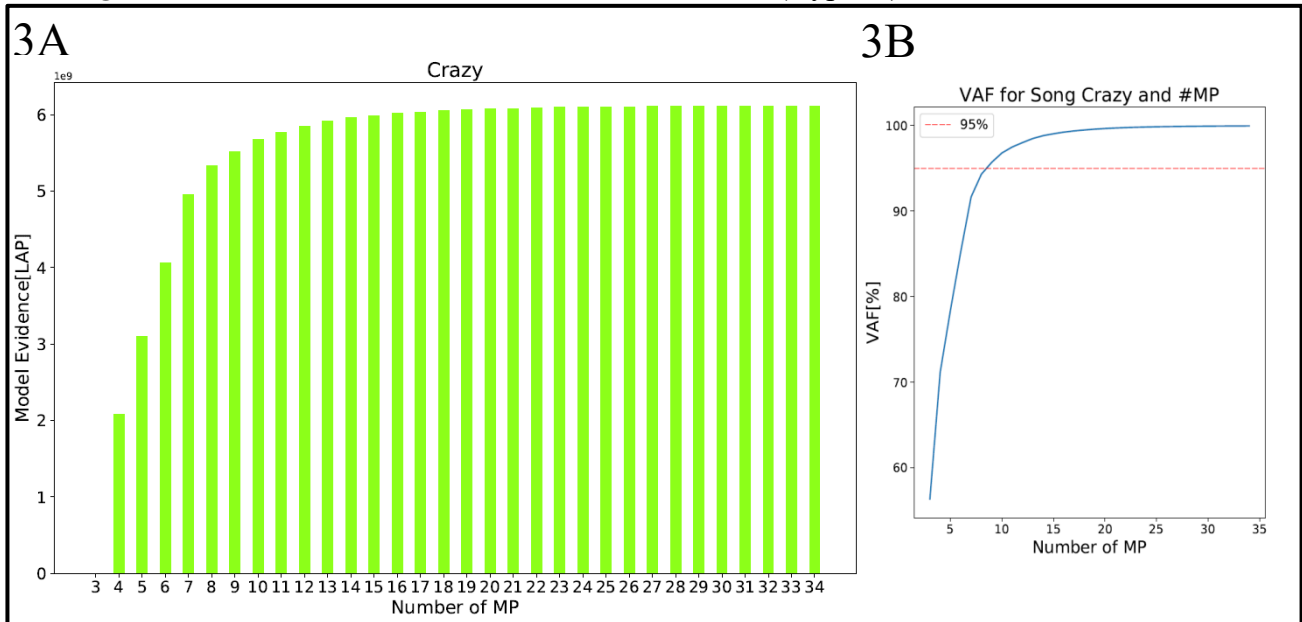
3.3.Hypothesis II): ‘Individual repertoire’

The results of the decomposition for the ‘individual repertoire’ model in the synchronous MP framework are illustrated in *Figure 2A-C*. On average, 9 synchronous MP accounted for 95.32% of the total data variance (SD=.94, *fig. 2A*), while a model with 3 components explains ~55.5% of the total variance (SD=5.69). LAP score is highest for a model with 34 components (see *Figure 2A*). AIC scores were highest for the most complex model with 34 MP for all participants. BIC scores, nevertheless, show differential courses depending on the participant (P): for P1, BIC score was maximized for a model with 17 MP; whereas for P2, a model with 29 synchronous MP maximizes BIC (see *Figure 9* in the Appendix)

3.4.Hypothesis III): ‘Stimulus-determined set of MP’

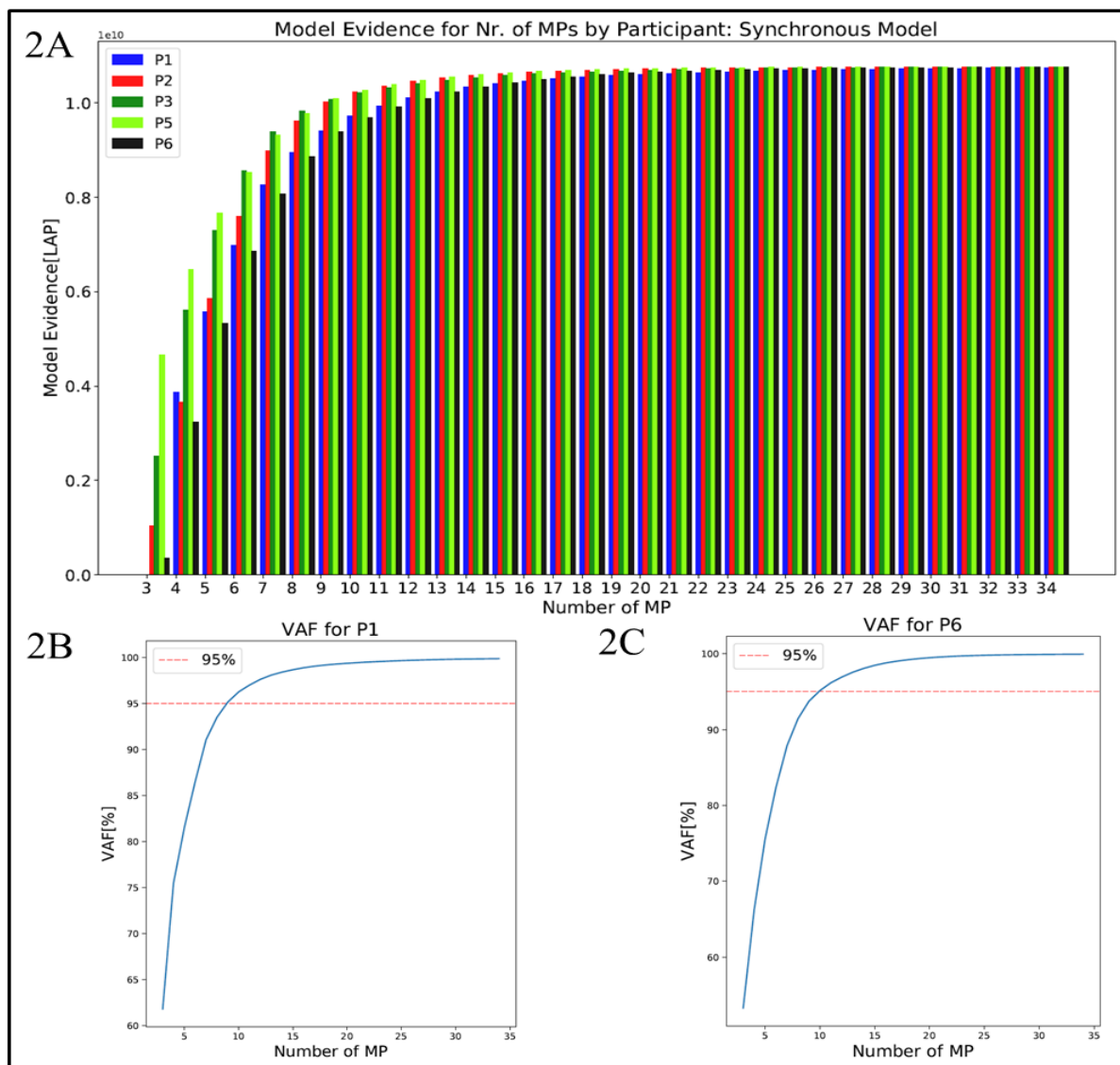
For the ‘stimulus determined model’, exemplary results of the model evidence evaluation are depicted in *Figure 3* and *Figure 10* (Appendix). A M_{Song} with 34 components achieved the highest LAP scores; and 9 components account for approximately 95% of the total variance in the data. AIC scores as well were highest for the most complex model with 34 MP for all songs. The BIC, however, was maximized for a less complex model for all songs (see *Figure 11*, Appendix, for an example). It is maximized for a model with ~26 MP across all songs.

Figure 3. Evidence for a ‘stimulus-determined’ model (Hyp. III)



Annotations. Model Evidence and explained variance for a stimulus-determined model of movements. **3A:** Model Evidence [LAP] for one song. For model evidence scores of the other stimuli, see *Figure 10* in the Appendix. For graphical purposes, all LAP scores were relativized with the minimal LAP score which was observed at a model with 3 MPs. The highest LAP score is observed for a synchronous model with 24 components. **3B:** Variance accounted for (VAF, %)

Figure 2. Evidence for an ‘individual repertoire’ (Hypothesis II) of synchronous MP

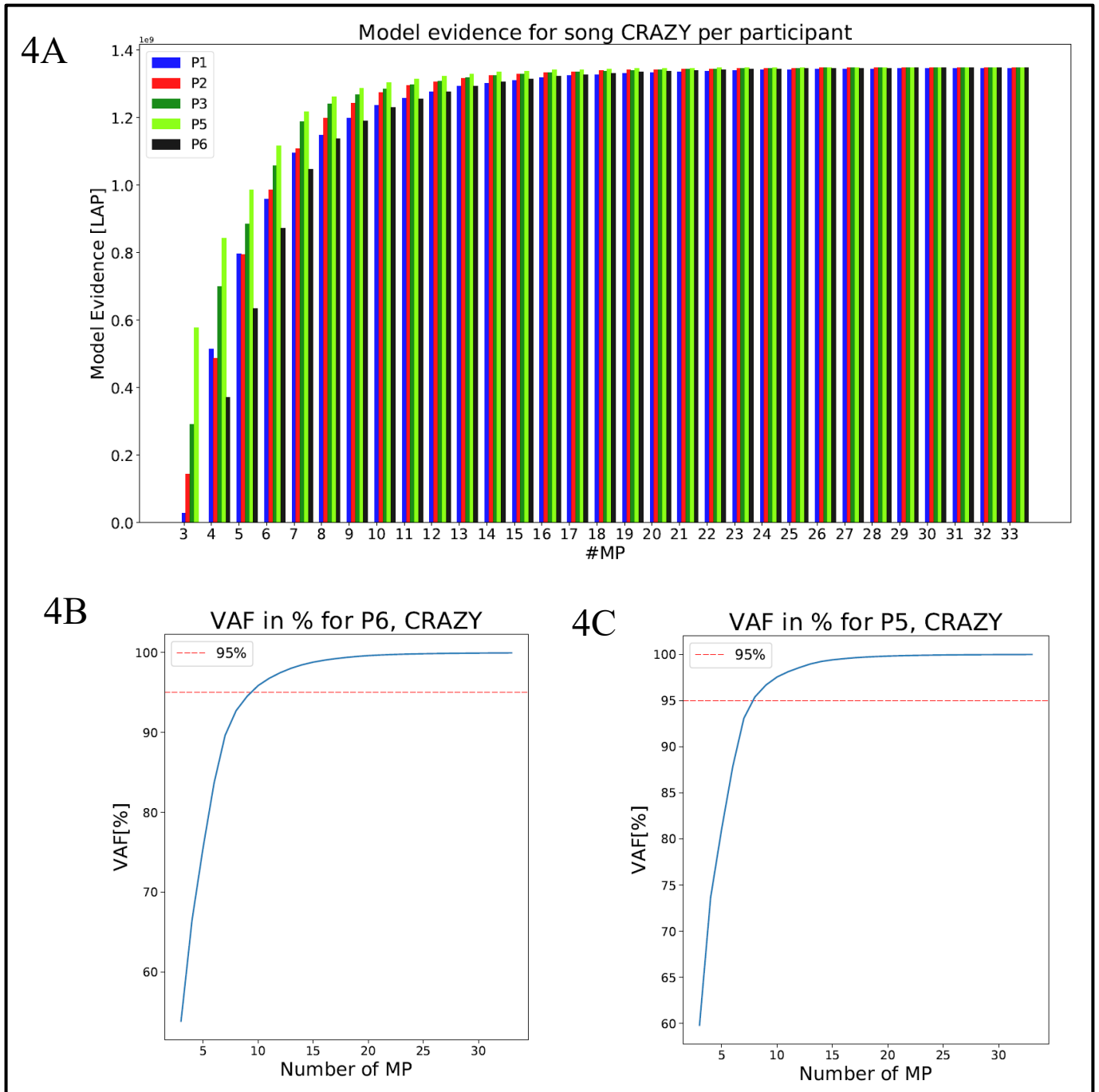


Annotations. 2A. The model evidence (LAP-criterion, here: relative to a minimum score) for the synchronous model. Here, the hypothesis of an ‘individual dance-repertoire’ is tested. LAP score was best for models with 34 components and worst for models with 3 MP. B,C. The variance accounted for by a synchronous model with the respective number of MP, each one specific for one participant (P1 and P6). 95.17% of the movements of participant 1 (P1) are explained by 9 MP. For participant 6 (P6), 10 MP explain 95.16% of the movement data.

3.5.Hypothesis IV): ‘Individual *and* stimulus-determined set of MP’

An exemplary result for the individual and stimulus interaction model are illustrated in Figure 4. Maximum LAP scores were observed for the most complex model with 33 MP across all stimuli and participant combinations (see also the supplementary material on the CD). VAF for one participant and one song led to the result that 9-10 MP lead to a satisfying decomposition of the data, explaining more than 95% of the total variance of the data.

Figure 4. Individual and stimulus interaction model: Model Evidences and VAF (hyp. IV)



Annotations. Model evidence scores (LAP) and VAF in % for a participant- and stimulus determined model of dance movements. **4A.** LAP scores relative to the minimal LAP score observed for the participant and stimulus model for the song ‘shook’. Large LAP scores indicate a good model fit. The lowest LAP score for all participants and this specific song was observed for a synchronous model with 3 MP. The highest relative LAP score was achieved for a model with 33 MP. **4B,C.** VAF for song ‘shook’ and participant 6 (P6, **4B**) and participant 5 (P5, **4C**). For P5, 9 MP explain 95% of the total data variance. 95% of the variance in the data for P6 are explained by 10 MP.

3.6. Model comparison using Bayes factor

The models were tested against each other using the log-Bayes factor introduced in section 2.4.2. An overview of the computation of the respective log-Bayes factor for a specific model comparison and the results of this comparison are summarized in *Table 1*. A comparison of $M_{All} - M_{Song}$ showed that the data suggests a higher evidence for M_{Song} than for M_{All} (Bayes factor -38 131 273.4749). When comparing $M_{All} - M_P$, M_P was identified as the model with higher evidences (Bayes-factor -33 919 960.8721). A Bayes factor of 4 211 312.6028 was observed for the comparison $M_{Song} - M_P$. For the comparison of $M_{Song} > M_{P\&Song}$, a Bayes factor of -1 618 926.4209 was observed; and a factor of -5 830 239.0237 for the comparison $M_P > M_{P\&Song}$. A further illustration of the model comparison results can be found in the Appendix (*Figure 12*).

Table 1. Results of the model comparison (log-Bayes factors)

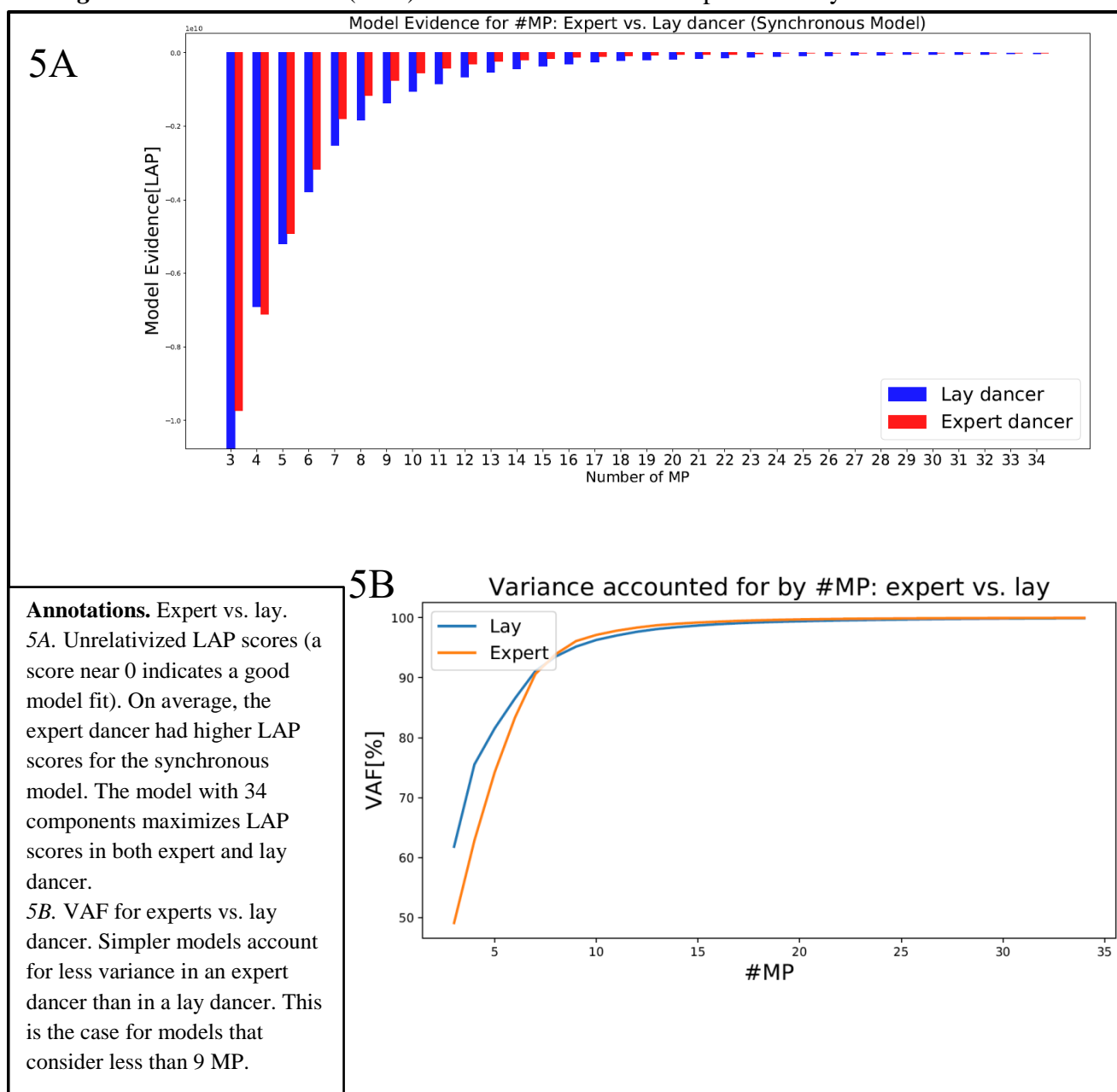
Models	Computation of (log-)Bayes factor	Bayes factor
$M_{All} > M_{Song}$	$\max(LAP) M_{All} - \sum_{song\ 1}^{song\ 7} \max(LAP) M_{song}$	-38131 273.47
$M_{All} > M_P$	$\max(LAP) M_{All} - \sum_{P1}^{P6} \max(LAP) M_P$	-33919 960.87
$M_{Song} > M_P$	$\sum_{song\ 1}^{song\ 7} \max(LAP) M_{song} - \sum_{P1}^{P6} \max(LAP) M_P$	4 211 312.60
$M_P > M_{P\&Song}$	$\sum_{P1}^{P6} \max(LAP) M_P - \sum_{song1}^{song7} \sum_{P1}^{P6} \max(LAP) M_{P\&Song}$	-5 830 239.02
$M_{Song} > M_{P\&Song}$	$\sum_{song\ 1}^{song\ 7} \max(LAP) M_{song} - \sum_{song1}^{song7} \sum_{P1}^{P6} \max(LAP) M_{P\&Song}$	-1 618 926.42

Annotations. Overview: computation of the Bayes-factor and results. M_{All} is the model that assumes a universal set of MP (hyp. I); M_P is the individual repertoire model (hyp II), M_{Song} is the stimulus-determined model (hyp. III), and $M_{P\&Song}$ is the model that assumes individually varying and stimulus-specific reactions towards every song (hyp. IV).

3.7. Additional Analysis: Expertise status

A comparison of lay- and expert dancers, categorized by their post-experimental questionnaire responses, was done. P2 identified as a ‘professional dancer’ as she actively danced ballet and took hip-hop dance classes. P1 reported never having participated in any sort of dance classes in the past. Results are illustrated in *Figure 5*. For both lay and expert dancer, LAP scores were highest for the most complex synchronous model with 34 MP; nevertheless, the observed LAP scores for simpler models varied depending on the expertise status: on average, higher LAP scores were observed for the expert dancer. The VAF for the lay dancer was higher for simpler models than it was for expert dancers (lay : expert; 3MP: 61.2% : 49.1%, 4MP: 75.5% : 62.8%, 5MP: 81.6% vs. 74.2%, 6MP: 86.5% vs. 83.5%). For both expert and lay dancer, 9 components accounted for 95% of the variance.

Figure 5. Model evidence (LAP) and VAF scores for an expert vs. a lay dancer



4. DISCUSSION

In the present work, the performance of the well-defined Synchronous MP approach is investigated in unconstrained human dance. Participants danced to 7 contemporary songs of varying genre and tempo without further instructions by the experimenter. The participants' naturalistic full-body movements were recorded using an unobtrusive inertial motion capture system that described rotations, abductions and flexions among adjacent body segments in Euler angles. The beats in a particular song were extracted and used for the pre-segmentation of the data and consequently, the model fit of the Synchronous model with the data was evaluated using VAF, AIC, BIC and LAP for 5 of the participants. Further, it was investigated whether the participants' movements were determined by one universal set of MP, by their individual repertoire; by the stimulus or song; or through the interaction of individual and stimulus-specific characteristics.

4.1. Summary and interpretation of results

The analysis of the motion capture data revealed that the Synchronous MP approach offers a feasible opportunity for the decomposition of the participants' movements when dancing to music. Regarding the variance accounted for (VAF), a very parsimonious decomposition of the dataset was suggested with approximately 9 to 10 MP accounting for the large proportion of the variance in the movement data. This was observed across all models conditions.

The LAP criterion for model selection was employed to identify the model that decomposes the data while still representing the data accurately. Almost all LAP scores observed in the analysis were highest for the most complex models across all model constraints. One reason for this is the noise level of 2% which was postulated a priori. This rather conservative noise level would be necessary for the animation of a humanoid robotics like the system described in Clever and colleagues (2017). However, a much lower noise level might have led to higher LAP-scores for less complex models. As a high proportion of variance is explained with models of a much lower order, it seems that LAP scores were still increasing as more complex models account for small jitters in the data that are not necessarily perceptually detectable. The fact that the observable LAP scores keep increasing when approaching the most complex models leads to the conclusion that potentially, even more complex models with a higher model order could finally maximize the LAP score under the present circumstances and assumptions. The same was true for the AIC score which in most cases was maximized for the most complex models, regardless of the model

type ($M_{All}, M_P, M_{Song}, M_{P\&Song}$). The BIC score, nevertheless, was often maximized for less complex models. This could be due to the nature of BIC's penalty term which acknowledges potential distortions by large N , which were given in the present study. Hence, employing BIC to this model selection process could lead to a rather conservative penalization of complex models.

Unfortunately, due to the unexpected computational load it was not possible to implement the tests for Hypothesis A: a comparison between synchronous MP and time-varying synergies (d'Avella & Tresch, 2002). Yet, regarding the universality of the MP humans rely on when dancing (Hyp. B), a model comparison process using log-Bayes factors led to interesting results. It appears that a model that assumes one universal set of MP all participants rely on regardless of situational factors is very inappropriate for the recorded data. This model turned out to be less supported by the data than the model that assumes an individual repertoire or a stimulus-determined model. Results indicated that dancing is emerging from an interaction between stimulus- and individual characteristics, suggesting an individual and situation-independent 'dance style' which is flexibly interacting with situational cues from the music. Interestingly, the model that assumes dancing a more stimulus-driven activity is suggested second-best by the present analysis as it turned out to be superior over the 'individual-repertoire'-model. This finding implies that the movements of the participants were more similar on an inter-individual level, depending on the specific song, than they were intra-individually: people seem to dance rather similarly to a specific song and these similarities are even stronger than the 'intraindividual consistency', or personal 'dance style'. This is in line with the notion that dance emerged as a part of group rituals in the human evolution (Laland, Wilkins & Clayton, 2016). People might intuitively align their movements according to implicit social practices rather than relying on personal style. However, future research will need to replicate this finding as it might have emerged artificially due to the specific methodological approach.

4.2.Limitations

Study Design. Unfortunately, the potential of the recorded dataset could not be recognized to the fullest. Only a small group of the assessed participants was considered for analyses, which is clearly not due to the quality of the data but purely due to timely restrictions. Further research with this dataset could focus on the extraction of the other MP model types from the full dataset (*Table 4* in the Appendix). Also, many comparisons as in section 3.7, where an expert dancer was compared to a lay dancer, are imaginable. For

example, a comparison of male vs. female participants could be of interest; especially as male participants evaluated their dance abilities significantly higher than female participants in the post-experimental questionnaire. Also, a comparison of the most vs. the least ‘danceable’ song, according to the judgement of the participants, could be informative.

It could be argued that speaking of naturally occurring, spontaneous dance in the case of the present data is not justified, as the situation of dancing in a laboratory setting with an experimenter observing the data recording is not a situation that generally encourages dance ‘as if nobody was watching’. Consequently, it is imaginable that participants felt inhibited and nervous during dancing, leading to stiffer-than-natural movements that do not generalize to dancing under natural conditions. The results from the questionnaire, however, point into a different direction: participants reported that they felt comfortable in the study situation (on a scale from 1 (very uncomfortable) to 7 (very comfortable): mean=6.00, SD=.866). Nevertheless, a validation of the model fit of the Synchronous model in a more natural ‘dancing environment’ could be of interest.

Finally, in the present analysis, beat sequences derived from the audio signals were used for the pre-segmentation of the movement data stream. This seemed plausible in the present case as it is certainly intuitive that humans adjust their movements so that they match the beats of the music; however, an explicit experimental validation of this assumption is still pending. The results from the model comparison (section 3.6) indicate that participants did indeed adjust their movements to the beats of music they heard; however, a finger-tapping experiment, similar to the procedure described in Janata and Grafton (2003) could be used for confirming that humans actually ‘move to the beat’.

Model selection procedures. Due to the unexpectedly heavy computational load the analysis of the present big dataset brings with it, we were unable to extract other components than the ones resulting from the Synchronous MP model. The extraction of time-varying synergies took approximately 29 hours on a modern computer cluster (‘marc2’) for the first $M_{P\&Song}$ entry, which is incompatible with a pending deadline. In the future, the adaption of the extraction algorithms for large datasets and the comparison of the models are indispensable, making the extraction of time-varying synergies from big data more feasible.

Also, the present analysis was limited to an RBF-kernel function with the width set to 30; which leads to a-priori assuming that movements are correlated to each other for the duration of 30% of one trial I whereas the movements of the beginning vs. towards the end

of one trial are uncorrelated. This assumption only imposes a soft constraint on the MP extraction and was overridden if the data suggested so; however, it would have been interesting to consider a range of settings for the kernel parameter. This process would allow identifying the best representation for the structure of the covariance-matrix underlying the data, potentially increasing the total model evidence scores when a good fit is achieved. Smaller settings for the width-parameter, for instance, would indicate that participants' trajectories were characterized by faster and more jerky movements within one trial I which is plausible. Also, considering different types of kernel functions would be interesting. A 'polynomial kernel', for instance, allows conclusions concerning the so-called 'Minimum-X' MP (Flash & Hogan, 1985). Depending on the order of the polynomial function considered, the MP could be characterized as Minimum-X with $X \in \{\text{acceleration, jerk, crackle, snap}\}$; i.e., the smoothness of the trajectories would be described as minimizing the acceleration or the jerk present in the movements (see Endres, Christensen, Omlor & Giese, 2011 for an investigation of polynomial kernel functions in movement perception). Also, 'wave kernels' have shown to be an appropriate choice when modelling human movement data (Endres, Chiovetto & Giese, 2013).

4.3.Future research and applications

While the extraction of MP from movement data constitutes an exciting opportunity for deciphering principles the motor system might rely on when planning and executing complex movements, research suggests that MP could also play an important role in perceptual mechanisms (Knopp, Schubert, Endres, 2017; Thill, Hemeren & Durán, 2011). In other words, a discrete repertoire of MP might not only simplify movement planning and execution, but also the perception of movements by forming a reliable basis of perceptual categories. Numerous findings suggest that perception and action might be closely intertwined, a notion also referred to as the 'Common Coding Theory' of perception and action (Prinz, 1990). Given the astonishing sensitivity of the perceptual system for the recognition of movements, it is assumed that movement perception has played an exceptional role in the human evolution (Johansson, 1975; Kozlowski & Cutting, 1977, Loula et al., 2005). Future research could attempt the perceptual validation of the extracted components, i.e.: are automatically derived synchronous MP consistent with a segmentation performed by humans? For choreographed and task-related actions, some studies that compare human and automatic MP derivation are available in the literature (see, e.g., Knopp, Schubert & Endres, 2017; Endres et al., 2011), suggesting evidence for the assumption that

action and perception indeed rely on similar representations of MP. Future research with the data recorded for this project could further investigate this notion, e.g., in a psychophysical experiment. Here, human observers could be instructed to segment the movements of an avatar animated with the recorded movement data. One of the numerous questions this experimental design could answer is whether human observers rely on the beats in the music for their segmentation decisions. As results of the model comparison suggest that dancers also rely on musical events when dancing, hearing the music should also determine the decisions of an observer who is segmenting the movements of an avatar.

Furthermore, the models investigated in the present analysis could provide a useful basis for the animation of naturalistic human movements. Computer graphics can be facilitated by a repertoire of distinct movement segments (Endres et al., 2011; Badler, Phillips & Webber, 1993). Also, the present movement decomposition could be used to animate a humanoid robotic system (Clever et al., 2016, 2017). It would be interesting to see whether the synchronous components extracted in this study suffice for making a robot convincingly look like it is ‘dancing’. For this purpose however, it would be necessary to find the K or number of MP for which the LAP is maximized under the present conditions (conservative noise level and large movement repertoire), as analyses suggest that LAP might be maximal for even more complex models than the ones considered here.

5. Conclusion

The present work demonstrated that the synchronous MP model is suitable for the extraction of a distinct number of primitive motions, or MP, from a continuous stream of a large unconstrained human movement dataset. Results indicated that dancing emerged from the interaction of individual and stimulus-specific features. This suggests that dance is a stimulus-driven behavior which is interacting with an individually stable style. Generally, complex models of higher orders account best for the dances in the framework of synchronous MP. Future research should focus on the replication of these findings as well as its application in perception research, computer graphics and robotics.

References

- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52(3), 317-332.
- Arbib, M. A. (1992). Schema theory. *The Encyclopedia of Artificial Intelligence*, 2, 1427-1443.
- Badler, N. I., Phillips, C. B., & Webber, B. L. (1993). Simulating humans: computer graphics animation and control. *Oxford University Press*.
- Ballard, D. H. (2015). *Brain computation as hierarchical abstraction*. MIT Press.
- Bernstein, N. (1967) *The coordination and regulation of Movement* (Pergamon, Oxford, 1967)
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Bizzi, E., Cheung, V. C. K., d'Avella, A., Saltiel, P., & Tresch, M. (2008). Combining modules for movement. *Brain research reviews*, 57(1), 125-133.
- Brown, S., Martinez, M. J., & Parsons, L. M. (2005). The neural basis of human dance. *Cerebral cortex*, 16(8), 1157-1167.
- Cheung, V. C., d'Avella, A., Tresch, M. C., & Bizzi, E. (2005). Central and sensory contributions to the activation and organization of muscle synergies during natural motor behaviors. *Journal of Neuroscience*, 25(27), 6419-6434.
- Chiovetto, E., & Giese, M. A. (2013). Kinematics of the coordination of pointing during locomotion. *PloS one*, 8(11), e79555.
- Chiovetto, E., d'Avella, A., & Giese, M. (2016). A Unifying Framework for the Identification of Motor Primitives. *Cornell University Library, quantitative biology. Bibliographic Code: 2016arXiv160306879C*
- Clever, D., Harant, M., Koch, H., Mombaur, K., & Endres, D. (2016). A novel approach for the generation of complex humanoid walking sequences based on a combination of optimal control and learning of movement primitives. *Robotics and Autonomous Systems*, 83, 287-298.
- Clever, D., Harant, M., Mombaur, K., Naveau, M., Stasse, O., & Endres, D. (2017). Cocomopl: A novel approach for humanoid walking generation combining optimal control, movement primitives and learning and its transfer to the real robot hrp-2. *IEEE Robotics and Automation Letters*, 2(2), 977-984.
- Danion, F., & Latash, M. L. (Eds.). (2011). *Motor control: theories, experiments, and applications*. Oxford University Press.
- d'Avella, A., & Tresch, M. C. (2002). Modularity in the motor system: decomposition of muscle patterns as combinations of time-varying synergies. *Advances in neural information processing systems*, 141-148.

- d'Avella, A., Saltiel, P., & Bizzi, E. (2003). Combinations of muscle synergies in the construction of a natural motor behavior. *Nature neuroscience*, 6(3), 300.
- Del Vecchio, D., Murray, R. M., & Perona, P. (2003). Decomposition of human motion into dynamics-based primitives with application to drawing tasks. *Automatica*, 39(12), 2085-2098.
- Ellis, D. P. (2007). Beat tracking by dynamic programming. *Journal of New Music Research*, 36(1), 51-60.
- Endres, D. M., Chiovetto, E., & Giese, M. A. (2013). Model selection for the extraction of movement primitives. *Frontiers in computational neuroscience*, 7.
- Endres, D., Chiovetto, E., & Giese, M. A. (2016). Bayesian Approaches for Learning of Primitive-Based Compact Representations of Complex Human Activities. In *Dance Notations and Robot Motion* (pp. 117-137). Springer International Publishing.
- Endres, D., Christensen, A., Omlor, L., & Giese, M. A. (2011). Emulating human observers with Bayesian binning: segmentation of action streams. *ACM Transactions on Applied Perception (TAP)*, 8(3), 16.
- Farnell, B. (1999). It goes without saying—but not always. In *Dance in the Field* (pp. 145-160). Palgrave Macmillan UK.
- Flash, T., & Hochner, B. (2005). Motor primitives in vertebrates and invertebrates. *Current opinion in neurobiology*, 15(6), 660-666.
- Flash, T., & Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *Journal of neuroscience*, 5(7), 1688-1703.
- Frigon, A. (2012). Central pattern generators of the mammalian spinal cord. *The Neuroscientist*, 18(1), 56-69.
- Good, I. J. (1985). Weight of evidence: A brief survey. In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, & A. F. M. Smith (Eds.), *Bayesian Statistics 2* (pp. 249–270). Elsevier.
- Graziano, M. S., Taylor, C. S., & Moore, T. (2002). Complex movements evoked by microstimulation of precentral cortex. *Neuron*, 34(5), 841-851.
- Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394(6695), 780.
- Hart, C. B., & Giszter, S. F. (2010). A neural basis for motor primitives in the spinal cord. *Journal of Neuroscience*, 30(4), 1322-1336.
- Hart, C. B., & Giszter, S. F. (2013). Distinguishing synchronous and time-varying synergies using point process interval statistics: motor primitives in frog and rat. *Frontiers in computational neuroscience*, 7.

- Ivanenko, Y. P., Cappellini, G., Dominici, N., Poppele, R. E., & Lacquaniti, F. (2005). Coordination of locomotion with voluntary movements in humans. *Journal of Neuroscience*, 25(31), 7238-7253.
- Janata, P., & Grafton, S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature neuroscience*, 6(7), 682.
- Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, 33(11), 1419-1432.
- Jenkins, O. C., & Mataric, M. J. (2003). Automated derivation of behavior vocabularies for autonomous humanoid motion. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems* (pp. 225-232). ACM.
- Jing, J., Cropper, E. C., Hurwitz, I., & Weiss, K. R. (2004). The construction of movement with behavior-specific and behavior-independent modules. *Journal of Neuroscience*, 24(28), 6315-6325.
- Johansson, G. (1975). Visual motion perception. *Scientific American*, 232(6), 76-89.
- Kaminski, T. R. (2007). The coupling between upper and lower extremity synergies during whole body reaching. *Gait & Posture*, 26(2), 256-262.
- Knopp, B., Schubert, M., Endres, D. (2017). Temporal Movement Primitive Perception under Naturalistic Conditions. *To be presented at SAP 2017, Cottbus*, doi: 10.475/123_4
- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Attention, Perception, & Psychophysics*, 21(6), 575-580.
- Kuha, J. (2004). AIC and BIC: Comparisons of Assumptions and Performance. *Sociological methods & research*, 33(2), 188-229.
- Laland, K., Wilkins, C., & Clayton, N. (2016). The evolution of dance. *Current Biology*, 26(1), R5-R9.
- Lavine, M., & Schervish, M. J. (1999). Bayes factors: what they are and what they are not. *The American Statistician*, 53(2), 119-122.
- Lee, W. A. (1984). Neuromotor synergies as a basis for coordinated intentional action. *Journal of motor behavior*, 16(2), 135-170.
- Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1), 210.
- Macpherson, J. M. (1991). How flexible are muscle synergies. *Motor control: concepts and issues*, 33-47.

- Marder, E., & Bucher, D. (2001). Central pattern generators and the control of rhythmic movements. *Current biology*, 11(23), R986-R996.
- McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015, 07). librosa: Audio and music signal analysis in python. In *Proceedings of the 14th python in science conference* (pp. 18-25).
- Moro, A. (2014). On the similarity between syntax and actions. *Trends in cognitive sciences*, 18(3), 109-110.
- MTw User Manual and Technical Documentation*; Xsens Technologies B.V.: Enschede, 2017. Available online: <http://www.xsens.com> (last accessed on 15 August 2017).
- Mussa-Ivaldi, F. A. (1999). Modular features of motor control and learning. *Current opinion in neurobiology*, 9(6), 713-717.
- Omlor, L., & Giese, M. (2007). Blind source separation for over-determined delayed mixtures. In *Advances in neural information processing systems* 1049-1056.
- Omlor, L., & Giese, M. A. (2006, June). Unsupervised learning of spatio-temporal primitives of emotional gait. In *International Tutorial and Research Workshop on Perception and Interactive Technologies for Speech-Based Systems* (pp. 188-192). Springer, Berlin, Heidelberg.
- Omlor, L., & Giese, M. A. (2007a). Extraction of spatio-temporal primitives of emotional body expressions. *Neurocomputing*, 70(10), 1938-1942.
- Overduin, S. A., d'Avella, A., Roh, J., Carmena, J. M., & Bizzi, E. (2015). Representation of muscle synergies in the primate brain. *Journal of Neuroscience*, 35(37), 12615-12624.
- Pinker, S. (2015). *Words and rules: The ingredients of language*. Basic Books.
- Prinz, W. (1990). A common coding approach to perception and action. In *Relationships between perception and action* (pp. 167-201). Springer Berlin Heidelberg.
- Rasmussen, C. E., & Williams, C. K. (2006). *Gaussian processes for machine learning* (Vol. 1). Cambridge: MIT press.
- Rosenbaum, D. A. (2009). *Human motor control*. Academic press.
- Schaal, S., Ijspeert, A., & Billard, A. (2003). Computational approaches to motor learning by imitation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 358(1431), 537-547.
- Schenau, G. J. V. I. (1989). From rotation to translation: Constraints on multi-joint movements and the unique action of bi-articular muscles. *Human Movement Science*, 8(4), 301-337.

- Schöner, G., & Kelso, J. S. (1988). Dynamic pattern generation in behavioral and neural systems. *Science*, 1513-1520.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), 461-464.
- Sherrington, C. S. (1910). Flexion-reflex of the limb, crossed extension-reflex, and reflex stepping and standing. *The Journal of physiology*, 40(1-2), 28-121.
- Stevenson, A. (Ed.). (2010). *Oxford dictionary of English*. Oxford University Press, USA.
- Sumbre, G., Gutfreund, Y., Fiorito, G., Flash, T., & Hochner, B. (2001). Control of octopus arm extension by a peripheral motor program. *Science*, 293(5536), 1845-1848.
- Thill, S., Hemeren, P. E., & Durán, B. (2011). Prediction of human action segmentation based on end-effector kinematics using linear models. In *European Perspectives on Cognitive Science: Proceedings of the European Conference on Cognitive Science 2011*. Sofia: NBU Press.
- Ting, L. H., & Macpherson, J. M. (2005). A limited set of muscle synergies for force control during a postural task. *Journal of neurophysiology*, 93(1), 609-613.
- Tresch, M. C., Saltiel, P., & Bizzi, E. (1999). The construction of movement by the spinal cord. *Nature neuroscience*, 2(2).
- Tresch, M. C., Saltiel, P., d'Avella, A., & Bizzi, E. (2002). Coordination and localization in spinal motor systems. *Brain Research Reviews*, 40(1), 66-79.
- Tresilian, J. (2012). *Sensorimotor control and learning: An introduction to the behavioral neuroscience of action*. Palgrave Macmillan.
- Vereijken, B., Emmerik, R. E. V., Whiting, H. T. A., & Newell, K. M. (1992). Free (z) ing degrees of freedom in skill acquisition. *Journal of motor behavior*, 24(1), 133-142.
- Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic bulletin & review*, 11(1), 192-196.
- Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of mathematical psychology*, 44(1), 92-107.
- Weakliem, D. L. (1999). A critique of the Bayesian information criterion for model selection. *Sociological Methods & Research*, 27(3), 359-397.
- Williams, B. H., Toussaint, M., & Storkey, A. J. (2006-09). Extracting motion primitives from natural handwriting data. In *International Conference on Artificial Neural Networks (pp. 634-643)*. Springer, Berlin, Heidelberg.
- Zucchini, W. (2000). An introduction to model selection. *Journal of mathematical psychology*, 44(1), 41-61.

III. APPENDIX. 1. Participant study information

Fachbereich Psychologie
Allgemeine und Biologische Psychologie
AE Theoretische Neurowissenschaft
Prof. Dr. D. Endres
Gutenbergstraße 18
35032 Marburg
Telefon: 06421-2823818
dominik.endres@uni-marburg.de



Teilnehmerinformation zur Studie

Produktion modularer Bewegungsprimitive beim Tanzen

Ziele der Studie

Das zentrale Nervensystem des Menschen ist imstande, mit anscheinender Leichtigkeit eine Vielzahl komplexer Bewegungen zu erzeugen und diese flexibel an unterschiedlichste Umweltbedingungen anzupassen.

Einer Vermutung zufolge entstehen diese flexiblen und teilweise sehr komplizierten Bewegungen durch die Kombination einfacher Bewegungseinheiten, den sogenannten Bewegungsprimitiven.

Wir möchten herausfinden, ob wir diese Bewegungsprimitive auch bei komplexen und kontinuierlich ablaufenden Bewegungsprozessen finden können, die beim Tanzen vorkommen.

Ablauf der Bewegungsaufzeichnung

Um diese Fragestellung zu beantworten, möchten wir die Ganzkörperbewegungen von Versuchspersonen aufzeichnen, die tanzen. Die so gewonnenen Daten werden dann im Hinblick auf ihre theoretische Passung zu den oben genannten Bewegungsprimitiven geprüft.

Die Aufzeichnung geschieht mit einem „XSENS“ System. Um Ihre Bewegungen aufzuzeichnen, werden Sie gebeten, zunächst ein Shirt, einen Gürtel und Handschuhe anzuziehen, welche die Anbringung der Sensoren ermöglichen. Außerdem werden an elastischen Bändern, die Sie mit Hilfe der Versuchsleiterin an Armen und Beinen befestigen, weitere Sensoren angebracht werden. Die Versuchsleiterin wird Sie dann nach der Kalibrierung darum bitten, möglichst frei und improvisiert zur Musik in einem gewissen Bereich des Labors zu tanzen. Hierbei gibt es keine weiteren Einschränkungen. Ihre Bewegungen werden für die an der Studie beteiligten Personen im Nachhinein lediglich abstrahiert (als eine Art Strichmännchen bzw. Punktwolke) zu sehen sein – es werden

ausdrücklich *keine* Videoaufzeichnungen gemacht!

Der gesamte Versuch dauert in etwa 1.5 Stunden. Dies beinhaltet auch die Vorbereitung der Messung und eine kurze Nachbefragung. Nach der ersten Hälfte der Lieder wird Ihnen angeboten, eine Pause zu machen, um zum Beispiel etwas zu trinken. Sie können aber auch immer, wenn Ihnen danach ist, um eine Pause bitten. Dies ist notwendig und hat keinen nachteiligen Effekt auf die Messungen. Sie können hierfür mit Versuchspersonenstunden oder einer finanziellen Vergütung von 8€ pro Stunde entlohnt werden.

Freiwilligkeit der Teilnahme

Sie können die Datenaufnahme jederzeit und ohne Nennung von Gründen abbrechen und die Löschung der bis dahin aufgezeichneten Daten verlangen, ohne dass Ihnen hierdurch Nachteile entstehen. Für bereits erbrachte Leistungen werden Sie dann anteilig vergütet.

Datenschutz

Zusätzlich zu Ihren Bewegungsdaten erheben wir einige persönliche Daten (Alter, Geschlecht, Gewicht), die jedoch ohne Verbindung zu Ihrem Namen gespeichert werden, so dass nach der Bewegungsaufzeichnung die Daten vollständig anonymisiert vorliegen. Das bedeutet auch, dass eine Löschung der Daten nach Ende der Aufzeichnung nicht möglich ist. Die aufgezeichneten Bewegungsdaten, die Ihrer Person nicht zugeordnet werden können, werden ausschließlich auf den passwortgeschützten Rechnern des Instituts gespeichert. Zugriff hierauf haben ausschließlich an der Untersuchung beteiligte WissenschaftlerInnen. Im Rahmen gesetzlicher Bestimmungen werden Ihre Daten 10 Jahre lang aufbewahrt und anschließend gelöscht.

Schließlich weisen wir darauf hin, dass Ihre Rohdaten im Rahmen von wissenschaftlichen Publikationen veröffentlicht werden könnten. Auch hierbei wird selbstverständlich kein Rückschluss auf Ihre Person möglich sein.

Voraussetzungen für die Teilnahme

- ✓ Sie sind mindestens 18 Jahre alt
- ✓ Sie leiden an keiner Erkrankung des Bewegungsapparates
- ✓ Sie sind nicht schwerhörig
- ✓ Sie sind normalsichtig, oder haben auf Normalsichtigkeit korrigiertes Sehvermögen

An wen können Sie sich bei Fragen wenden?

Für Fragen oder Anmerkungen steht Ihnen unser Studienteam gern telefonisch unter der 06421 / 2823818 oder per Mail unter dominik.endres@uni-marburg.de zur Verfügung. Wenden Sie sich auch gerne an Ihre Versuchsleiterin unter eckerta@students.uni-marburg.de

Wir würden uns freuen, Sie als TeilnehmerIn begrüßen zu dürfen!

Dominik Endres

Anna-Lena Eckert

III. APPENDIX: 2. Informed Consent Form

Einverständniserklärung

Produktion modularer Bewegungsprimitive beim Tanzen

Ich, Frau/ Herr _____

bin über das Ziel und Ablauf der Bewegungsaufzeichnungen für die Studie “Produktion modularer Bewegungsprimitive beim Tanzen“ informiert worden. Ich habe alle Informationen vollständig gelesen und verstanden. Sofern ich Fragen zu dieser Studie hatte, wurden sie vom Studienleiter vollständig und zu meiner Zufriedenheit beantwortet.

Ich hatte genügend Zeit für eine Entscheidung und bin bereit, an der o.g. Studie teilzunehmen. Ich weiß, dass die Teilnahme an der Studie freiwillig ist und ich die Teilnahme jederzeit ohne Angaben von Gründen beenden kann. Im Falle des Widerrufs entstehen mir keine Kosten, Verpflichtungen oder andere Nachteile.

Ich bin dazu bereit, Bewegungssensoren an meinem Körper anzubringen bzw. anbringen zu lassen und im Labor unter geschützten Bedingungen zu tanzen. Die Studie wird, inklusive Kalibration, Anbringung der Bewegungssensoren und einer kurzen Nachbefragung insgesamt ca. 1.5 Stunden dauern.

Mir ist bekannt, dass die erhobenen Daten ohne Nennung meines Namens vollständig anonymisiert gespeichert werden. Eine Löschung meiner Daten nach Ende der Bewegungsaufzeichnung ist deshalb nicht möglich. Ich bin einverstanden, dass meine anonymisierten Daten zu Forschungszwecken weiter verwendet werden können und mindestens 10 Jahre gespeichert werden. Ich bin einverstanden, dass die Rohdaten, die in dieser Erhebung entstanden sind, in anonymisierter Form als Teil einer Publikation eingereicht werden könnten.

Eine Ausfertigung der Teilnehmerinformation über die Untersuchung habe ich erhalten. Die Teilnehmerinformation ist Teil dieser Einwilligungserklärung.

_____ (Unterschrift
Teilnehmer)

Ich bin daran interessiert, etwas über die Ergebnisse der Studie zu erfahren, und bitte hierzu um Übersendung entsprechender Informationen. Zu diesem Zweck gebe ich meine Email-Adresse an (optional)

E-Mail: _____

Bei Fragen oder Anliegen kann ich mich an die folgenden Personen wenden:

Versuchsleiterin	Projektleiter
Anna-Lena Eckert Gutenbergstraße 18 35032 Marburg eckerta@students.uni-marburg.de	Prof. Dr. Dominik Endres Gutenbergstraße 18 35032 Marburg dominik.endres@uni-marburg.de

Als Versuchsleiterin verpflichte ich mich hiermit, die Bedingungen zur Teilnahme an der Studie sowie zum Datenschutz einzuhalten.

Unterschrift Versuchsleiterin

III. APPENDIX: 3. Post-questionnaire

VP Nr. _____

Alter: _____

Geschlecht: _____

Körpergewicht: _____

Wie wohl haben Sie sich in der Studiensituation gefühlt?

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sehr unwohl	Unwohl	Eher unwohl	neutral	Eher gut	gut	Sehr gut

In einer durchschnittlichen Woche, an wie vielen Tagen tanzen Sie? (allein zu Hause, mit Freunden, ausgehen...)

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nie (0)	Selten (0-1)	Manchmal (2-3)	Oft (6-7)	Täglich (7)

Wie gut hat Ihnen die Musik im Durchschnitt gefallen?

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Extrem schlecht	schlecht	Eher schlecht	Neutral	Eher gut	gut	Extrem gut

Bitte bewerten Sie, wie gut Sie zu jedem der Lieder tanzen konnten.

1 = sehr schlecht, 2 = eher schlecht, 3= geht so/ Neutral, 4= eher gut, 5= sehr gut.

Bitte fragen Sie die Versuchsleiterin, wenn Sie das Lied nicht mit Namen kennen!

Lied	Einschätzung 1-5
Move on Up – Curtis Mayfield	
Give it up to me – Sean Paul	
Bonfire – Felix Jaehn & Alma	
Could You Be Loved – Bob Marley	
Dickes B - Seeed	
She drives me crazy – Fine Young Cannibals	
You shook me all night long - ACDC	

Inwiefern treffen diese Beschreibungen auf Ihre Stimmung *vor* dem Tanzen zu?

	Gar nicht	Eher nicht	Neutral	Eher schon	Total
Ängstlich					
Traurig					
Aufgeregt					
Lustlos					
Müde					
Motiviert					

Haben Sie Erfahrungen als „professionelle/r“ TänzerIn?*z.B. Tanzkurse oder Turniere, Freizeitkurse, Sportkurse, Erfahrungen als TanzlehrerIn...*☐ Ja☐ Nein

Anmerkungen: Art, Dauer,...

Verglichen mit einer durchschnittlichen Person in Ihrem Alter und Ihres Geschlechts, wie schätzen Sie Ihre Fähigkeiten als TänzerIn ein?

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schrecklich	Sehr schlecht	schlecht	Genauso (Durchschnitt)	gut	Sehr gut	Hervorragend

Hat Ihnen während des Tanzens etwas weh getan?

Platz für weitere Anmerkungen:

Table 2: Descriptive results post-questionnaire

Item	Minimum	Maximum	Mean	Std. Deviation
Wellbeing in study situation	4	7	6.00	.866
Dance frequency	2	5	2.76	.903
Liked music on average	3	7	5.06	1.249
How danceable was...				
1. 'Move On Up'	2	5	3.53	1.068
2. 'Give It Up To Me'	2	5	3.94	.899
3. 'Bonfire'	2	5	3.65	.996
4. 'Could You Be Loved'	2	5	3.29	.996
5. 'Dickes B'	2	5	3.35	1.272
6. 'She Drives Me Crazy'	2	5	3.53	.874
7. 'You Shook Me...'	1	5	3.12	1.219
Mood				
○ Scared	1	4	1.94	.748
○ Sad	1	4	1.41	.870
○ Excited	1	4	3.24	1.033
○ Lethargic	1	4	1.65	.931
○ Tired	1	5	2.59	1.278
○ Motivated	2	5	4.18	.809
Social comparison dance skills	2	6	4.35	1.057
Professional experiences	0	1	5=yes 12=no	-

Annotations. The items were coded as follows: Wellbeing in study situation from 1 (=very uncomfortable) to 7 (=very comfortable); Dance frequency in a normal week from 1 (=Never) to 5 (=every day), average evaluation of the music from 1 (=extremely bad) to 7 (=extremely good), evaluation of how danceable a specific song is (how well one could dance to it) from 1 (=very bad) to 5 (=very good); mood during study participation (scared, sad, excited, lethargic, tired, motivated) from 1 (=not at all) to 5 (=totally); professional experiences with dance, including courses, experiences as a teacher, etc. either yes (1) or no (0), social evaluation of one's own dancing skills ('compared to an average person of your gender and your age') from 1 (=horrible) to 7 (=excellent).

Two-tailed t-tests for independent samples were performed for the identification of gender differences, showing a significant difference between male and female participants for the 'social comparison' item: male participants evaluated their dance abilities significantly higher than female participants (mean difference: -1.22, $t=-2.866$, $df=15$, $p=.011$)

Table 3. Full list of stimuli.

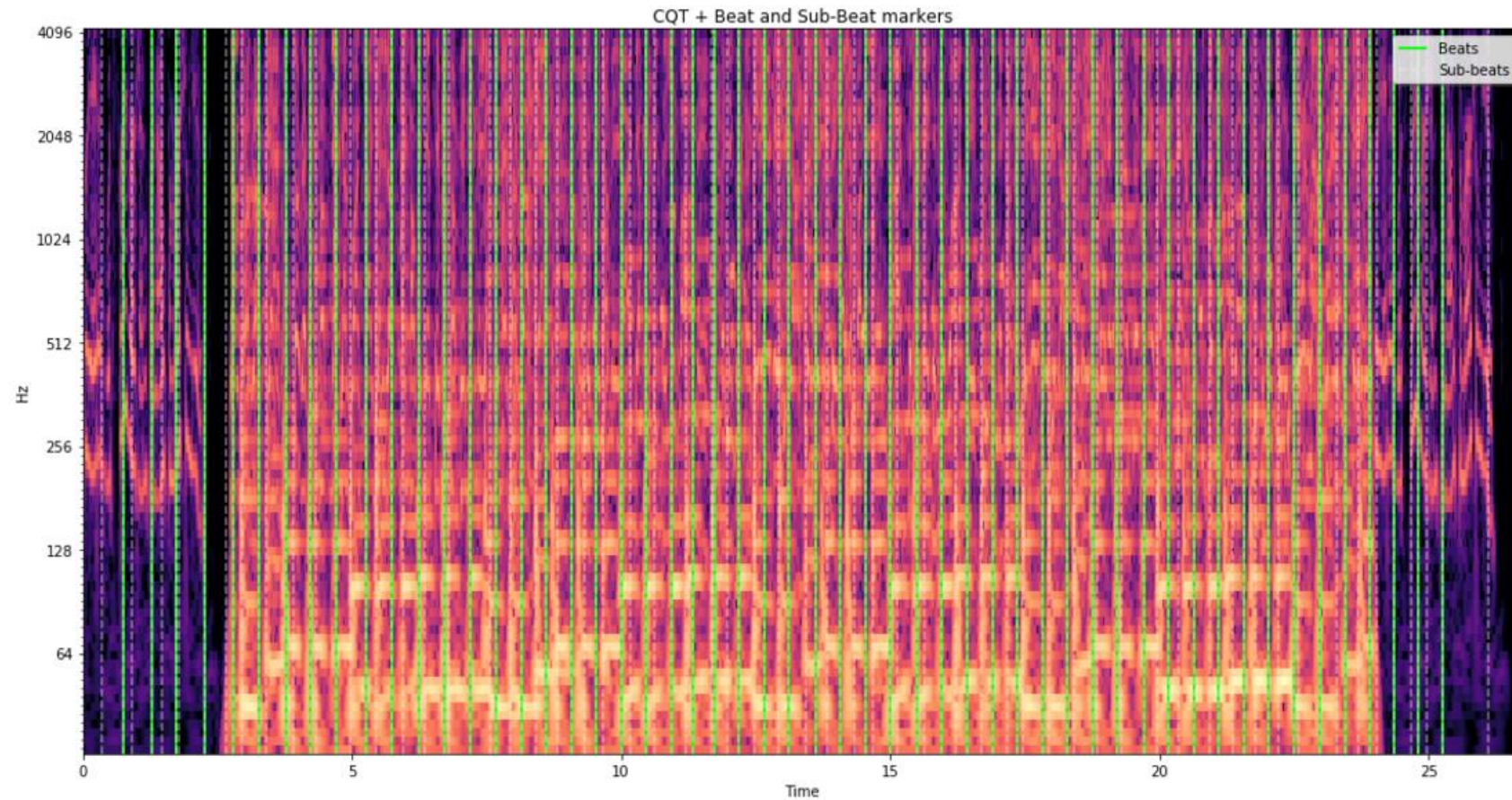
Song	Genre⁽¹⁾	Artist, year⁽¹⁾	duration	bpm⁽²⁾	<i>I</i>⁽⁴⁾
‘(When you gonna) Give It Up To Me’	Dancehall	Sean Paul feat. Keysha Cole	4:04	96.02	124
‘Dickes B’	Hip-Hop	Seed	4:00	98.00	92
‘Could You Be Loved’	Reggae	Bob Marley,	3:57	103.31	99
‘She drives me crazy’	Pop	Fine Young Cannibals, 1989	3:35	108.81	95
‘Bonfire’	House	Felix Jaehn feat. Alma	3:02	111.02	79
‘You shook me all night long’	Rock	AC/DC,	3:30	128.31	109
‘Move On Up’	Funk	Curtis Mayfield	2:46	138.31	94
			Total: 24:54	mean ⁽³⁾ : 111.96	mean ⁽⁵⁾ : 98,85

Annotations: All songs were presented in a random order using the shuffle function of the media player. The songs were selected so that they represent a wide range of musical genres and number of beats per minute. ⁽¹⁾ The genre of the song and its year of appearance as indicated in the album description. ⁽²⁾ Beats per minute (bpm) were detected using the python library ‘librosa’ for music analysis (function: `librosa.beat.beat_track` Ellis, 2007) ⁽³⁾ Mean-bpm=111,96; SD=15,79. ⁽⁴⁾ is *I*, or the number of trials that are associated with the song. The detected beats were used for the purpose of presegmenting the data into ‘trials’ (*I*) which is necessary for decomposition purposes. ⁽⁵⁾ Mean number of trials across all songs was M=98,85, SD=14,23.

Table 4. Overview: Prevalent conceptualizations of movement primitives (MP) and underlying generative models

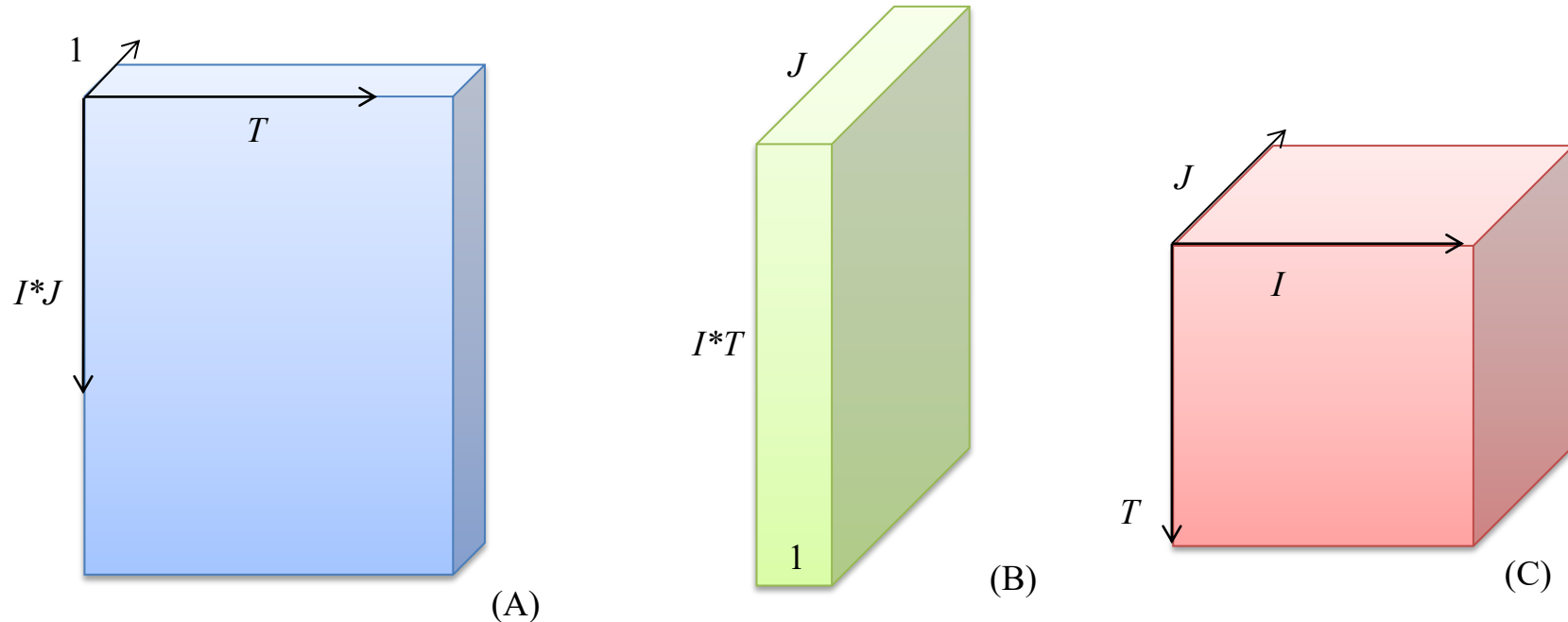
<i>Synchronous MP</i> (Ivanenko et al., 2005)	$x_{i,j}(t) = \sum_{s=1}^S w_{(i,j),s} MP_s(t)$	<ul style="list-style-type: none"> • $x_{i,j}(t)$ is the signal for joint j in trial i at time point t. s number of extracted temporal components $MP_s(t)$ are weighted with the coefficient $w_{(i,j),s}$; the indices of w and MP imply that the MP is temporally stereotypic and can be weighted so that it affects different body parts. • Ex.: Grasping and kicking: equal temporal processes at different effector locations.
<i>Spatial MP</i> (Tresch, Saltiel & Bizzi, 1999)	$x_{i,j}(t) = \sum_{s=1}^S w_{i,s}(t) MP_{s,j}$	<ul style="list-style-type: none"> • $x_{i,j}(t)$ is the recorded signal that can be decomposed into the sum of joint-independet weights for time point t * the spatially invariant pattern $MP_{s,j}$ • Ex.: arm-wrestling and lever-pushing: the same muscular features for pushing and balancing are observed, varying only over time.
<i>Time-varying synergies model</i> (d'Avella & Tresch, 2002)	$x_{i,j}(t) = \sum_{s=1}^S w_{i,s} MP_{s,j}(t - \tau_{i,j})$	<ul style="list-style-type: none"> • The signal $x_{i,j}(t)$ can be decomposed into a weight w and the df-dependet MP that is executed with potential timely delays $(t - \tau_{i,j})$ • An episode is characterized by coordinated muscle activation amplitude and coordinated activation timings • Extraction originally with NNMF method
<i>Anechoic Mixture Model</i> (Omlor & Giese, 2007)	$x_{i,j}(t) = \sum_{s=1}^S w_{(i,j),s} MP_s(t - \tau_{i,j})$	<ul style="list-style-type: none"> • MP with temporal delays τ_{ij} approximate the signal $x_{i,j}(t)$ after weighing with a weighting coefficient $w_{(i,j),s}$ which is bound to a specific joint j. • Requires no a priori assumptions on the size of τ_{ij}

Figure 6. An illustration of the beat-tracking system.



Annotations. For illustration purposes, the results of the beat-tracking function `librosa.beat.track_beat` performing on a 25 seconds long music segment is depicted. Beats are represented with green vertical lines; sub-beats are marked with green dashed lines. The colors represent the presence of a specific wavelength (in Hz) in the music signal, while bright yellow stands for a strong proportion of this specific wavelength and dark purple stands for a small relative presence of the wavelength. From this representation of the musical signal, the beat tracking system derives the onset strength that is in a consequent step used for the identification of potential beat events in the audio signal.

Figure 7. Illustration for the hypothetical data matrices underlying the MP conceptualizations.



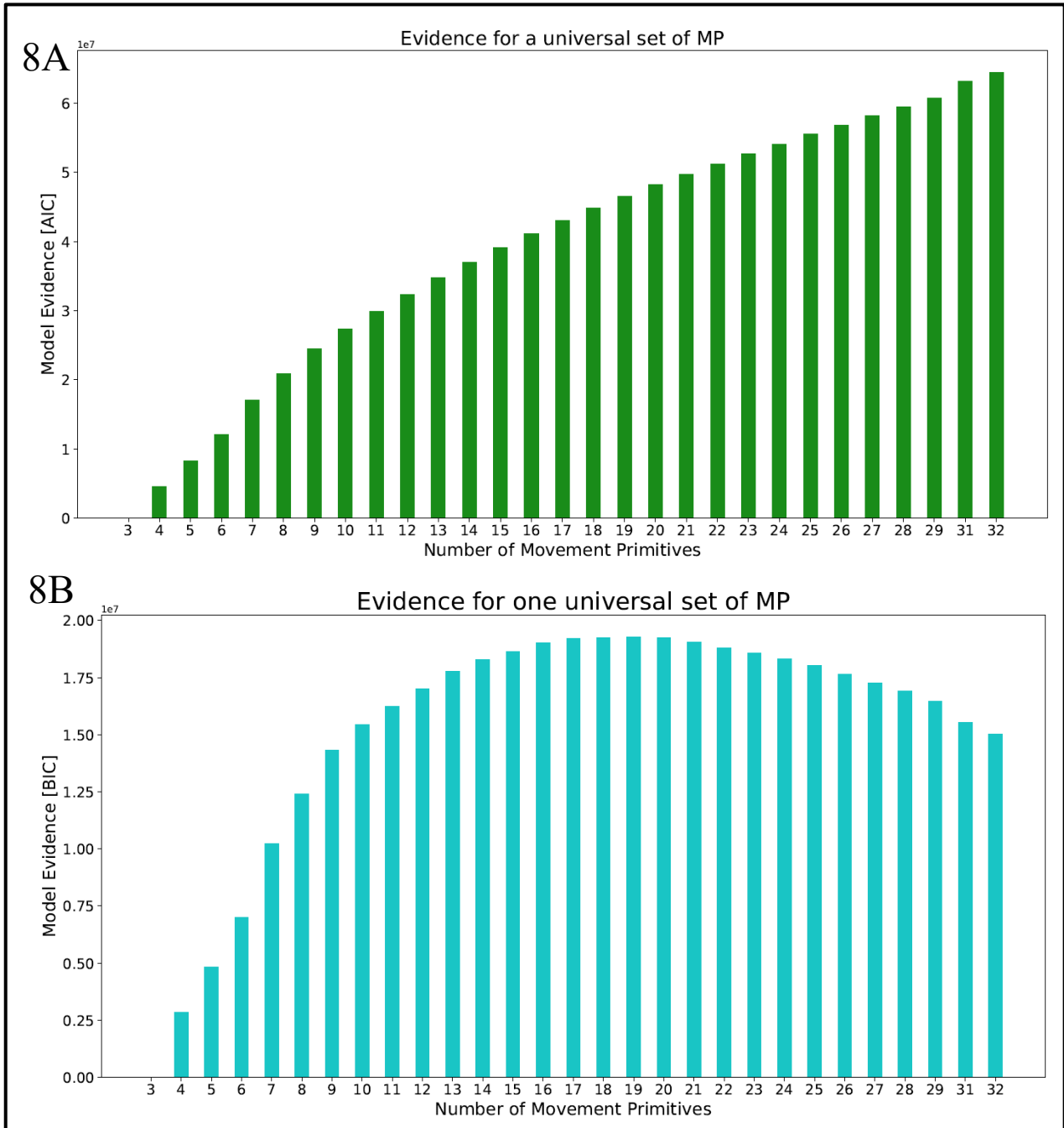
Annotations. For illustration purposes, the arrangement of the data that underlies the different MP conceptualizations is depicted. I are all trials, which in the present case are determined by the number of beats in the song (range: 92-124, $M=98.85$; $SD=14,23$). J are the 66 joints. T are the time steps that, for the purpose of later decomposition, are normalized to 100 (using Gaussian processes).

(A). The cube represents the arrangement of the data underlying ‘synchronous MP’. It postulates that the movements that belong to one MP can be expressed as a movement with temporal stereotypic features. This stereotypical time course can affect various effectors; e.g. a grasping motion and a kicking motion with the same temporal features would be considered belonging to one MP in this model. For the data analysis, the data is required to be of the shape $(I*J), T$. The 1 in panel (A) reflects that one synchronous MP is independent of all other synchronous MP. One row in the matrix therefore represents one temporally stereotypical movement.

(B) In the concept of ‘Spatial MP’, it is assumed that movements that belong to one MP are defined by the synergetic coordination of joints or muscles; therefore, this concept works with a $(I*T), J$ data matrix. This implies a stereotypical spatial coactivation is varied in the time-dimension. An example is the postural control when expecting impedances or armwrestling vs. lever-pushing (see text)

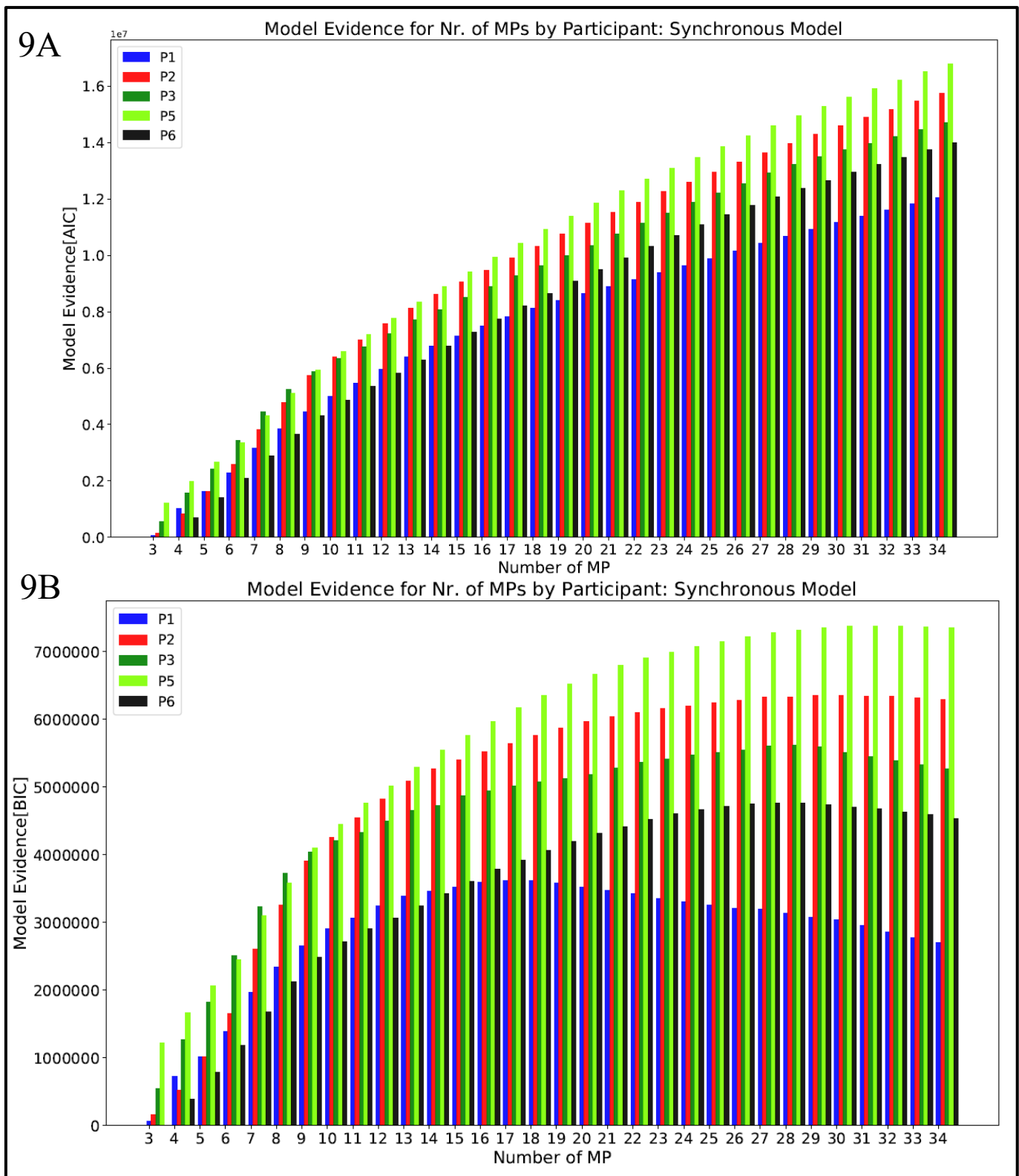
(C) Time-varying synergies allow for the analysis of the data without assuming independences of the MP from temporal or spatial features. One time-varying synergy is not only characterized by a coordinated amplitude weighting, but also by coordinated activation timings.

Figure 8. AIC and BIC scores for the ‘universal set of MP’ model (synchronous MP)



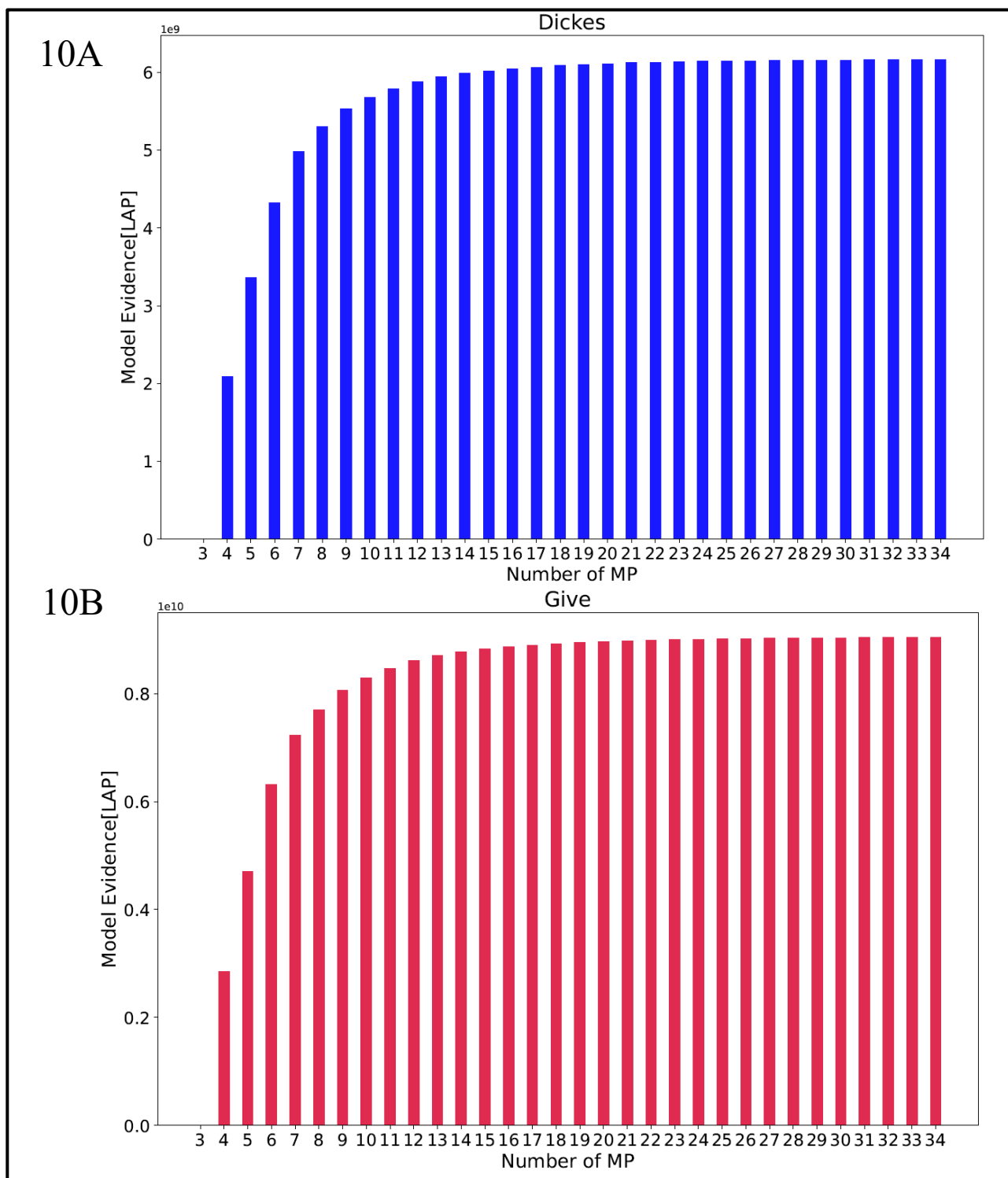
Annotations. AIC and BI scores for the ‘universal MP set’ model. *Figure 8A* shows that the best AIC score is achieved for a synchronous model that includes 32 components. As it still increases when approaching 32 MP, it can be assumed that the AIC would further increase for models with >32 MP. Following *Figure 8B*, nevertheless, BIC score is maximized for a model with 19 components. BIC scores hence would prefer a model with 19 components whereas AIC (and LAP, *Figure 1*) prompt the selection of a model with more than 32+ MP in the framework of synchronous MP.

Figure 9. AIC and BIC scores for the ‘individual repertoire’ model (synchronous MP).



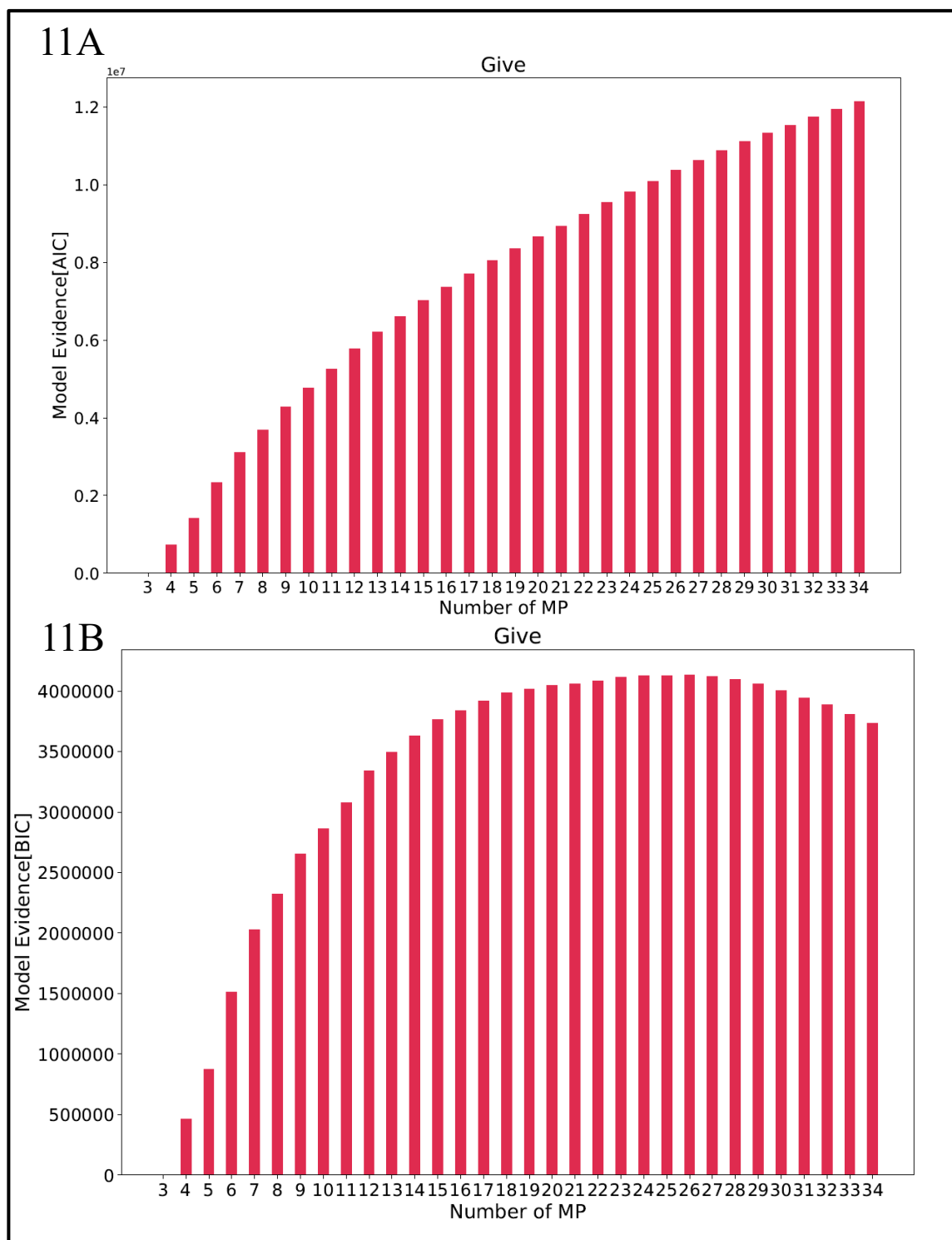
Annotations. AIC and BIC scores for the ‘individual repertoire’ hypothesis (nr. 2). *Figure 9A* AIC scores for the individual repertoire model are maximized for a model with 34 components. As scores are still increasing when approaching 34 MP, the maximum AIC score might be reached with 34+ MP. *Figure 9B* shows the BIC score for participant-determined model. Differential courses, depending on the participant, are observable. For example, BIC score for participant 1 is maximized for a synchronous model with 17 MP, whereas for participant 2, (P2) a model with 29 components scores highest.

Figure 10. LAP scores for other stimuli, testing the ‘stimulus determined’ model (hyp. 3)



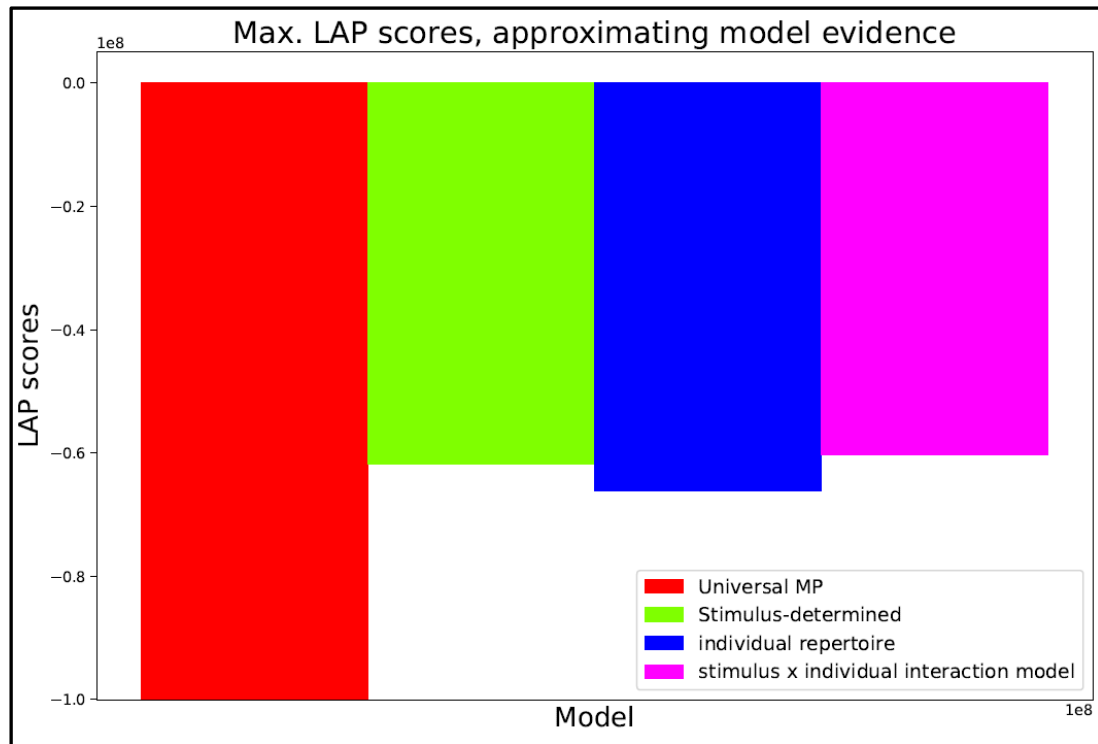
Annotations. LAP scores for the song ‘dickes’ and for the song ‘give’. LAP scores for all songs that were tested for the stimulus-determined model were highest for models with 34MP. All other song-specific LAP-plots can be found in the supplementary material on the disk.

Figure 11. AIC and BIC scores for the stimulus-determined model (hyp. 3)



Annotations. AIC (11A) and BIC (11B) scores for the song 'Give' for the test of a 'stimulus-determined' model (hyp. 3). AIC is maximized when approaching the most complex model with 34 MP, while BIC is maximized for a model with 26 components and the score decreases when considering more complex models. All other song-specific plots can be found in the supplementary material on the disk.

Figure 12. Maximum LAP scores (LAP_{max}) for each of the four models under consideration.



Annotations. Unrelativized maximum LAP scores for every model; a score near zero indicates a good model fit. The lowest score was observed for a model that assumes one universal set of MP for all participants across all stimulus conditions. The second lowest LAP score is observed for the model where an individual repertoire per participant is assumed. The highest LAP score was observed for an interaction model between stimulus and individual, followed by a model in which dancing is assumed to be stimulus-determined.

Eckert, Anna-Lena
Name / Vorname

Einverständniserklärung

Ich erkläre mich damit einverstanden, daß die vorliegende
Arbeit:

'Dance like nobody is watching you'
The production of movement primitives in ...
(Kurztitel)

in Bibliotheken allgemein zugänglich gemacht wird.
Dazu gehört, daß sie

- von der Bibliothek der Einrichtung, in der ich meine Arbeit angefertigt habe, zur Benutzung in ihren Räumen bereit gehalten wird;
- in konventionellen und maschinenlesbaren Katalogen, Verzeichnissen und Datenbanken verzeichnet wird;
- der UB für die lokale Benutzung und für Fernleihzwecke zur Verfügung steht;
- im Rahmen der urheberrechtlichen Bestimmungen für Kopierzwecke genutzt werden kann.

Marburg/L., den 29.05.17



(Unterschrift der Autorin/
des Autors)



(Unterschrift der Betreuerin /
des Betreuers der Arbeit)

Hiermit versichere ich/versichern wir, die vorliegende Arbeit selbstständig, ohne fremde, unerlaubte Hilfe und mit keinen anderen als den angegebenen Hilfsmitteln verfasst zu haben.

Marburg, 29.09.17

Unterschrift/Unterschriften des/der Studierenden:

Anna Lea Eick

Hiermit bestätige ich, die vorliegende Arbeit betreut und angenommen zu haben.

Marburg, 29.9.17

Unterschrift der Leiterin/des Leiters:

O. D. B.