Chimpanzee play sequences are structured hierarchically as games

Alexander Mielke[1,2] and Susana Carvalho[1,3]

[1] Primate Models for Behavioural Evolution Lab, School of Anthropology and Museum Ethnography, University of Oxford, Oxford, United Kingdom
[2] School of Psychology and Neuroscience, University of St Andrews, St Andrews, United Kingdom
[3] Interdisciplinary Centre for Archaeology and Evolution of Human Behaviour (ICArEHB), Universidade do Algarve, Faro, Portugal

ABSTRACT

Social play is ubiquitous in the development of many animal species and involves players adapting actions flexibly to their own previous actions and partner responses. Play differs from other behavioural contexts for which fine-scale analyses of action sequences are available, such as tool use and communication, in that its form is not defined by its function, making it potentially more unpredictable. In humans, play is often organised in games, where players know context-appropriate actions but string them together unpredictably. Here, we use the sequential nature of play elements to explore whether play elements in chimpanzees are structured hierarchically and follow predictable game-like patterns. Based on 5,711 play elements from 143 bouts, we extracted individual-level play sequences of 11 Western chimpanzees (*Pan troglodytes verus*) of different ages from the Bossou community. We detected transition probabilities between play elements that exceeded expected levels and show that play elements form hierarchically clustered and interchangeable groups, indicative of at least six games that can be identified from transition networks, some with different roles for different players. We also show that increased information about preceding play elements improved predictability of subsequent elements, further indicating that play elements are not strung together randomly but that flexible action rules underlie their usage. Thus, chimpanzee play is hierarchically structured in short games which limit acceptable play elements and allow players to predict and adapt to partners' actions. This "grammar of action" approach to social interactions can be valuable in understanding cognitive and communicative abilities within and across species.

INTRODUCTION

Animal lives take place in time—actions happen sequentially in response to changing environments and the behaviour of other individuals. Particularly in social interactions, each action is a decision based on the social context, the actor's previous actions, the partner's reactions, and intended outcomes. Sequential social decisions are therefore an important window into the complexity of animal decision-making abilities (*Gygax, Zeeland & Christina, 2021*). Sequences can be considered complex for participants and bystanders if contingencies between actions are increasingly removed in time or with increasing

Corresponding author
Alexander Mielke,
mielke.alexand@gmail.com,
am261@st-andrews.ac.uk

flexibility and decreased determinism of transitions between actions (*Kershenbaum et al., 2016*). In many animal species, social play involves rapid exchanges of actions between several participants that often appear random to observers, making it one of the most complex social contexts individuals are regularly involved in. Multiple individuals combine distinguishable and discrete actions ('play elements') in temporal patterns ('play sequences') and adapt to partners' actions. This complexity allows us to unravel fast-paced decision-making in sequential exchanges between players. Indices have been developed to quantify how asymmetric play behaviour is (*Bauer & Smuts, 2007*) and how complex bouts are (by quantifying the number of distinct play elements performed in about; *Cordoni & Palagi, 2011*). However, we currently lack a framework to understand the predictability of sequences in this context.

Social play behaviour, at least during some parts of development, is common in most mammals and birds (*Diamond & Bond, 2003*; *Fagen, 1981*). Species can have large repertoires of distinct elements that are exchanged between players (*Petru et al., 2009*). Play signals are used to prevent play from breaking down when intentions are unclear or risk is high (*Cordoni & Palagi, 2012*), following specific rules determined by the social context (*Flack, Jeannotte & De Waal, 2004*), and extending the length of play bouts (*Waller & Dunbar, 2005*). We have yet to learn how coordinated other play actions are. In human play, there are specific, socially learned arbitrary rules that govern what we call 'games' (*Suits, 1967*): playful encounters in which a certain 'state' has to be achieved using only actions permitted by a set of rules that are more restrictive than necessary. In a game, certain actions are allowed, but their order can be flexible, while rules at the same time limit permittable actions, even excluding actions that would be mechanically useful in this situation (*Suits, 1967*). However, there is no prescribed outcome. Games therefore achieve a 'generative balance between the open-endedness of contingencies and the reproducibility of conditions for action' (*Malaby, 2007*). For example, in hide-and-seek, hiding is allowed, and where and how to hide is up to the player; leaving the play area or hiding for too long might be counter-productive despite achieving the intended goal of not being found. There is evidence that apes have standardised games and play them with each other and human partners (*Costa et al., 2019*; *Pika & Zuberbühler, 2008*; *Tanner & Byrne, 2010*). However, these examples focus on special contexts (*e.g.*, playing in water, playing socially with objects), and we do not have a method to determine how widespread predictable behavioural rules are.

Play is a prime candidate to understand sequential actions in animals because its form is not necessarily the results of a specific function—play has been hypothesised to have evolved as practice for future challenges facing individuals, so it is defined by its unpredictability compared to 'real' interactions (*Fagen, 1981*; *Palagi et al., 2004*; *Smith, 1982*). In other domains, such as tool use and communication, the form of sequences is defined partially by their function: in tool use, an 'optimal' sequence exists to solve a problem (*Estienne, Stephens & Boesch, 2017*). In communication, complexity is limited by the need to be understood. Songs are not constrained the same way, often containing hundreds of hierarchically structured elements (*Berwick et al., 2011*). Given that most species do not create song-like vocalisations, understanding sequences in social interactions

would allow for a broader perspective on sequence production and comprehension. There is ample evidence in humans that the cognitive processes underlying sequence generation and comprehension in different domains, such as language, music, tool use, and manual actions, are linked by an underlying 'grammar of action' (*Fitch & Martins, 2014*; *Pastra & Aloimonos, 2012*; *Stout et al., 2018*). Thus, we can use complex hierarchical sequences outside of the communicative context to understand the ability of animals to generate nested and embedded sequences necessary for human-like syntax (*Steele, Ferrari & Fogassi, 2012*)—with play and its long and 'random' exchanges a strong contender.

Predictable, hierarchical sequential decision-making processes in animals have been investigated in some detail in tool use and communication. For example, New Caledonian crows (*Hunt & Gray, 2004*; *Wimpenny et al., 2009*) as well as several primate species (*Boesch et al., 2020*; *Carvalho et al., 2008*; *Deblauwe et al., 2006*; *Estienne, Stephens & Boesch, 2017*; *Hihara et al., 2003*; *Martin-Ordas, Schumacher & Call, 2012*) use sequences of steps, often involving multiple objects, to solve problems using tools and tool sets. In chimpanzees, stone tool use (*Carvalho et al., 2008*; *Sirianni, Mundry & Boesch, 2015*), termite fishing (*Deblauwe et al., 2006*), and digging for underground bee nests (*Estienne, Stephens & Boesch, 2017*) have been analysed as complex decision sequences. Complex tool use has been linked repeatedly to the evolution of human language and syntax (*Steele, Ferrari & Fogassi, 2012*). Similarly, vocal patterns of bats (*Bohn et al., 2009*), birds (*Engesser, Ridley & Townsend, 2016*; *Sasahara et al., 2012*), cetaceans (*Allen et al., 2019*), rock hyraxes (*Kershenbaum et al., 2012*), and primates (*Arcadi, 1996*; *Girard-Buttoz et al., 2022*; *Leroux et al., 2021*; *Ouattara, Lemasson & Zuberbühler, 2009*) have been described as temporal sequences with different degrees of predictability, combinatorial complexity, and hierarchical structure. This has often been related to the evolution of syntax (*Leroux & Townsend, 2020*). Increasingly, communication sequences are found for other communicative modalities, such as gestures and facial signals (*e.g.*, *Aychet, Blois-Heulin & Lemasson, 2021*; *Graham, Furuichi & Byrne, 2020*; *Liebal, Call & Tomasello, 2004*; *McCarthy, Jensvold & Fouts, 2013*). Different species show turn-taking in exchanges and adapt their signals as sequential response to a partner's actions (*Demartsev et al., 2018*; *Fröhlich, 2017*).

The Bossou Western chimpanzees (*Pan troglodytes verus*) have been studied since 1976 (*Matsuzawa & Humle, 2011*). An 'outdoor laboratory' was created in 1988 as a clearing in the territory of the community where stones and nuts are provided to study tool use, with standardised video recordings available for over 30 years. Because the chimpanzees spend considerable time there, social and object play can be observed regularly (*Myowa-Yamakoshi & Yamakoshi, 2011*). In chimpanzees, infants and juveniles play more than older subadults and adult individuals (*Cordoni & Palagi, 2011*), but chimpanzees are among the species where adult play seems common (*Fernandez-Duque et al., 2000*) and fulfils several functions, especially in conflict regulation and stress reduction (*Palagi et al., 2004*). Chimpanzees play with and without objects (*Koops, Furuichi & Hashimoto, 2015*), and solitary and socially, often involving more than two players (*Cordoni et al., 2018*; *Shimada, 2013*). Play signals are used to advertise willingness initiate play bouts and increase their duration (*Davila Ross et al., 2009*; *Matsusaka, 2004*; *Waller & Dunbar, 2005*), and there is good evidence that chimpanzees show matching or mimicry of partners' play face and

laughter (*Davila-Ross et al., 2011*; *Ross, Bard & Matsuzawa, 2014*) and that play signals are adapted to the social context of the bout (*Flack, Jeannotte & De Waal, 2004*). Gestures can occur in sequences during play (*Bard et al., 2014*), especially if partners fail to respond initially, with tactile and audible gestures usually occurring early in the sequence (*McCarthy, Jensvold & Fouts, 2013*) and younger individuals producing more tactile gestures (*Fröhlich, Wittig & Pika, 2016*). The cooperative and coordinated nature of play (multiple individuals adapting their behaviour in real-time to sustain the interaction) has been used to study higher socio-cognitive skills such as joint intention and shared intentionality with varying results (*Bekoff & Allen, 1998*; *Pika & Zuberbühler, 2008*; *Tomasello et al., 2005*), and joint commitment and joint action (*Heesen et al., 2021a*; *Heesen et al., 2017*; *Heesen et al., 2021b*). Anecdotal evidence from the Bossou chimpanzees has repeatedly indicated that chimpanzee play might involve aspects of pretence or imagination (*Matsuzawa, 2020*; *Nakamura, 2012*). Our focus is on the form of play, how elements are strung together, which has its own implication for cognitive evolution.

For this study, we tested whether sequences of play elements are predictable for players or are largely random, and whether we can identify hierarchical structure in sequence patterns. To do this, we ask two main questions: if I know the previous action ('antecedent'), can I predict the subsequent action ('consequent')? And are there higher-order connections between elements, in the form of network clusters of interchangeable elements? This study specifically looks at transitions within individuals—partner behaviour is considered 'noise'. This will reduce predictability, because actions that appear 'unexpected' here are possibly expected responses to partner actions. We hypothesize that some play elements are consistently more likely to follow specific antecedents than would be expected at random. Using the probabilities of each element and each transition to 'predict' which element will appear next, we expected classification accuracy that exceeds random assignment, and that higher-order sequences (AB, rather than B alone, to predict C) further improved prediction accuracy. If we find these patterns, we would talk about a 'rule-based system', with the understanding that rules here are descriptive transition rules rather than normative rules that are enforced by partners (*Flack, Jeannotte & De Waal, 2004*), as we do not include information about partner behaviour at this point. We also hypothesized that, like communication in some species (*Allen et al., 2019*), we can detect hierarchical structures in transition networks ('games') as clusters of elements that are often used together and can be used interchangeably. We use this operationalisation to reflect the fluid nature of games: once we are playing a game, certain elements are permissible and follow each other regularly with high probability, but their order and exact usage can vary (*Suits, 1967*); other actions that would be physically possible as consequent are not used. They are recognisable by the use and sequence of certain central play elements but are not 'reproducible' in the sense that no two games of the same kind will look the same (*Malaby, 2007*). Using the transition probabilities of each element to each other element, we can identify clusters of elements that have similar transition patterns (*i.e.*, act like 'synonyms'). The network structure allows us to identify elements that were essential to a game (in the sense that they occurred at higher rates than other elements in the cluster and connected other elements in the sequence; *Carvalho et al., 2008*). Lastly, we predict that the similarity and transition

clusters overlap—*i.e.*, we have clusters of elements are interchangeable and tightly linked in time.

METHODS

Sample

We scanned 116 h of video material from the Bossou video database (*Matsuzawa & Humle, 2011*), collected between 2009 and 2013. While footage from the Bossou outdoor lab has high video quality and filming consistency, the social composition of the group limits generalisability. The Bossou community at the time was small (around 13 individuals) (*Matsuzawa & Humle, 2011*). Due to the age distribution, there was only one infant, one juvenile, and one subadult individual in the group during data collection—making it difficult to differentiate between age effects and individual preferences (*Fröhlich, Wittig & Pika, 2016*). Eleven individuals were observed playing at least once; however, the distribution of observations was highly skewed, with the two juvenile/subadult players each participating in about 75% of all play bouts, while none of the adults participated in more than 20% of play bouts. Thus, most play elements and transitions were provided by two individuals, often playing with each other. In this study, we do not control for individual or age differences in play behaviour and sequences, due to the limited sample. These could make play transitions more predictable (individuals or specific age groups might have standardised ways of reacting that other group members know). Considerably more data would be necessary to control for individual- or dyad-level effects in transition patterns. We identified 143 bouts of social play across 35 videos—defined as play involving at least two individuals, with a new bout started if both individuals stopped playing for at least 5 s continuously (*Liebal, Call & Tomasello, 2004*). Bouts consisted of between three and 181 individual play elements (mean = 30.3), including between two and four players at any given time. For analyses, the bouts were split into individual-level bouts (every play element an individual performed during a bout), resulting in 306 individual-bouts.

Coding scheme

The coding scheme, with detailed definitions of all play elements and coding conventions, can be found in the associated repository. Potential play elements were identified from several sources—primarily, every behaviour indicated in *Nishida et al. (2010)* as potential play behaviour, the literature on ape gestural repertoires (*Genty et al., 2009*; *Graham, Furuichi & Byrne, 2017*; *Hobaiter & Byrne, 2011*; *Hobaiter & Byrne, 2014*), previous chimpanzee play literature (*Fröhlich, Wittig & Pika, 2016*), and descriptions of play elements in primates more widely (*Petru et al., 2009*). Often, these sources use different terms for similar play elements, and the definitions used here do not always overlap perfectly with those used previously. To our knowledge, the ethogram used here is the most detailed ethogram for chimpanzee play to date. Play elements can roughly be categorised as contact or non-contact, and as events (countable, one-off or repeated actions) or states (continuous behaviour with a clear start and end point). Social object play formed its own category, with multiple different ways of interacting with detached objects (mainly stones, nuts, and sticks) available. We coded play faces and vocalisations (where audible) when they were

clearly visible—however, the distance and video quality did not enable us to code facial expressions reliably, and we did not include their placement in the analysis. In total, our ethogram contained 118 different play elements, of which 106 were observed at least once. We assumed that the elements we defined are meaningfully different from each other. This might not be the case: the difference between *Retreat* (walking away from partner), *Flee* (running away from partner), and *Retreat Backwards* (walking away from partner while looking at them) might be an artifact of the coding scheme.

Coding was done using BORIS v.7.9 video coding software (*Friard & Gamba, 2016*). We coded bouts one player at a time and marked the start of every change in play element and mark all active play elements at that time point. For example, if an individual goes *bipedal*, this is marked. If, while bipedal, the individual approaches the partner, we would mark *bipedal/approach*. If they would then raise their arm while performing those actions, we would mark *bipedal/approach/arm raise*, and so on. This leaves us with a string of play elements with a time stamp for initiation. If any player stopped playing (*i.e.*, no play element was active), a Break was coded. The duration of play elements was available but was not considered in this study—we focus entirely on the sequential order.

Video coding of entire play bouts is slow, due to fast changes of behaviour and movements, and researchers usually focus only on play initiation and re-initiations (*Heesen et al., 2021a*; *Hobaiter & Byrne, 2011*). Due to the challenges of this detailed coding approach, no inter-rater reliability was performed, and results must be viewed with this limitation. Predictability should be higher in studies using simpler coding schemes, so if we can show high predictability using the current ethogram, we have taken the conservative approach. The dataset currently contains 5,711 play elements. Where possible, we present results including uncertainties, and used permutation and bootstrapping approaches to discriminate between spurious and reliable transition patterns.

Pre-processing

All pre-processing and analyses were conducted in R statistical computing software (*R Development Core Team, R Core Team, 2020*). The video coding data needed pre-processing to deal with three main problems inherent to the coding process: rare elements; some artificially common elements; and establishing the sequential order of co-occurring elements.

To robustly establish probabilities of transitions between elements, rare elements are a problem (*Silge & Robinson, 2017*). For example, if an element only occurs three times, and each time transitions into a different element, we do not know if the high transition probability would disappear with increasing sample size. We set the threshold at 20 occurrences per play element. However, removing these cases completely (as is often done in linguistic studies; *Silge & Robinson, 2017*) would be wasteful given the sample size of this study. For most play elements, we defined *a priori* with which other play element they would be combined if too few occurrences were observed (see Supplemental Information). Replacement elements were chosen based on similarity of movement. If the combination after this lumping process failed to reach the threshold, we nevertheless retained it. Thus, our rarest element had nine occurrences (see associated repository for

occurrence probabilities of all play elements before and after pre-processing). After this step, 68 play elements remained.

Some elements occurred at much higher frequencies than others. The seven most common elements (Bipedal, Hold, Follow-Other, Approach, Retreat-backwards, Retreat, Flee) were all coded continuously and therefore were noted every time a change occurred while they were active. Imagine a musical piece on the piano: sometimes one note is held while others are played. In play, a chimpanzee could go bipedal, but then perform other actions while the *Bipedal* was marked at every change in event. These elements potentially skew transition probabilities and mask transitions between other elements. Ideally, we want a sequence that reflects when individuals made the choice to use a specific element. We addressed this by detecting cases where one of those seven elements occurred multiple times in a row, and only retained the first case. If players stopped the continuous action (*e.g.*, stopped fleeing, then started again), the element was counted again.

In play, it is possible to go *Bipedal*, *Arm Swing* with one arm and *Hit* the partner with the other arm. This is problematic in terms of the transitions - does *Bipedal* lead to *Arm Swing*; or *Arm Swing* to *Bipedal*? This problem also occurs because some elements (*e.g.*, *Flee* or *Bipedal*) are continuous states, while others (*e.g.*, *Kick*) have a clearly defined beginning and end. We used permutations—randomly assigning order within co-occurring elements and repeating all analyses 1,000 times with different orders—as there was no *a priori* reason to assign primacy to one co-occurring element over another. Thus, all described transition probabilities are averages over multiple permutations, which is why transition counts are not integers. Two alternative approaches (random sampling of only one of the co-occurring elements, bag-of-words) can be performed using the attached R scripts and generally showed similar results.

Transition probabilities

The transition probability between antecedent and consequent were defined by the number of times the consequent followed the antecedent, divided by the number of times any element followed the antecedent (conditional probability). The antecedent could be a single element (used to establish first-order n-grams, networks, and transition similarities), but also n-grams of different order (*e.g.*, first order: *Hit*; second order: *Hit/Slap*; third order: *Hit/Slap/Tickle* etc.). The latter approach was taken to determine whether increased information about antecedents increases prediction accuracy. Current sample size prevents us from analysing long sequences, as the number of possible transitions increases exponentially with each new level. We limited the analyses to a maximum of three antecedent elements. We restricted ourselves to one-element consequents and did not consider non-adjacent contingencies (*Sonnweber, Ravignani & Fitch, 2015*; *Watson et al., 2020*).

The large number of possible combinations combined with a small dataset and the small number of individuals leads necessarily to overfitting: some combinations will only occur a few times and adding new information could influence our understanding of their function. We did two things to counter this: rare elements were combined, as described above. Where possible, we report some measure of robustness to give the

reader an understanding of how reliable results were. Robustness was established using bootstrapping procedures—randomly selecting 1,000 subsets of the data and establishing transition probabilities within those subsets.

Randomisation procedures

To test which elements followed which antecedent, we created a null model of 'expected' transitions using permutations of observed patterns. We chose this resampling approach over collocation analysis (*Bosshard et al., 2022*) to account for the regular co-occurrence of play elements that is usually not seen in single-modularity communication; however, using collocation analysis instead did not alter the results. We repeatedly randomized the order of elements across bouts: while the number of elements per bout, the probability of elements to occur across bouts, and the position of Breaks and missing data in each bout were kept the same, we randomly assigned element positions. Thus, transitions are considered significant if they were observed more often than would be expected if play elements were just strung together randomly given their base probabilities. We ran 1,000 randomisations to create the expected distribution for each transition and compare whether the observed transition probability fell within this distribution or not. To compare the observed and expected values, we provide a *p*-value (how many of the 1,000 randomisations show higher transition probabilities than observed; *Mielke et al., 2021*). We report transitions that occurred at least five times and that were significant at 0.01 level (*i.e.*, the observed value was higher than for 990 permutations). These calculations also constituted the basis for the network clusters described below.

Prediction accuracy

To understand the predictability of transitions rules, we applied the probabilities derived from a subset of the data to 'unknown' test data and explored how well the former predicted the latter (*Chollet & Allaire, 2018*). We tested the predictability of elements within bouts by calculating transition probabilities for 95% of all other bouts, then predicting each element in the remaining 5% of bouts based on their antecedents (k-fold validation). This was repeated 1,000 times per bout. We tested the expected correct classification if the consequent element was only determined by base occurrence probabilities (null model). The difference between this value and the observed prediction accuracy of the models tells us how much knowledge of the antecedent increases our predictions. Aside from using one element as antecedent (describing a simple Markov process), we repeated the process with two or three elements as antecedents (n-gram prediction; *Eisenstein, 2019*). For higher-order antecedents, the probabilities of the lower-order antecedents were combined (interpolation)—therefore, for *Approach/Stare At/Hit* as third-order antecedent, the probability is the product of the probabilities of the triad, *Stare At/Hit*, and *Hit*. This was done because many higher-order antecedents only occurred infrequently, and no information would otherwise be available as to which consequent was appropriate. For transitions that were never observed, Laplace smoothing was applied, assigning them one occurrence, and adapting all transitions accordingly (*Eisenstein, 2019*). If the prediction accuracy under those conditions was higher than for one element, this indicated hierarchical

processes - for example, if *Hit* correctly predicts to *Hold* 10% of the time, but *Stare At/H it* leads to *Hold* in 80% of the time, then the sequence order added information. We present the mean correct classification rate across all bouts and elements. In addition to predictions based on the transition probabilities, we implemented a naïve Bayes classifier using the 'e1071' package in R (*Meyer et al., 2021*). Naïve Bayes classifiers use vectors of feature values (in our case, the previous play element, two previous play elements, etc.) to predict the correct consequent using Bayes theorem (*Eisenstein, 2019*). Using an established classifier offers the advantage that classification is optimised and faster than the above-described prediction based on raw transition probabilities. However, naïve Bayes classifiers make a strong independence assumption, effectively assuming that the antecedents are independent from each other given the consequent class (*Eisenstein, 2019*). Therefore, while increased performance of the classifier with increasing number of antecedents would indicate that information about previous play actions increases predictability of what happens next, performance cannot be interpreted as based on sequential information.

Similarity

We determined whether there were play elements that resembled each other in which elements followed them and tested whether we could find clusters of similar elements. This is similar to the identification of synonyms in language (*Levshina, 2015*), and we did it both to test whether our assignment of distinct elements during coding was meaningful and to see whether there were clusters of interchangeable elements. Each element was represented by a vector of transition probabilities with all elements. We applied Uniform Manifold Approximation and Projection (UMAP; *McInnes, Healy & Melville, 2018*) to achieve two-dimensional representation for each vector using the 'umap' package (*Konopka, 2022*). We established similarity between play elements by calculating the Euclidean distances between UMAP projections. To identify the optimal number of clusters for the hierarchical clustering, we used K-Means clustering as implemented in the 'cluster' R package (*Maechler et al., 2022*) to determine (a) the optimal number of clusters, and (b) the quality of the cluster solution. We present the silhouette value (*Rousseeuw, 1987*) to detect the best cluster solution, indicating an acceptable distance between clusters and coherence within clusters—any solution above 0.3 can be considered to show that there is more similarity within than between clusters. As cluster solutions differ based on the outcome of the UMAP dimension reduction, we repeated the dimension reduction and cluster detection 50 times with varying numbers of epochs for the UMAP (on average 7000 epochs) and continue using the most likely cluster solution. We plot the dendrogram for the optimal cluster solution and saved cluster memberships for later comparison with network clusters.

Networks

Networks can be useful tools to visualise the connections between elements in communication networks and to identify clusters of elements that have above-expected connections with each other (*Allen et al., 2019*; *Aychet, Blois-Heulin & Lemasson, 2021*; *Barceló-Coblijn et al., 2017*; *Mielke et al., 2021*; *Weiss et al., 2014*). Here, we created a network using all play elements as nodes and the transition probabilities between them as

edges (*Newman, 2010*). Only transitions that were significantly more likely than expected and occurred at least five times were considered, to make the network intelligible despite the large number of elements and ensure biological relevance. Edges were weighted, representing the transition probabilities between elements; and directed, meaning that each dyad of elements was represented with two values (A to B, B to A). We used the 'igraph' and 'ggraph' R packages (*Csardi & Nepusz, 2006*; *Pedersen, 2021*) to create and visualise networks. To test whether distinct 'clusters' of play elements existed in the network (indicating groups of play elements that have strong connections with each other but weak connections to the outside), we used the 'cluster_optimal' community detection algorithm in igraph, which maximises modularity of clusters (*Csardi & Nepusz, 2006*). Clusters were considered to represent stronger connections within than between clusters if the modularity value of the cluster solutions was larger than 0.3. Cluster solutions were compared to those produced by the similarity measure above.

RESULTS

(a) Non-random transitions

There were 1,622 transitions that were observed at least one time. The histogram (Fig. 1) shows that most elements are followed by several different consequents with low probabilities. In only four cases did a consequent constituted more than 30% of all possible transitions of an antecedent, with two of those (*Drum Tree* and *Kick Dirt*) being loops—the element was repeated sequentially. At the same time, each element was observed to be followed by between seven and 53 elements. Thus, there was no tight coupling between any two elements. This might indicate random assignment—any elements could be followed by any other. However, it might also mean situation-specific responses that were tailored to the players' own previous action and the partners' reaction, or predictability at a higher order (*e.g.*, based on multiple antecedent).

We also visualized how robust transitions were (Fig. 2). Using bootstraps, we created an interval around the observed transition probabilities. We plotted the range of values for each transition for the 1,000 bootstraps (calculated as the highest transition probability minus the lowest transition probability of A to B in the set) against the number of times the antecedent was observed. For some rare elements, transition probabilities remained volatile. Transition probabilities of rare elements therefore must be interpreted with caution, and elements will be filtered to exclude rare transitions—in all descriptions of 'significant' transitions and in the networks, only transitions that occurred at least 5 times were considered and reported.

In total, 146/1622 transitions (9%) were significantly more likely than expected. More detailed depictions of these patterns can be seen in the network below and in the associated repository. When analysing the non-random transitions in detail, we found that many elements significantly followed themselves (21 out of 147 significant transitions). Several of the elements used here—for example, rocking or drumming on an object—are repeated actions and each occurrence was marked as independent event. In contrast to all observed transitions described above, many elements (17/68 elements) had no significant consequent,
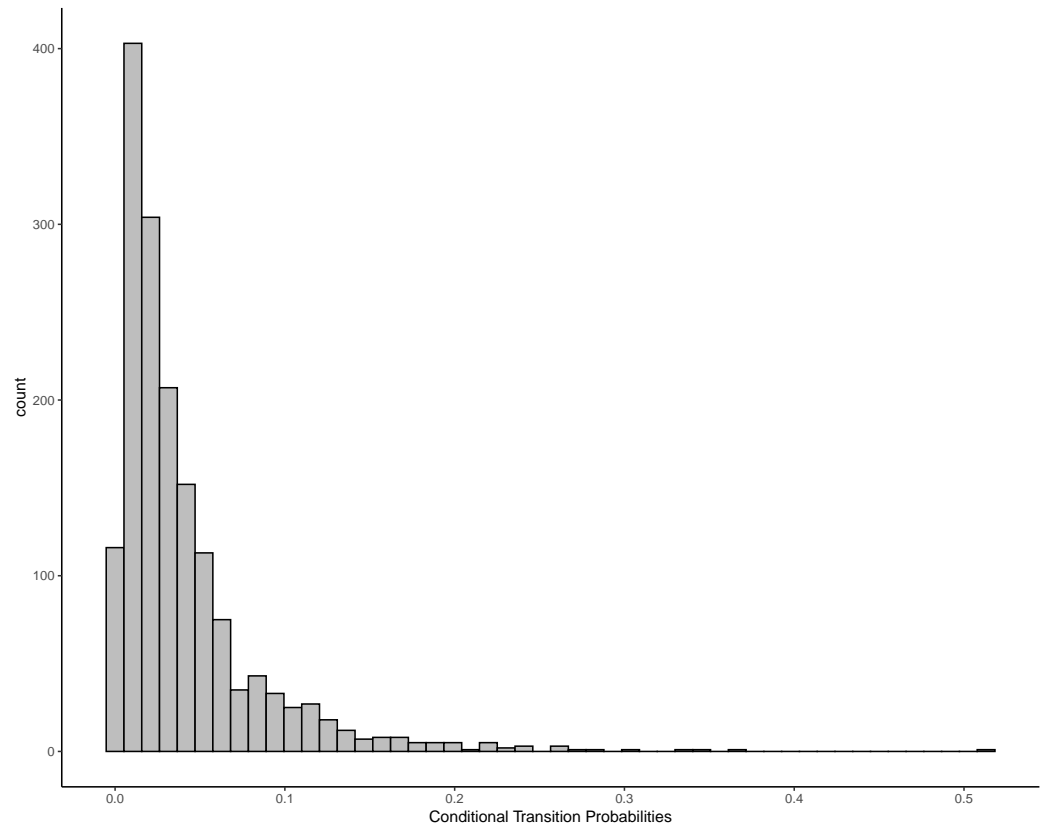
Figure 1 **Transition probabilities for each simple antecedent—consequent pair.** Each bar represents the number of transitions between antecedent-consequent pairs with a specific conditional transition probability.

14/68 had only one significant consequent, with the maximum number of significant transitions in one antecedent being 10 consequents (for *Holding* the partner and *Bipedal*).

(b) Next-element predictions

When applying the transition probabilities as predictions, increased information about antecedents increased predictability (Table 1). The basic probability of correctly predicting an element based on its occurrence probability (zero-order) was 0.03. By applying the probability of one antecedent (unigram; *e.g.*, *Hit*) we increased the probability to 0.06—almost a doubling of correct classification. When adding two antecedents (bigram; *e.g.*, *Bipedal/Hit*), there was another rise to 0.11—again, almost a doubling of correct classifications, and almost four times higher than having no information about antecedents. At the third order, we do not achieve further improvement. For the naïve Bayes classifier, using a more optimised approach that however assumes independence of antecedent elements, we achieve correct classification results of 0.07 as baseline, 0.13 for the first order, 0.16 for the second order, and 0.14 for the third order. Thus, additional information about preceding elements improved prediction accuracy. However, there was still a lot of unexplained variation.
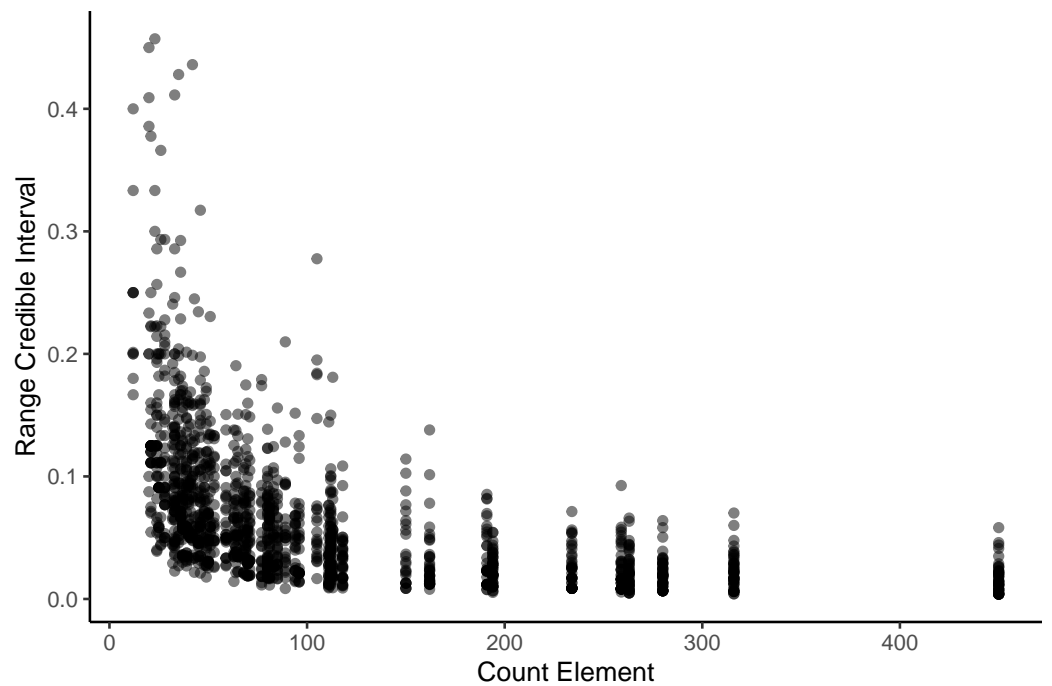
Figure 2 **Range of bootstrapped transition probabilities compared to the occurrence of the antecedent.** Each data point indicates the range between the 2.5th percentile and 97.5th percentile of bootstrapped transition probabilities, plotted against the occurrence frequency of the antecedent. Transitions of rare antecedents are volatile.

Table 1 **Correct prediction ability of consequent elements based on antecedents of different orders for the probability distribution and nave Bayes classifier.**

Order	Antecedent example	Accuracy probability	Accuracy Naïve Bayes
0	–	0.030	0.067
1	Hit	0.057	0.128
2	Stare At/Hit	0.105	0.155
3	Approach/Stare At/Hit	0.105	0.144

(c) Similarity between elements

In Fig. 3, we can see the dendrogram representation of hierarchical clusters of distances between transition probability vectors of all play elements. Elements connected through shorter branches and assigned the same cluster membership (same colour of branches) are considered more similar than those further away and with different colours. The best cluster solution, with silhouette value of 0.68 (indicating a well-distinguished cluster solution) contained 12 clusters. The cluster allocation can be seen in Table 2, and we will discuss their potential classification together with the network. What we can see here is that there were many elements that were similar in consequents. For example, *Kicking* the partner and *Jumping on* them were close, indicating that they could have been defined as a single

play element. Similarly, *Retreating Backwards* and *Retreating* were closely connected. A lot of similarity between elements can be explained by their frequent co-occurrence—for example, *Retreat* and *Bipedal* showed high similarity because chimpanzees often retreat from the play partner while bipedal.

(d) Network structure

In contrast to the similarity clusters, which assess whether two elements are used at similar points in a sequence, the transition network (Fig. 4) describes which consequent follows which antecedent. The network only depicts transitions that occurred at higher-than-expected rates and occurred at least five times in the dataset. Colours indicate community membership. As the high modularity of the network community detection algorithm (modularity = 0.65) indicates, there were seven clearly distinguished communities in the network. If community assignment was random, we would expect around 32% of transitions between the elements within communities, but we observed 48% of transition within communities—a 1.5-fold increase. Connections between communities were often due to elements that can be used in different situations. For example, *Shake Off* is used when playing *wrestling* with a partner to get away, but equally when the player is *hanging* off a branch or *retreating*—hence, the element is connected to three communities. Individuals *stomp* when initiating play in combination with *Bop* and *Bow*, but also when they were *bending* a small tree and holding onto it.

For the interpretation of communities, in combination with the similarity clusters, see Table 2. There was considerable overlap between the two approaches, with small variation arising mainly because several elements did not have any significant transitions above threshold level, and the combination of object-related and movement elements resulted in overlap between the chase and object clusters. The different cluster combinations ('games') can be categorised broadly by whether they involved climbing by either partner, had physical contact, involved chasing, involved objects, or were play invitations. For the latter, one clear cluster emerged, consisting of *Bop*, *Bow*, *Stomp*, and *Slap Ground*, which individuals often combined and repeated in quick succession to indicate that they were willing to play. Some other, rarer elements (*Present Body Part*, *Rock*, *Kick Dirt*, *Stare At*) can fulfil a similar function. Play elements routinely used when one or both individuals were in a tree transitioned into each other at high rates, depending on the role of the focal individual. When the player was on the ground and the partner in the tree, individuals would often *bend the tree* (the most central element of this cluster), and then *pull* or *shake* it, sometimes while *jumping*. Players regularly *hide swing* (swinging around a tree at speed) before climbing up. While players were in the tree, they *climb* up and then *hang* while *swinging*, *kicking* the partner, *shaking them off*, and ultimately *falling*.

Most contact play formed one large community in the network, with elements transitioning into each other at high rates. Based on the similarity of transition probabilities, we could differentiate two groupings: contact play that involves players to stay in one spot (*Bite*, *Wrestle*, etc.), centred on *holding* the partner in place; and those that involve one player trying to get away from their partner while still in contact (*Push, Trip* etc.).
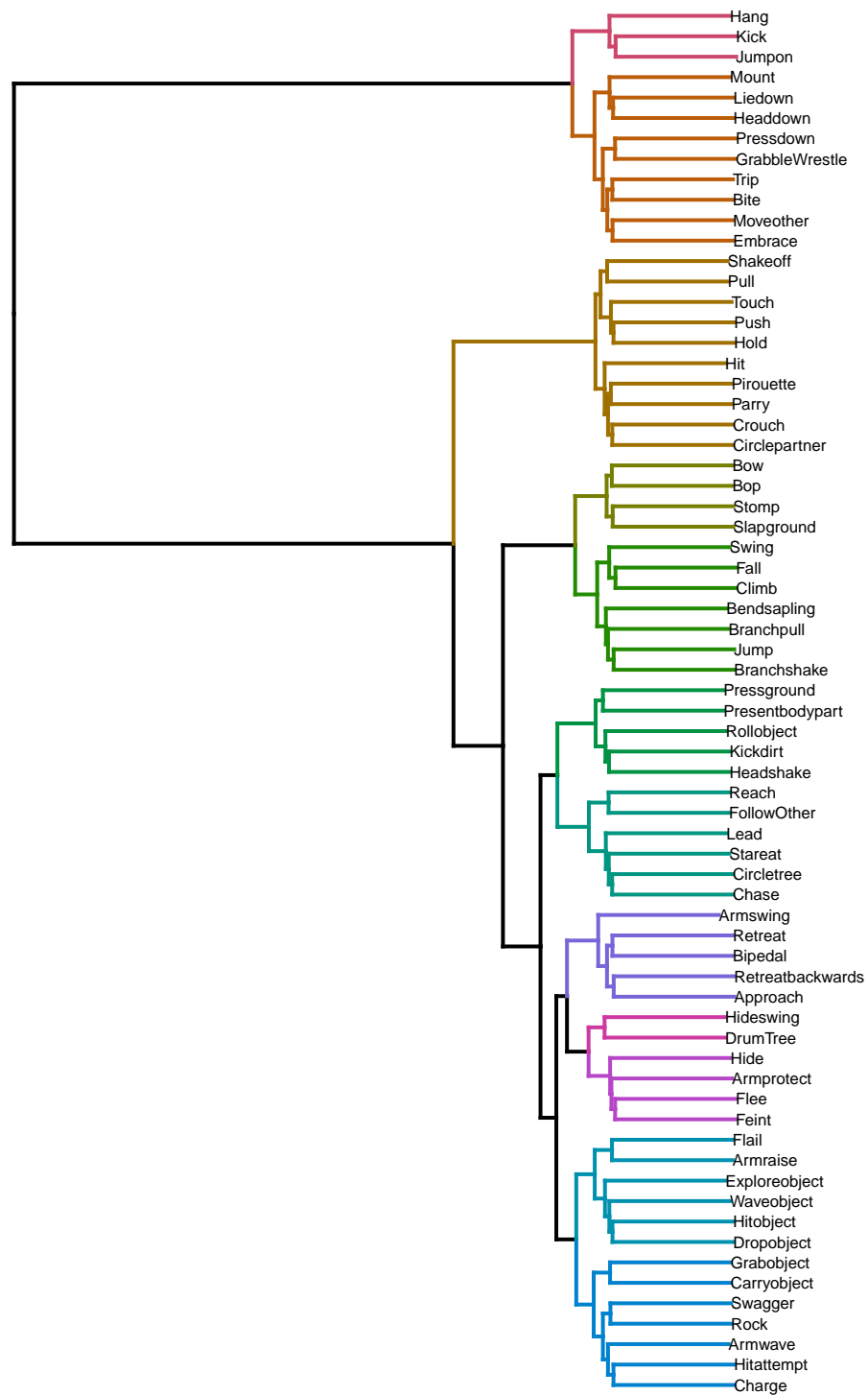
Figure 3 **Dendrogram of hierarchical clustering of distances between play elements.** Branch colours indicate established cluster membership. Optimal cluster solution based on k-means clustering: 12 clusters.

Table 2 Play elements with their cluster/community assignment for both the similarity and network of transition probabilities.

Play element	Similarity cluster	Network community	Game description
BranchPull	5	3	
Jump	5	3	*Tree Play: Focal on Ground*
BranchShake	5	3	
BendSapling	5	3	
Climb	5	6	
Swing	5	6	*Tree Play: Focal on Tree, No Contact*
Fall	5	6	
Kick	11	6	
Hang	11	6	*Tree Play: Focal on Tree, Contact*
ShakeOff	11	6	
HideSwing	3	6	*Tree Play: Transition to Tree*
Stomp	7	5	
Bop	7	5	*Initiation: Playful Approach*
Bow	7	5	
SlapGround	7	5	
Bite	6	4	
GrabbleWrestle	6	4	
Embrace	6	4	
Mount	6	4	
PressDown	6	4	*Contact Play: Wrestling*
LieDown	6	4	
Trip	6	4	
MoveOther	6	4	
HeadDown	6	4	
Hold	9	4	
Pull	9	4	
Touch	9	4	*Contact Play: Moving*
Hit	9	4	
Push	9	4	
Crouch	9	4	
DrumTree	3	2	*Object Interaction: Drum Tree*
RollObject	12	2	*Object Interaction: Initiation*
CarryObject	4	2	
ArmWave	4	2	*Object Interaction: Movement*
Swagger	4	2	
Flail	10	2	*Object Interaction: Play Threat*
ArmRaise	10	2	

(continued on next page)

Table 2 (*continued*)

Play element	Similarity cluster	Network community	Game description
ExploreObject	10	7	*Object Interaction: Explore Object*
HitObject	10	7	
RetreatBackwards	1	7	*Avoidance: Retreat*
Retreat	1	7	
Flee	2	7	*Avoidance: No Contact*
ArmProtect	2	7	
Hide	2	7	
Feint	2	7	
Parry	9	7	*Avoidance: Contact*
CirclePartner	9	7	
Approach	1	1	*Approach*
Bipedal	1	2	
ArmSwing	1	2	
CircleTree	8	1	*Follow: Chase*
Lead	8	1	
Chase	8	1	
Reach	8	1	
FollowOther	8	1	
HitAttempt	4	1	*Follow: Charge*
Charge	4	1	
GrabObject	4	1	
Pirouette	9	1	*Follow: Pirouette*
HeadShake	12	1	*Follow: Initiation*
PressGround	12	1	
PresentBodyPart	12	–	*No Significant Transitions: Initiations*
KickDirt	12	–	
Rock	4	–	*No Significant Transitions: Charge*
StareAt	8	–	*No Significant Transitions: Chase*
WaveObject	10	–	*No Significant Transitions: Explore Object*
DropObject	10	–	
JumpOn	11	–	*No Significant Transitions: Focal on Tree*

The different object-related play elements were connected, including detached objects and trees. Chimpanzee players held onto objects once they had grabbed them and then manipulated them in different ways. Object contact was the defining element of this type of play. A common way for the Bossou chimpanzees to initiate play with object contact was to *roll objects* towards the partner or *press the ground*. Individuals will often *drum trees*. Players *wave objects* about while *swaggering* towards the partner and *flailing* or *waving* their arms. Many of the social object elements were connected to *retreating* movements,

Transition Network Play Elements; Modularity = 0.66
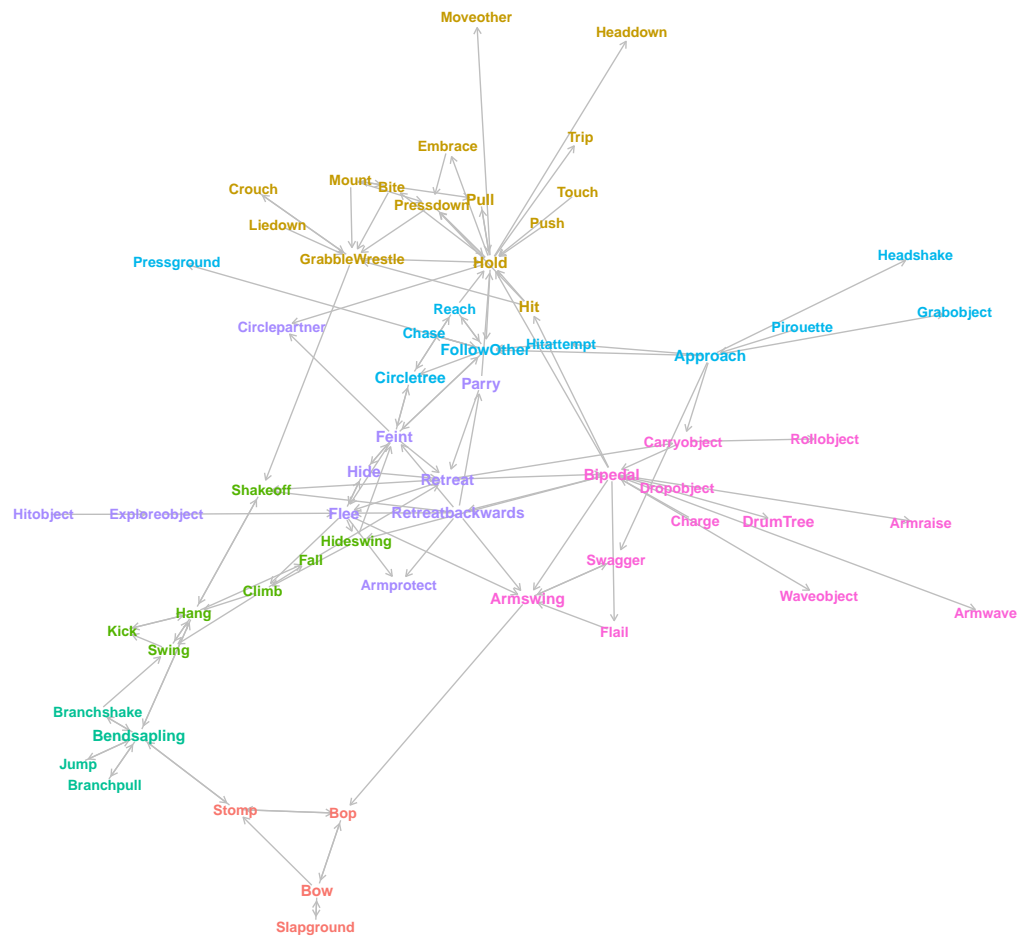
Figure 4 **Network plot of weighted transition probabilities between play elements.** Play elements are nodes, significant transitions that occur at least 5 times are edges (directed), and colour indicates cluster membership.

with the player retreating while holding an object, which explains the community overlap of object interactions and avoidance movements.

The remaining cluster combinations were related to chasing play on the ground. Again, we can identify different roles of the player: One community were those elements strongly connected to movements used to avoid the partner *retreating* or *retreating backwards* from them and *hiding* behind trees or *feinting* directional changes, often lifting their *arm protectively*. They will *circle the partner* while *parrying hits*. The last cluster combination involved the opposite, with the individual *approaching* the partner (sometimes following a *pirouette* as play initiation, often combined with *bipedal* movements and *arm swings*), *chasing*, and trying to make physical contact while the partner flees (*Reach, Hit Attempt*).

DISCUSSION

In this study, we explored the sequence structure of Western chimpanzee play behaviour for the Bossou community. We were interested in how predictable play was, and whether we find distinct 'games' with clear rules for sequences used by each player. Despite the large number of play elements and of transitions that were observed infrequently, only a small number of transitions occurred at higher-than-expected rates. Information about the preceding play element allowed for more accurate predictions than random choice, and the predictions became more accurate when including more antecedent elements—however, play retained its unpredictability, as the accuracy of predictions remained low. The reason for this can be found in the patterns of different 'games': we showed that there were several clusters of highly connected play elements with similar transition patterns. Thus, when a player was climbing in a tree, there were only few play elements available to them (partially, but not exclusively due to the substrate), but the exact order cannot be predicted. This appears to be similar to human games—if two children play tag, there is a finite number of play elements that each of them can use to keep the game going, but it is not in either players interest to let the partner know which one is next. The conditions and rules of the interaction are reproducible, but the sequence of used actions is not (*Malaby, 2007*). Importantly, the clusters we detected indicated clear roles for at least some of the games, with play partners on the ground acting different from the one in the tree and avoiding play elements clearly distinguished from approaching elements in chases.

Animal play is characterised by its unpredictable nature compared to other contexts, leading to theories that it has evolved for young individuals to learn social and motor skills (*Fagen, 1981*; *Smith, 1982*). We show that, at least for the Bossou chimpanzees, play is a mix of predictability and unpredictability—while knowledge of previous actions allows us to improve predictive accuracy, play sequences are not simple Markov chains, where one or few antecedent actions allow for accurate reactions. However, that does not mean that play is random, as clear games emerged from our bottom-up, data driven approach. We detect clusters of elements that are used together and interchangeably, indicating a rule-based system where the game limits the number of appropriate responses. It is not clear whether transition rules are descriptive (typical response to specific situations that are learned by individuals) or prescriptive (perceived and enforced by players; *Flack, Jeannotte & De Waal, 2004*). Further studies will have to determine whether non-linear prediction methods, *e.g.*, deep learning (*Chollet & Allaire, 2018*) could increase predictive accuracy, and whether action sequences are better described using non-Markov processes (*Kershenbaum et al., 2014*). Using a naïve Bayes classifier improved predictive accuracy, and more complex machine learning algorithms and a larger dataset could further extend our ability to detect transition patterns.

Some of the games have previously been identified by researcher when coding primate play—for example, many studies code 'rough-and-tumble' play as an overarching category for physical play in close contact, especially if individual actions are not discernible (*Palagi et al., 2016*). Our results show that this category can be established with a data driven approach. The same is true for chasing games. Another overarching context is tree-related

play, either with the player climbing or on the ground. Lastly, we identified social object play as its own context, which equally has attracted research in the past as a possible window into game-like behaviour (*Shimada, 2006*; *Tanner & Byrne, 2010*). Each of those games consisted of some central elements—holding the partner, moving towards them, moving away from them, holding an object, hanging from a tree etc.—that defined the context, with other elements added more freely, similar to tool use sequences in chimpanzees (*Carvalho et al., 2008*). We found clear evidence of role-reversal between players, as has long been described for play across species (*Fagen, 1981*), including in chimpanzees—players on the ground have a clear role in tree play that differs from those of the partner in the tree, and chasing players use different play elements than those fleeing. However, it needs to be kept in mind that the small sample size for many of the elements makes some of these clusters unreliable and dependent on researcher choices for the clustering algorithms.

This study demonstrates the power of a 'grammar of actions' approach (*Pastra & Aloimonos, 2012*), where methods from natural language processing and syntactical analysis are employed to understand the sequential nature of behavioural actions. In humans, there is strong evidence that hierarchical sequence processing across different domains, such as language and manual actions, is linked (*Fitch & Martins, 2014*). Our study presents evidence that the ability of chimpanzees to produce hierarchically structured sequences is not limited to their communication (*Arcadi, 1996*; *Girard-Buttoz et al., 2022*; *Leroux et al., 2021*; *Liebal, Call & Tomasello, 2004*) and tool-related behaviour (*Carvalho et al., 2008*; *Estienne, Stephens & Boesch, 2017*; *Vale et al., 2017*), but is also prevalent in fast-paced social interactions that require adaptation to multiple partners in real time (*McCarthy, Jensvold & Fouts, 2013*). As play behaviour is arguably practice for other contexts (*Palagi, 2018*) and tool use (*Myowa-Yamakoshi & Yamakoshi, 2011*), it would be of interest to investigate how the complexity of play sequences changes with age and if these changes are reflected in more complex sequences in other contexts (*e.g.*, *Soldati et al., 2022*). Importantly, as play is much more widespread in the animal kingdom than combinatorial communication, the detailed analysis of action sequences in this context might be a promising approach to trace the evolution of cognitive abilities that would eventually lead to human syntax.

One aspect currently missing from the picture is partner behaviour: while within-player behaviour shows limited predictability, it might be more predictable when knowing what the partner did. Chimpanzees and other primates engage in turn-taking when communicating (*Chow, Mitchell & Miller, 2015*; *Fröhlich, 2017*), and play has been described as a context that elicits joint commitment between players, with clear evidence that they re-establish that commitment after breaks (*Heesen et al., 2021a*; *Heesen et al., 2021b*). Thus, we need an approach that understands social interactions (including play) as a complex system of decisions taken by all involved individuals. Importantly, we do not currently know the effect of rule violations on play bouts—play is frequently suspended (*Heesen et al., 2021b*) and facial expressions are used to extend play in situations that might otherwise be interpreted as agonistic (*Palagi et al., 2016*; *Waller & Dunbar, 2005*), but whether these are preceded by inappropriate responses from play partners is not currently known. One question is whether play is indeed more complex in its sequential structure than other social contexts, such as grooming or aggressions, or communicative
exchanges, especially when we include the constraints afforded by the play context—for example, playing in trees removes the possibility to slap the ground. The statistical analyses underlying this study can be replicated using any data consisting of sequences of discrete elements.

The specific research context of this study, using video footage of the Bossou chimpanzees while they are in the forest clearing of the outdoor lab, constrains the number of different games that could be observed—for example, water games (*Costa et al., 2019*) cannot be observed in this environment. The physical substrate around the outdoor lab limits the amount of arboreal play. The Bossou chimpanzees are a very small group and subadult players lack same-aged play partners. Results were based on a small number of players who had an outsized impact on the dataset, and accounting for individual-level idiosyncrasies and age-dependent contingencies in transitions between play elements might dramatically improve predictability (*Cordoni & Palagi, 2011*). Many play elements were rare, and we had to make choices on how to combine them; there was still considerably uncertainty for some of the transition probability estimates. Thus, while we describe a method to detect games, larger datasets and more varied collection contexts would be necessary to characterise chimpanzee games more broadly. We are not trying to describe species-specific play patterns for chimpanzees in general (which probably include strong developmental, individual, dyadic, and group-level effects), but show that in this fairly standardised sample, chimpanzee play shows complex sequential patterns. Importantly, our approach would allow direct comparisons between different communities of chimpanzees, based on transition probabilities and network patterns. As the form of play is less defined by its function than for example tool use, this might be a useful approach to study cultural differences in a meaningful way (*Boesch et al., 2020*).

The data collection, pre-processing, and analytical choices of this study introduce several researcher degrees of freedom that limit generalisability of results (*Wicherts et al., 2016*). Thus, we are interpreting all results regarding the structure of play element transitions conditional on the coding scheme and group. We set strict cut-offs for significance levels and the minimum number of observed transitions to err conservatively, but an increased dataset or different thresholds might influence results. Another choice we had to make was regarding co-occurring play elements. We chose to use permutations to randomly assign which elements occurred at what point in the sequence, but this approach necessarily increases noise in the data. Lastly, every study of play behaviour is using a different ethogram, with different levels of complexity. We would predict that a simpler coding scheme would result in higher predictability. Because of the complexity of the coding scheme chosen here, no inter-rater reliability was performed, thus results should be interpreted as conditional on the coding process.

In summary, we show that chimpanzee play behaviour in Bossou is a complex sequential process with an identifiable hierarchical structure—chimpanzees play games consisting of play elements that are interchangeable in their sequence position and transition into each other at higher rates than they transition into play elements that are representative of other games. Information about previous actions allows for prediction of subsequent elements and including more antecedent elements improves accuracy. Our results show

that there is considerable leeway to study decision-making and cognitive complexity in animal social interactions on the micro-level (*Gygax, Zeeland & Christina, 2021*), but this process, like the study of communication, requires detailed video analysis of long-term data (*Hobaiter & Byrne, 2011*). In the future, being able to achieve reliable behaviour recognition from video databases, as has been demonstrated for the Bossou chimpanzees (*Bain et al., 2021*), could be a valuable tool in reducing the coding effort involved. As it stands, our results further highlight the special place play behaviour holds in the cognitive and behavioural development of chimpanzees—by creating a save environment to explore and train fast-paced behavioural sequences, it allows young individuals to learn to predict how a partner will react in different social situations.

ACKNOWLEDGEMENTS

We thank Professor Tetsuro Matsuzawa for the long-term efforts to develop and support the research on wild chimpanzees, at the Bossou Field Station, in Guinea. We thank Daniel Schofield and Misato Hayashi for their continuous help in curating the Bossou video archive, and Sophie Berdugo for help identifying the chimpanzees. We are also very grateful to Dora Biro, Catherine Hobaiter, David R. Braun, and all the KUPRI researchers who helped to collect field data at Bossou between 2009 and 2013. Special thanks are due to Direction General de la Recherche Scientifique et l'innovation Technologique (DGERSIT) and to Dr. Ali Gaspard Soumah and the Institut de Recherche Environnementale de Bossou (IREB), République de Guinée, for the research permission for researchers to conduct field work at Bossou that generated the data that we used; as well as research assistants Boniface Zogbila, Gouanou Zogbila, Henry Didier Camara, and Gilles Doré, Pascal Goumy, for their invaluable work at Bossou.

Funding

This research was funded by the British Academy through Alexander Mielke's Newton International Fellowship. Alexander Mielke also received funding from the Leverhulme Trust. The long-term efforts to develop and support the research on wild chimpanzees, at the Bossou Field Station, in Guinea, were supported by grants from MEXT (#12002009, #16002001, #20002001, #24000001, #16H06283) and JSPS (Core-to-core CCSN and U04-PWS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
British Academy through Alexander Mielke's Newton International Fellowship.
Leverhulme Trust.
MEXT: #12002009, #16002001, #20002001, #24000001, #16H06283.
JSPS.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Alexander Mielke conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Susana Carvalho conceived and designed the experiments, authored or reviewed drafts of the article, provided Database, and approved the final draft.

Data Availability

All data, scripts, and instructions are available at GitHub:

https://github.com/AlexMielke1988/Mielke-Carvalho_Chimpanzee-Play.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.14294#supplemental-information.

REFERENCES

Allen JA, Garland EC, Dunlop RA, Noad MJ. 2019. Network analysis reveals underlying syntactic features in a vocally learnt mammalian display, humpback whale song. *Proceedings of the Royal Society B: Biological Sciences* **286(1917)**:20192014 DOI 10.1098/rspb.2019.2014.

Arcadi AC. 1996. Phrase structure of wild chimpanzee pant hoots: patterns of production and interpopulation variability. *American Journal of Primatology* **39(3)**:159–178 DOI 10.1002/(SICI)1098-2345(1996)39:3<159::AID-AJP2>3.0.CO;2-Y.

Aychet J, Blois-Heulin C, Lemasson A. 2021. Sequential and network analyses to describe multiple signal use in captive mangabeys. *Animal Behaviour* **182**:203–226 DOI 10.1016/J.ANBEHAV.2021.09.005.

Bain M, Nagrani A, Schofield D, Berdugo S, Bessa J, Owen J, Hockings KJ, Matsuzawa T, Hayashi M, Biro D, Carvalho S, Zisserman A. 2021. Automated audiovisual behavior recognition in wild primates. *Science Advances* **7(46)**:4883 DOI 10.1126/SCIADV.ABI4883.

Barceló-Coblijn L, Serna Salazar D, Isaza G, Castillo Ossa LF, Bedia MG. 2017. Netlang: a software for the linguistic analysis of corpora by means of complex networks. *PLOS ONE* **12(8)**:e0181341 DOI 10.1371/journal.pone.0181341.

Bard KA, Dunbar S, Maguire-Herring V, Veira Y, Hayes KG, Mcdonald K. 2014. Gestures and social-emotional communicative development in chimpanzee infants. *American Journal of Primatology* **76(1)**:14–29 DOI 10.1002/AJP.22189.

Bauer EB, Smuts BB. 2007. Cooperation and competition during dyadic play in domestic dogs, Canis familiaris. *Animal Behaviour* **73(3)**:489–499 DOI 10.1016/j.anbehav.2006.09.006.

Bekoff M, Allen C. 1998. Intentional communication and social play: how and why animals negotiate and agree to play. In: *Animal play.* Cambridge, UK: Cambridge University Press, 97–114 DOI 10.1017/CBO9780511608575.006.

Berwick RC, Okanoya K, Beckers GJL, Bolhuis JJ. 2011. Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences* **15(3)**:113–121 DOI 10.1016/j.tics.2011.01.002.

Boesch C, Kalan AK, Mundry R, Arandjelovic M, Pika S, Dieguez P, Ayimisin EA, Barciela A, Coupland C, Egbe VE, Eno-Nku M, Michael Fay J, Fine D, Adriana Hernandez-Aguilar R, Hermans V, Kadam P, Kambi M, Llana M, Maretti G, Morgan D, Murai M, Neil E, Nicholl S, Jayne Ormsby L, Orume R, Pacheco L, Piel A, Sanz C, Sciaky L, Stewart FA, Tagg N, Wessling EG, Willie J, Kühl HS. 2020. Chimpanzee ethnography reveals unexpected cultural diversity. *Nature Human Behaviour* **4(9)**:910–916 DOI 10.1038/s41562-020-0890-1.

Bohn KM, Schmidt-French B, Schwartz C, Smotherman M, Pollak GD. 2009. Versatility and stereotypy of free-tailed bat songs. *PLOS ONE* **4(8)**:e6746 DOI 10.1371/JOURNAL.PONE.0006746.

Bosshard AB, Leroux M, Lester NA, Bickel B, Stoll S, Townsend SW. 2022. From collocations to call-ocations: using linguistic methods to quantify animal call combinations. *Behavioral Ecology and Sociobiology* **76(9)**:1–8 DOI 10.1007/S00265-022-03224-3/TABLES/6.

Carvalho S, Cunha E, Sousa C, Matsuzawa T. 2008. Chaînes opératoires and resource-exploitation strategies in chimpanzee (Pan troglodytes) nut cracking. *Journal of Human Evolution* **55(1)**:148–163 DOI 10.1016/J.JHEVOL.2008.02.005.

Chollet F, Allaire JJ. 2018. *Deep Learning in R.* Shelter Island, NY: Manning Publications.

Chow CP, Mitchell JF, Miller CT. 2015. Vocal turn-taking in a non-human primate is learned during ontogeny. *Proceedings of the Royal Society B: Biological Sciences* **282(1807)**:20150069 DOI 10.1098/rspb.2015.0069.

Cordoni G, Norscia I, Bobbio M, Palagi E. 2018. Differences in play can illuminate differences in affiliation a comparative study on chimpanzees and gorillas. *PLOS ONE* **3(3)**:e0193096 DOI 10.1371/journal.pone.0193096.

Cordoni G, Palagi E. 2011. Ontogenetic trajectories of chimpanzee social play: similarities with humans. *PLOS ONE* **6(11)**:e27344 DOI 10.1371/journal.pone.0027344.

Cordoni G, Palagi E. 2012. Fair play and honest signals in immature chimpanzees. *Atti Della Societa Toscana Di Scienze Naturali, Memorie Serie B* **119**:97–101 DOI 10.2424/ASTSN.M.2012.14.

Costa RFP, Hayashi M, Huffman MA, Kalema-Zikusoka G, Tomonaga M. 2019. Water games by mountain gorillas: implications for behavioral development and flexibility—a case report. *Primates* **60(6)**:493–498 DOI 10.1007/s10329-019-00749-6.

Csardi G, Nepusz T. 2006. The igraph software package for complex network research. *InterJournal: Complex System* **1695(5)**:1–9.

Davila-Ross M, Allcock B, Thomas C, Bard KA. 2011. Aping expressions? Chimpanzees produce distinct laugh types when responding to laughter of others. *Emotion* **11(5)**:1013–1020 DOI 10.1037/a0022594.

Davila Ross M, Owren MJ, Zimmermann E, Ross MD, Owren MJ, Zimmermann E. 2009. Reconstructing the evolution of laughter in great apes and humans. *Current Biology* **19(13)**:1106–1111 DOI 10.1016/j.cub.2009.05.028.

Deblauwe I, Guislain P, Dupain J, Van Elsacker L. 2006. Use of a tool-set by Pan troglodytes troglodytes to obtain termites (Macrotermes) in the periphery of the Dja Biosphere Reserve, southeast Cameroon. *American Journal of Primatology* **68(12)**:1191–1196 DOI 10.1002/AJP.20318.

Demartsev V, Strandburg-Peshkin A, Ruffner M, Manser M. 2018. Vocal turn-taking in meerkat group calling sessions. *Current Biology* **28(22)**:3661–3666.e3 DOI 10.1016/j.cub.2018.09.065.

Diamond J, Bond AB. 2003. A comparative analysis of social play in birds. *Behaviour* **140(8–9)**:1091–1115 DOI 10.1163/156853903322589650.

Eisenstein J. 2019. *Introduction to natural language processing.* Cambridge, MA, USA: MIT Press.

Engesser S, Ridley AR, Townsend SW. 2016. Meaningful call combinations and compositional processing in the southern pied babbler. *Proceedings of the National Academy of Sciences of the United States of America* **113(21)**:5976–5981 DOI 10.1073/PNAS.1600970113/-/DCSUPPLEMENTAL.

Estienne V, Stephens C, Boesch C. 2017. Extraction of honey from underground bee nests by central African chimpanzees (Pan troglodytes troglodytes) in Loango National Park, Gabon: techniques and individual differences. *American Journal of Primatology* **79(8)**:e22672 DOI 10.1002/AJP.22672.

Fagen RM. 1981. *Animal play behavior.* Oxford, UK: Oxford University Press.

Fernandez-Duque E, Valeggia CR, Mason WA, Pellis SM, Iwaniuk AN. 2000. Adult-adult play in primates: comparative analyses of its origin, distribution and evolution. *Ethology* **106(12)**:1083–1104 DOI 10.1046/j.1439-0310.2000.00627.x.

Fitch WT, Martins MD. 2014. Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences* **1316(1)**:87–104 DOI 10.1111/nyas.12406.

Flack JC, Jeannotte LA, De Waal FBMM. 2004. Play signaling and the perception of social rules by juvenile chimpanzees (Pan troglodytes). *Journal of Comparative Psychology* **118(2)**:149–159 DOI 10.1037/0735-7036.118.2.149.

Friard O, Gamba M. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* **7(11)**:1325–1330 DOI 10.1111/2041-210X.12584.

Fröhlich M. 2017. Taking turns across channels: Conversation-analytic tools in animal communication. *Neuroscience & Biobehavioral Reviews* **80**:201–209 DOI 10.1016/j.neubiorev.2017.05.005.

Fröhlich M, Wittig RM, Pika S. 2016. Play-solicitation gestures in chimpanzees in the wild: flexible adjustment to social circumstances and individual matrices. *Royal Society Open Science* **3(8)**:e160278 DOI 10.1098/rsos.160278.

Genty E, Breuer T, Hobaiter C, Byrne RW. 2009. Gestural communication of the gorilla (Gorilla gorilla): repertoire, intentionality and possible origins. *Animal Cognition* **12(3)**:527–546 DOI 10.1007/s10071-009-0213-4.

Girard-Buttoz C, Zaccarella E, Bortolato T, Friederici AD, Wittig RM, Crockford C. 2022. Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Communications Biology* **5(1)**:410 DOI 10.1038/S42003-022-03350-8.

Graham KE, Furuichi T, Byrne RW. 2017. The gestural repertoire of the wild bonobo (Pan paniscus): a mutually understood communication system. *Animal Cognition* **20(2)**:171–177 DOI 10.1007/s10071-016-1035-9.

Graham KE, Furuichi T, Byrne RW. 2020. Context, not sequence order, affects the meaning of bonobo (Pan paniscus) gestures. *Gesture* **19(2–3)**:335–364 DOI 10.1075/GEST.19028.GRA/CITE/REFWORKS.

Gygax L, Zeeland YRA, Christina DAR. 2021. Fully flexible analysis of behavioural sequences based on parametric survival models with frailties—a tutorial. *Ethology* **128(2)**:183–196 DOI 10.1111/ETH.13225.

Heesen R, Bangerter A, Zuberbühler K, Iglesias K, Neumann C, Pajot A, Perrenoud L, Guéry J-P, Rossano F, Genty E. 2021a. Assessing joint commitment as a process in great apes. *IScience* **24(8)**:102872 DOI 10.1016/J.ISCI.2021.102872.

Heesen R, Genty E, Rossano F, Zuberbühler K, Bangerter A. 2017. Social play as joint action: a framework to study the evolution of shared intentionality as an interactional achievement. *Learning & Behavior* **45(4)**:390–405 DOI 10.3758/s13420-017-0287-9.

Heesen R, Zuberbühler K, Bangerter A, Iglesias K, Rossano F, Pajot A, Guéry J-P, Genty E. 2021b. Evidence of joint commitment in great apes' natural joint actions. *Royal Society Open Science* **8(12)**:211121 DOI 10.1098/RSOS.211121.

Hihara S, Obayashi S, Tanaka M, Iriki A. 2003. Rapid learning of sequential tool use by macaque monkeys. *Physiology & Behavior* **78(3)**:427–434 DOI 10.1016/S0031-9384(02)01006-5.

Hobaiter C, Byrne RW. 2011. The gestural repertoire of the wild chimpanzee. *Animal Cognition* **14(5)**:745–767 DOI 10.1007/s10071-011-0409-2.

Hobaiter C, Byrne RW. 2014. The meanings of chimpanzee gestures. *Current Biology* **24(14)**:1596–1600 DOI 10.1016/j.cub.2014.05.066.

Hunt G, Gray R. 2004. The crafting of hook tools by wild New Caledonian crows. *Proceedings of the Royal Society of London B* **271(Supp 3)**:S88–S90.

Kershenbaum A, Blumstein DT, Roch MA, Akçay Ç, Backus G, Bee MA, Bohn K, Cao Y, Carter G, Cäsar C, Coen M, De Ruiter SL, Doyle L, Edelman S, Ferreri-Cancho R, Freeberg TM, Garland EC, Gustison M, Harley HE, Huetz C, Hughes M, Hyland Bruno J, Ilany A, Jin DZ, Johnson M, Ju C, Karnowski J, Lohr B, Manser MB, McCowan B, Mercado 3rd E, Narins PM, Piel A, Rice M, Salmi R, Sasahara K, Sayigh L, Shiu Y, Taylor C, Vallejo EE, Waller S, Zamora-Gutierrez V.

2016. Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biological Reviews* **91(1)**:13–52 DOI 10.1111/brv.12160.

Kershenbaum A, Bowles AE, Freeberg TM, Jin DZ, Lameira AR, Bohn K. 2014. Animal vocal sequences: not the Markov chains we thought they were. *Proceedings of the Royal Society B: Biological Sciences* **281(1792)**:20141370 DOI 10.1098/rspb.2014.1370.

Kershenbaum A, Ilany A, Blaustein L, Geffen E. 2012. Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B: Biological Sciences* **279(1740)**:2974–2981 DOI 10.1098/RSPB.2012.0322.

Konopka T. 2022. umap: Uniform Manifold Approximation and Projection (0.2.8.0)..

Koops K, Furuichi T, Hashimoto C. 2015. Chimpanzees and bonobos differ in intrinsic motivation for tool use. *Scientific Reports* **5(1)**:1–7 DOI 10.1038/srep11356.

Leroux M, Bosshard AB, Chandia B, Manser A, Zuberbühler K, Townsend SW. 2021. Chimpanzees combine pant hoots with food calls into larger structures. *Animal Behaviour* **179**:41–50 DOI 10.1016/J.ANBEHAV.2021.06.026.

Leroux M, Townsend SW. 2020. Call combinations in great apes and the evolution of syntax. *Animal Behavior and Cognition* **7(2)**:131–139 DOI 10.26451/ABC.07.02.07.2020.

Levshina N. 2015. *How to do linguistics with R.* Amsterdam, Netherlands: John Benjamins Publishing Company.

Liebal K, Call J, Tomasello M. 2004. Use of gesture sequences in chimpanzees. *American Journal of Primatology* **64(4)**:377–396 DOI 10.1002/ajp.20087.

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. 2022. cluster: Cluster Analysis Basics and Extensions (2.1.3).

Malaby TM. 2007. Beyond play. *Games and Culture* **2(2)**:95–113 DOI 10.1177/1555412007299434.

Martin-Ordas G, Schumacher L, Call J. 2012. Sequential tool use in great Apes. *PLOS ONE* **7(12)**:e52074 DOI 10.1371/JOURNAL.PONE.0052074.

Matsusaka T. 2004. When does play panting occur during social play in wild chimpanzees? *Primates* **45(4)**:221–229 DOI 10.1007/s10329-004-0090-z.

Matsuzawa T. 2020. Pretense in chimpanzees. *Primates* **61(4)**:543–555 DOI 10.1007/S10329-020-00836-Z/FIGURES/5.

Matsuzawa T, Humle T. 2011. Bossou: 33 years. In: *The chimpanzees of Bossou and Nimba.* Tokyo: Springer, 3–10 DOI 10.1007/978-4-431-53921-6_2.

McCarthy MS, Jensvold MLA, Fouts DH. 2013. Use of gesture sequences in captive chimpanzee (Pan troglodytes) play. *Animal Cognition* **16(3)**:471–481 DOI 10.1007/s10071-012-0587-6.

McInnes L, Healy J, Melville J. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *Available at https://onikle.com/articles/284481*ArXiv preprint. arXiv:1802.03426.

Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F. 2021. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (1.7-9). TU Wien.

Mielke A, Waller BM, Pérez C, Rincon AV, Duboscq J, Micheletta J. 2021. NetFACS: using network science to understand facial communication systems. *Behavior Research Methods* **54**:1912–1927 DOI 10.3758/S13428-021-01692-5.

Myowa-Yamakoshi M, Yamakoshi G. 2011. Play behaviors involving the use of objects in young chimpanzees at bossou. In: *The Chimpanzees of Bossou and Nimba.* Tokyo: Springer, 231–240 DOI 10.1007/978-4-431-53921-6_25.

Nakamura M. 2012. Playing with his leg: a case of imaginary social play by an adolescent male chimpanzee at bossou? *Pan Africa News* **19(1)**:1–3 DOI 10.5134/157939.

Newman M. 2010. Networks: an introduction. In: *Networks: an introduction.* Oxford, UK: Oxford University Press DOI 10.1093/acprof:oso/9780199206650.001.0001.

Nishida T, Zamma K, Matsusaka T, Inaba A, McGrew WC. 2010. Chimpanzee behavior in the wild. In: *Chimpanzee behavior in the wild.* Tokyo, Japan: Springer Science & Business Media DOI 10.1007/978-4-431-53895-0.

Ouattara K, Lemasson A, Zuberbühler K. 2009. Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences of the United States of America* **106(51)**:22026–22031 DOI 10.1073/PNAS.0908118106.

Palagi E. 2018. Not just for fun! Social play as a springboard for adult social competence in human and non-human primates. *Behavioral Ecology and Sociobiology* **72(6)**:1–14 DOI 10.1007/s00265-018-2506-6.

Palagi E, Burghardt GM, Smuts B, Cordoni G, Dall'Olio S, Fouts HN, Řeháková Petrů M, Siviy SM, Pellis SM. 2016. Rough-and-tumble play as a window on animal communication. *Biological Reviews* **91(2)**:311–327 DOI 10.1111/brv.12172.

Palagi E, Cordoni G, Borgognini Tarli SM, Tarli SMB, Borgognini Tarli SM. 2004. Immediate and delayed benefits of play behaviour: new evidence from Chimpanzees (Pan troglodytes). *Ethology* **110(12)**:949–962 DOI 10.1111/j.1439-0310.2004.01035.x.

Pastra K, Aloimonos Y. 2012. The minimalist grammar of action. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367(1585)**:103–117 DOI 10.1098/rstb.2011.0123.

Pedersen TL. 2021. ggraph: an implementation of grammar of graphics for graphs and networks. R package version 2.0.5.

Petrů M, Charvátová MS, Lhota S, Špinka M, Charvátová V, Lhota S. 2009. Revisiting play elements and self-handicapping in play: a comparative ethogram of five old world monkey species. *Journal of Comparative Psychology* **123(3)**:250–263 DOI 10.1037/a0016217.

Pika S, Zuberbühler K. 2008. Social games between bonobos and humans: evidence for shared intentionality? *American Journal of Primatology* **70(3)**:207–210 DOI 10.1002/ajp.20469.

R Development Core Team, R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing Vienna Austria. DOI 10.1038/sj.hdy.6800737.

Ross KM, Bard KA, Matsuzawa T. 2014. Playful expressions of one-year-old chimpanzee infants in social and solitary play contexts. *Frontiers in Psychology* **5**:741 DOI 10.3389/fpsyg.2014.00741.

Rousseeuw PJ. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20**(**C**):53–65 DOI 10.1016/0377-0427(87)90125-7.

Sasahara K, Cody ML, Cohen D, Taylor CE. 2012. Structural design principles of complex bird songs: a network-based approach. *PLOS ONE* **7**(**9**):e44436 DOI 10.1371/journal.pone.0044436.

Shimada M. 2006. Social object play among young japanese macaques (macaca fuscata) in Arashiyama, Japan. *Primates* **47**(**4**):342–349 DOI 10.1007/s10329-006-0187-7.

Shimada M. 2013. Dynamics of the temporal structures of playing clusters and cliques among wild chimpanzees in Mahale Mountains National Park. *Primates* **54**(**3**):245–257 DOI 10.1007/s10329-013-0348-4.

Silge J, Robinson D. 2017. *Text mining with R. Available at* https://www.tidytextmining.com/.

Sirianni G, Mundry R, Boesch C. 2015. When to choose which tool: multidimensional and conditional selection of nut-cracking hammers in wild chimpanzees. *Animal Behaviour* **100**:152–165 DOI 10.1016/J.ANBEHAV.2014.11.022.

Smith PK. 1982. Does play matter? Functional and evolutionary aspects of animal and human play. *Behavioral and Brain Sciences* **5**(**1**):139–155 DOI 10.1017/S0140525X0001092X.

Soldati A, Muhumuza G, Dezecache G, Fedurek P, Taylor D, Call J, Zuberbühler K. 2022. The ontogeny of vocal sequences: insights from a newborn wild Chimpanzee (Pan troglodytes schweinfurthii). *International Journal of Primatology* DOI 10.1007/s10764-022-00321-y.

Sonnweber R, Ravignani A, Fitch WT. 2015. Non-adjacent visual dependency learning in chimpanzees. *Animal Cognition* **18**(**3**):733–745 DOI 10.1007/s10071-015-0840-x.

Steele J, Ferrari PF, Fogassi L. 2012. From action to language: comparative perspectives on primate tool use, gesture and the evolution of human language. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**(**1585**):4–9 DOI 10.1098/RSTB.2011.0295.

Stout D, Chaminade T, Thomik AAC, Apel J, Faisal AA. 2018. Grammars of action in human behavior and evolution. *BioRxiv* 281543 DOI 10.1101/281543.

Suits B. 1967. What is a game? *Philosophy of Science* **34**(**2**):148–156 DOI 10.1086/288138.

Tanner JE, Byrne RW. 2010. Triadic and collaborative play by gorillas in social games with objects. *Animal Cognition* **13**(**4**):591–607 DOI 10.1007/s10071-009-0308-y.

Tomasello M, Carpenter M, Call J, Behne T, Moll H. 2005. Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences* **28**(**5**):675–691 DOI 10.1017/S0140525X05000129.

Vale GL, Davis SJ, Lambeth SP, Schapiro SJ, Whiten A. 2017. Acquisition of a socially learned tool use sequence in chimpanzees: implications for cumulative culture.

Evolution and Human Behavior **38(5)**:635–644
DOI 10.1016/J.EVOLHUMBEHAV.2017.04.007.

Waller BM, Dunbar RIM. 2005. Differential behavioural effects of silent bared teeth display and relaxed open mouth display in chimpanzees (Pan troglodytes). *Ethology* **111(2)**:129–142 DOI 10.1111/j.1439-0310.2004.01045.x.

Watson SK, Burkart JM, Schapiro SJ, Lambeth SP, Mueller JL, Townsend SW. 2020. Nonadjacent dependency processing in monkeys, apes, and humans. *Science Advances* **6(43)**:eabb0725 DOI 10.1126/sciadv.abb0725.

Weiss M, Hultsch H, Adam I, Scharff C, Kipper S. 2014. The use of network analysis to study complex animal communication systems: a study on nightingale song. *Proceedings of the Royal Society B: Biological Sciences* **281(1785)**:20140460 DOI 10.1098/rspb.2014.0460.

Wicherts JM, Veldkamp CLS, Augusteijn HEM, Bakker M, Van Aert RCM, Van Assen MALM. 2016. Degrees of freedom in planning, running, analyzing, and reporting psychological studies: a checklist to avoid p-hacking. *Frontiers in Psychology* **7**:1832 DOI 10.3389/fpsyg.2016.01832.

Wimpenny JH, Weir AAS, Clayton L, Rutz C, Kacelnik A. 2009. Cognitive processes associated with sequential tool use in new caledonian crows. *PLOS ONE* **4(8)**:e6471 DOI 10.1371/JOURNAL.PONE.0006471.