This is the Wordfile of the manuscript that appeared in
Nonlinear Dynamics, Psychology and Life Sciences

Collaborative play in young children as a complex dynamic system; revealing gender related differences

Henderien Steenbeek, *University of Groningen, the Netherlands; University of Applied Studies, Groningen, the Netherlands*
Diny van der Aalsvoort, *Saxion University of Applied Sciences, the Netherlands*
Paul van Geert, *University of Groningen, the Netherlands*

RUNNING HEAD: Complex dynamic systems approach of learning-teaching trajectories

Correspondence should be addressed to: H.W. Steenbeek, Heijmans Institute, University of Groningen; Grote Kruisstraat 2/1, 9712 TS Groningen, +31 50 363 7993; h.w.steenbeek@rug.nl

ABSTRACT

This study was focused on the role of gender-related differences in collaborative play, by examining properties of play as a complex system, and by using micro-genetic analysis techniques. A complex dynamic systems model of dyadic play was used to make predictions with regard to duration and number of contact-episodes during play of same-sex dyads, both on the micro- (i.e., per individual session), meso- (i.e., in smoothed data), and macro time scale (i.e., the change over six consecutive play sessions). The empirical data came from a study that examined the collaborative play skills of children who experienced six twenty minute play sessions within a three week period of time. Monte Carlo permutation analyses were used to compare model predictions and empirical data. The findings point to strongly asymmetric distributions in the duration and number of contact episodes in all dyads over the six sessions, as a direct consequence of the underlying dynamics of the play system. The model prediction that girls-dyads would show longer contact episodes than boys-dyads was confirmed, but the prediction regarding the difference in number of peaks was not confirmed.

In addition, the majority of the model predictions regarding changes over the course of six sessions were consistent with the data. That is, the average duration and the maximum duration of contact-episodes increases both in boys-dyads and girls-dyads, but differences occur in the strength of the increase. Contrary to expectation, the number of contact-episodes decreases both in boys-dyads and in girls-dyads.

KEYWORDS: Collaborative play; gender differences; micro genetic analyses; dynamic modeling, complex dynamic systems approach, intra-individual variability

COLLABORATIVE PLAY IN YOUNG CHILDREN AS A COMPLEX DYNAMIC SYSTEM; REVEALING GENDER RELATED DIFFERENCES

Collaborative play is a major developmental task as children become part of a classroom and school community (Fantuzzo et al., 1995; Lutz, Fantuzzo & McDermott, 2002; Wentzel & Looney, 2007). During collaborative play interactions (Broadhead, 2006) children jointly develop rules that guide their activities without having the experience of being forced into the propositions of their partners (DeLisi & Golbeck, 1999; Robinson, Anderson, Porter, Hart, & Wouden-Miller, 2003). Collaborative play develops from asymmetrical interactions into increasingly symmetrical collaboration (Verba, 1994, 1998). As there has been little research on the relationship between opportunities to engage in collaborative play and the development of skills that allow increase of in-depth play, an experimental study was carried out to reveal emergence of collaborative play over time. . The experimental condition included same-sex dyads that were offered six play sessions within three weeks as opposed to the control condition which consisted of a single play session at the beginning of the first week only. On average, the dyads in the experimental condition showed a linear increase of time spent on collaborative play over the six points in time, was somewhat stronger for girl-dyads than for boys (Van der Aalsvoort & Van der Leeden, 2010). However, the latter study considered only averaged findings of group data. An important question that remained concerns the time serial or process nature of differences within sessions in individual dyads.

In order to answer this question the current study re-analyzes the data from the abovementioned study, following a complex dynamic systems approach (Steenbeek & van Geert, 2013). The aim of this study is to focus on the role of gender-related differences in collaborative play, by examining properties of play as a system, and by using microgenetic analysis techniques in addition to dynamical modelling. This focus on play patterns over the time frame of a single play session and changes in play sessions over a few weeks potentially reveals important information about the development of collaborative play in children over longer time frames (Granott & Parziale, 2002), and emerging gender-related differences.

The outline of this article is as follows: first, current insights into gender differences in play interactions are briefly discussed, followed by the central properties of a nonlinear dynamical systems approach to play processes. Then, a conceptual model for the short-term dynamics of play is briefly presented, as well as the output from the corresponding simulation model, which yields predictions about gender-related differences during play. In the method and results section, empirical data are presented. Section four concludes with a discussion about the added value of this approach for analyzing empirical data, for theory building and practice.

Gender related differences in collaborative play

What makes boys-dyads operate in a different way than girls-dyads? According to Fabes, Martin and Hanish (2003) the communication between boys generally involves more active and forceful physical contact, whereas girls tend to emphasize collaboration among play partners and use enabling ways of communication (Jordan, 2009). From infancy on boys are mainly engaged in cooperative, rule-governed activities, whereas girls mainly engage in associative, social conversation behaviors (Pellegrini, Blatchford, & Kato, 2001; Sondell, 2002). Differences in play styles occur far into middle childhood, where boys seem to be more likely involved in ball games and fantasy play through a more rough-and-tumble fashion, and girls more likely in conversations, sedentary play, and verbal games (Blatchford, Barnes, & Pellegrini, 2003; Buhrmester, 1996; Lansford & Parker, 1999; Mathieson & Banerjee, 2011; Moller, Hymel, & Rubin, 1992; van der Aalsvoort, Van Tol, & Karemaker, 2004). Moreover, girls engage in more prolonged interactions than boys (Benenson, Aposteloris, & Parnass, 1997). In addition, marked gender differences in empathy and related capacities are reported, which are noticeable from an early age on (Auyeung, Wheelwright, Allison, Atkinson, Samarawickrema, & Baron-Cohen, 2009; Eisenberg & Lennon, 1983; Romer, Ravitch, Tom, Merrell, & Wesley, 2011; Rueckert & Naybar, 2008). These differences are observable in the style and properties of peer interaction in young children (Rose & Rudolph, 2006). However, gender differences in a more or less obligatory interaction situation such as the current play sessions that are based on the demand of an adult experimenter could also be explained by a greater tendency to conformity in girls than in boys (Haun & Tomasello, 2011). That is, if asked to play together, girls will tend to comply with this request more consistently than boys. Nevertheless, one can argue that this tendency to conformity is an expression of the same underlying child-in-context characteristic, namely the level of empathy with which the child reacts to the demands of others (e.g. the teacher in the classroom). For the purpose of this article, we confine ourselves to focusing on the widely reported gender differences in *empathy during play*, in the general sense of the word. Finally, girls are more effective in sending and receiving emotional information than boys (Berk, 2013). The preceding results, which are based on group comparisons, share the assumption that play is like a sum of independent variables, for instance style of behavior plus less effective emotional information sending plus rougher interaction equals boys play (plus or minus random variation). However, the starting point in this study is that the play process emerges as a result of a dynamic process, which is inspired by our earlier work on building and empirically validating a dynamic systems model of play (Steenbeek, 2006; Steenbeek & van Geert, 2005, 2006, 2007, 2008, 2013). Differences between boys and girls emerge over time and can only be understood by studying the actual unfolding of play interaction in real time. Children form a system that constitutes, over the length of a particular play session, a certain pattern in the play and non-play actions of the system, i.e., episodes in which play partners actively interact alternate with episodes in which they focus on their own activities.

Complex Dynamic Systems and Application to Dyadic Play

A collaborative play session of two children can be conceived of as a complex dynamic system (Steenbeek & van Geert, 2005; Howe & Lewis, 2005). Such a complex dynamic system can be defined as a set of many components that interact with one another over time and that change as a consequence of those interactions. These changes self-organize into characteristic higher order patterns (Van Geert, 2008). The dynamic aspect implies that one state of a system changes into another state (Weisstein, 1999). That is, the components of the play system change over time, as a consequence of the interactions between those components. In this case, components are for instance the children's play actions, which together form the alternating contact- and non-contact episodes. The play system operates in a *developmental state space*. This state space comprises the whole of possible, developmentally relevant states distinguishable in the system; the states are described by the set of dimensions or variables needed to specify the system as being a *developmental* system (Van Geert, 2008, p.1). In a play session, possible states are: The state in which both children are playing together, or the state in which both children are playing on their own, or the state in which one child directs its behavior towards playing together whereas the other child wants to play alone. Over time, the system occupies a particular succession of states in its state space.

The first property of the play system is the *iterative character* of the process (Van Geert & Steenbeek, 2005; Steenbeek, 2006). That is, within a single collaborative play session each previous (play- or non-play-) action of one of the children of the dyad has an influence on the subsequent (re-)action of his peer, and vice versa. In the long run, each play session has an influence on the subsequent session. The reciprocal pattern of influences between the participants of play sessions occur at various time scales, resulting in the second property, namely that of the *occurrence of patterns and mechanisms of change on various* *interdependent time scales*. The first is the micro-genetic scale, representing changes on a micro-level, i.e. in seconds, in the form of 'having contact' and 'having no contact'. Secondly, the meso-level incorporates changes in patterns of these short-term interactions. For instance, changes in the frequency of alternating periods of contact/no-contact give information about patterns in the play dynamics, and how these patterns possibly change over time. Thirdly, the macro-level represents changes over a number of play sessions, showing in e.g. the amount and number of contact-episodes over several play sessions over a period of a number of weeks. In a dynamic system all levels of the developing system interact and consist of nested processes that unfold over many time scales (Van Geert & Steenbeek, 2005). A third property is the *occurrence of intra-individual variability,* defined as "differences in the behavior within the same individuals (or dyads), at different points in time" (Van Geert & Van Dijk, 2002, p. 341). Although it is a common belief that intra-individual variability reflects measurement error and that it is noise imposed upon the real underlying latent variable, there is now strong support for the idea that intra-individual variability is a central feature of complex systems, in particular with regard to development. For instance, Flynn and Siegler (2007) state that over the course of development, performance often oscillates between less and more variable periods. Changes in variability can indicate and support developmental transitions (e.g. Bassano & Van Geert, 2007; Van Dijk & Van Geert, 2007). Kello et al. (2010) summarize a host of studies that, contrary to expectation, show that many behavioral phenomena are in fact strongly asymmetrically distributed as a consequence of their underlying dynamics. In our play data intra-individual variability is interpreted as 'the change in the duration and number of contact-episodes'. We expect to find intra-individual variability, both on micro-, meso-, and macro-level, and we also expect that it will differ from mere random noise.

A dynamic systems based model of play

In order to make predictions about gender-related differences, we used our complex dynamic systems model of dyadic play (for details of this model, see Steenbeek & van Geert, 2005, 2006, 2007, 2008, 2013). This model simulates the real time course of dyadic play situations, including the children's other-directed behavior, the real play interaction and the emotional expressions. The model is a so-called agent model (Elsenbroich, 2012), in which each agent, i.e., a playing child, is represented as a network of interacting concerns, behaviors or action tools, emotional appraisals and emotional expressions. The dyadic activity results from the real-time interactions between the behaviors and emotional expressions of the players. The model generates interaction patterns that are qualitatively similar to the interaction patterns observed in real playing dyads, and can be used to predict various characteristics of play, such as the distribution of contact episodes, the intensity of emotional expressions, or the occurrence of ‘peaks’ in the intensity of playing together across a play session. The model was used in earlier work to predict differences in dyadic play patterns in dyads composed of children of different sociometric levels (individuals with a rejected, average and popular status; see Steenbeek & van Geert, 2007). However, it can also be used to predict gender-related differences in play.

The conceptual basis of the model is shown in Fig. 1. Children’s behavior and emotions are driven by concerns, the most basic of which are the concern for Autonomy and the concern for Relatedness (Cobb & Yackel, 1998; Cobb, Yackel, & Wood, 1998; Cziko, 2000; Deci & Ryan, 2002; Frijda, 1986, 2001). Each moment, the child appraises – in a process that occurs automatically – to what extent the current context provides a good balance between the concerns. This appraisal leads to a specific emotion and emotional expression, in which the child expresses her or his emotion about the concern gap (Meyer &Turner, 2006; Pekrun, 2006). This appraisal also leads to a certain behavior (note that expressions such as "leads to" are simplifications of processes that are to a considerable extent cyclical). The main behaviors during play – being directed towards 'playing together', or directed towards 'playing alone', – serve the balance between Relatedness-oriented and Autonomy-oriented concerns (Steenbeek & van Geert, 2013)

Insert Figure 1 about Here

The main parameter values in the simulation model were set based on our earlier work on play interaction in dyads (Steenbeek & Van Geert, 2006, 2007). In the present study, these parameters were fine-tuned on the basis of theoretical assumptions regarding gender differences. First, we assumed that the concern parameter, governing the concern for playing together or not (the relatedness parameter) should be similar for boys and girls, since all dyads received the same instruction, namely to play together for a while. The relatedness parameter affects the short term dynamics of the play session, including the number of playing-together sequences, maximum duration of those sequences et cetera. Although we expect boys and girls to differ on these indicators, this expectation should be an emergent property of the play dynamics, and not the result of a predefined difference in a parameter that directly affects these indicators.

Second, on the basis of the literature, we assumed that girls would show a higher level of empathy. In the model, empathy corresponds with the symmetry parameter, which governs the degree of mirroring the behavior of the other person in one's own behavior (Steenbeek & Van Geert, 2006, 2007). The values were set to 0.2 for the girls and 0.1 for the boys[1].

Third, long-term change (over the 6 repeated play sessions) was simulated by means of changes in the relatedness parameter. During the simulated play session, each agent (child) keeps track of the emotional evaluation (appraisal) of the moments of playing together and playing alone. If the playing together sequences were more positively evaluated than the playing alone sequences, the relatedness parameter value is increased by a small proportional value (and decreased in the other case). This new value will be the starting value of the next play session, and this occurs 5 times in a row, covering all six play sessions. The model output typically consists of a together versus alone play behavior per child per second, for a duration of 1000 seconds. In total, we simulated six consecutive play sessions for 500 boy and 500 girl dyads. This output was expressed in core variables of the play interaction, which were also examined in the empirical data.

Questions and hypotheses

The model output was used to yield predictions regarding the empirical data. These predictions can be divided into methodological predictions about the form of the distribution of contact-episodes which coincide with system characteristics of play, and predictions about emerging gender-related differences. A first methodological prediction based on the model simulations is that, as a direct consequence of the underlying dynamics of the play system, strongly asymmetric distributions in the duration and number of contact-episodes will occur (hypothesis 1). Secondly, the model predicts that contact-moments are not randomly dispersed over the total play time, but come in bursts of activity. Thus, there will be relatively few periods of intensive play contact, which can be referred to as *peaks of joint activity* (hypothesis 2).

Insert Table 1 about Here

With regard to the gender differences, 5000 simulation runs based on the dynamic model predicted that boys will have less contact episodes during one play session than girls (boys-19.5 versus girls-22.2), which last less long (average duration: boys-42,7 seconds versus girls-

86,1 seconds; maximum duration: boys-198,6 seconds versus girls-297,3 seconds (hypothesis 3). On a meso-level, the model simulations predict that girls-dyads and boys-dyads will on average have the same number of 'peaks of joint activity' during their play session (hypothesis 4). On a macro-level (six sessions), the model simulations predict that the average duration of contact-episodes and the duration of the longest contact-episode (the maximum episode) will increase both in boys-dyads and girls-dyads. In addition, the model predicts that the number of contact-episodes will decrease for boys-dyads and increase for girls-dyads. With regard to the fluctuations in number and duration of the contact-episodes over the six sessions, the model predicts that both boys-dyads and girls-dyads increase in the strength of the fluctuations, but boys-dyads show a smaller increase than girls-dyads (hypothesis 5).

METHOD

Participants

A multiple-case study was conducted, consisting of 18 same-sex dyads in which 24 boys (12 dyads) and 12 girls (6 dyads) were included (mean age 68 months ranging from 67 to 74 months). These dyads were chosen from the experimental group of the original dataset (Van der Aalsvoort & van Leeden, 2010). The students who participated in the play study, all of whom attended regular schools, were selected based upon a teacher questionnaire about the children's task orientation, agreeableness, well-being and externalizing behavior (van der Aalsvoort & van der Leeden, 2010). The children were selected only when they belonged to the normal population. To illustrate the temporal patterns, a representative boy and a girl dyad were retrieved. In Table 2 it can be seen that both dyads match the group means, and that they are therefore good representatives of the gender group to which they belong.

Insert Table 2 about Here

Procedure

Session description

Each play-session took place in a room separate from the classroom, and was recorded with a video camera. Prior to the first session each dyad was informed about the study and invited to take part, which none of the children refused. Their interest in the activity was elicited as follows. Wooden blocks and toy animals were placed in front of the dyad. The researcher asked: “Do you see the blocks and toy-animals? Can you build a zoo for me? It’s up to you to think about how to build the zoo and which animals are going to live in it. I will film what you play”. A dyad was invited to stop playing after 20 minutes or when a child asked permission to leave. The researcher was present during the sessions but she did not participate or interfere in the activities. At the end of each session the researcher made a positive comment on the play and took a photo, explaining this would enable her to re-build the zoo before the children’s next session. She then escorted the dyads back to their classroom. In the next session the configuration was re-built by the researcher. She asked the dyad whether this had been done correctly, and invited play to re-commence. This procedure was repeated for play sessions 2 to 6.

Transcripts as data

Transcripts were available of all the play sessions. All play sessions were categorized by a four-step procedure that was an adjustment of Verba’s (1994) procedure. The current study uses the assessment of episodes in the transcripts. An episode starts when one child initiates collaboration and lasts until the end of that collaboration. An episode was initiated by means of either a child’s verbal and nonverbal actions which were clearly directed towards the other child, and which then were followed by a reply, in the form of a comment or gesture that suggested communication. Collaboration was defined as the children having contact, either verbally (talking towards each other), or nonverbally (actions directed towards each other, e.g. giving an object to the other child). The seconds of time of each episode were added to compute the amount of time that a dyad collaborated. Episode time was computed as the percentage of session time. With regard to inter-rater reliability the mean Kappa and range of the Kappa scores with regard to episode were satisfying in that all scores were .75 and more (Van der Aalsvoort & Van der Leeden, 2010).

Variables

The contact-episodes were used as a main variable in order to examine the quality of the play interaction. Different aspects of the contact-episodes were analyzed, by examining the empirical data in three rounds, first by using the raw data of each play session separately (micro-level), secondly, by averaging the smoothed data (meso-level), and thirdly, by examining the change over the course of six play sessions (macro-level). The following variables were examined: on the micro level, the children's '*average duration of contact-episodes*', their '*maximum duration of contact-episode'*, and their '*number of contact-episodes'*, as they showed in each play session. For the methodological questions, we calculated these variables over sessions, and used the maximum value (or an average of such values) as an indicator of the extremeness of the left tail of the distribution.

With regard to the smoothed data at the meso-level the '*average number of peaks*' was examined. A peak represents a period in the play session in which the chance that 'contact' will take place between play partners was more than 50%, meaning that the raw data contained more or longer parts at level 1 (contact) than at level 0 (non-contact).

At the macro-level over six play sessions, the change in the variables '*average mean duration of contact-episodes*' (the '*average maximum duration of contact-episodes'* and the '*average* *number of contact-episodes')* were used (see Table 1 for an overview of the most important variables, on the micro-, meso-, and macro-level, as predicted by the model).

Data-analyses

Due to the labor-intensive nature of the data coding, the number of dyads in the sample is relatively small. In addition, the data are characterized by relatively idiosyncratic distributions (see results section). For these two reasons, we decided to perform our statistical analyses by using Monte Carlo permutation methods (Todman & Dugard, 2001) to simulate the null hypothesis that the relationship or property is based on chance. Such methods make no assumptions as to the size of the samples and the distribution of the data and test the null hypothesis by generating surrogate data that are randomly distributed across the conditions one wishes to compare. The null hypothesis holds that the properties of a play session are similar for boys and for girls, that there is no systematic change in those properties over the sequence of six play sessions, and, in general that all play sessions are interchangeable among children. First, we provide statistical tests for the observed differences between girls' and boys' data. In order to provide an idea of the importance of the empirically observed difference, values are expressed in terms of effect size, defined as the difference between the average scores of girls and boys, divided by the standard deviation of the scores. These effect sizes are reported as follows (Cohen, 1988): $d < 0.1$= no meaningful difference; $0.1 < d < 0.3$= small; $0.3 < 0.8$ =middle; $d > 0.8$ = big.

P-values are based on random permutation tests generating 5000 cases from the null hypothesis. Iin order to show the fit between the dynamic model and the data, we bootstrapped a confidence interval for the model predictions, by randomly sampling 12 boy and six girl dyads from the simulated cases (500 each for the boy and girl dyads). These random samples have a similar composition as the empirical sample.

As to the calculation of the variables, indicators of intra individual variability are alternation in contact and non-contact episodes during a play session (short-term variability), and changes in the average number or the length of contact episodes over the six repeated play sessions (long-term variability). Frequencies were determined by means of a kernel density frequency estimator, which produces smoothed histograms (Scheer, 2014). The smoothed curves were used to count the number of peaks, both in the model runs as in the empirical data. The dividing-value for counting a peak was set on the theoretical median value, which is 0.5 or a lower value (i.e., the chance that children have contact is lower or equal to 50% percent). That is, two peaks must be divided by a dip of 0.5 or lower; a peak must be higher than 0.5 in order to count as a peak.

On the macro level, the slope of the changes in number and length of contact-episodes (across the six sessions) indicates the direction of change and the strength of the direction (whether it is a significant change or not). The amount of fluctuations was calculated by means of the variable ‘absolute average number of fluctuations’ of the contact episodes, across six play sessions. This variable gives an indication about the strength of fluctuation, regardless of the direction of the fluctuation at hand. Stronger fluctuations across play sessions can be an indicator of higher adaptability or flexibility in the players.

RESULTS

Distribution of Contact Episodes

General distribution of the contact episodes

In order to test the prediction that strongly asymmetric distributions in the duration and number of contact-episodes will occur (hypothesis 1), it was examined whether the durations of contact-episodes in specific dyads follow a symmetric, e.g. normal distribution or a strongly asymmetric distribution with a tail to the right (e.g. a lognormal distribution). First, calculations of the durations of all contact-episodes during six consecutive play sessions in a representative girls-dyad (dyad 1) and a representative boys-dyad (dyad 2) were made. Episodes were collected for all the play sessions for the two dyads separately. Frequencies were determined by means of a kernel density frequency estimator (see Method section). Figure 2 shows that the distributions in these individual dyads are indeed strongly skewed to the right as predicted. In addition, the tail of the distribution for the girls-dyad lies around 500 seconds, whereas the tail of the distribution of the boys-dyad around the 250 seconds, which is considerably shorter.

Insert Figure 2 about Here

The dyad-specific smoothed histograms were then fitted to a lognormal distribution model. It appears that the lognormal models provided an excellent fit of the data (Fig. 2). Such patterns are characteristic of processes in which interacting components show a form of preferential attachment, that is, long episodes tend to become much longer, short episodes tend to be cut off earlier (Kello et al., 2010).

In addition, the *average duration of contact-episodes*, the *maximum duration of contact*-episodes, and the *standard deviation* of the contact-episodes for the girls-dyad (dyad 1) and the boys-dyad (dyad 2) were calculated over all six sessions, as a long term indicator of the play behaviors of boys-dyads and girls-dyads (see Table 3). In the girls-dyad the contact-episodes are considerably longer than in the boys-dyad, and they are also considerably more variable in terms of respective standard deviations (girls-dyad: m = 51.8, max = 527; me = 24.5, sd 76.8; boys-dyad: m = 31.9, max = 276, me = 15.5, sd = 43).

Insert Table 3 about Here

The lognormal model was then fitted to the contact episodes in each of the 18 dyads. The lognormal model described the distribution of every dyad, without exception. Second, we checked whether the difference between the above-mentioned boys- and girls-dyad was characteristic of boys and girls dyads in general. In order to do so, we concentrated on the variable '*maximum*' (representing the maximum duration of contact-episodes across all six sessions), and calculated this variable for all dyads. It was found that boys-dyads and girls-dyads significantly differed (p = 0.01) with regard to their maximum duration of contact-episodes. This means that these two dyads can be considered as representative cases. Thus, the first prediction of our simulation model that strongly asymmetric distributions in the duration and number of contact episodes would emerge in all dyads (hypothesis a), is confirmed. In addition, in girls-dyads the distributions are considerably more asymmetric than in boys-dyads, as evidenced by the difference in duration of the longest contact moment.

Distribution of peaks

Recall that a peak represents an extended period during the play session in which the dyad shows a higher than average intensity of playing together, eventually with intermittent short episodes of noncontact (see explanation in Method section). This is typical of the meso level timescale. The hypothesis is that there will be relatively few periods of intensive play contact, or peaks of joint activity (hypothesis 2).

In figure 3 the smoothed data of the first play session of the representative girls-dyad and the boys-dyad is shown (for a description of the smoothing technique, see Method section). The figure must be read as follows: The line represents the probability that the dyad will have contact (play together) on this particular moment during the play session; i.e., it roughly corresponds with the changing probability that a contact-episode can be found at that particular moment in time.

Insert Figure 3 about Here

Figure 3 clearly shows that the level of the contact intensity is higher in the girls-dyads than in the boys-dyad: in the girls it reaches the value 1 twice, and in the boys the maximum level of the contact probability is about 0.6. In the girls-dyad (Fig. 3), the huge drop around seconds 541 to 625 refers to girl 1sitting down and watching girl 2 building a cage. Girl 1 also looks around and looks at the camera. When she stands up, girl 2 looks up and another episode starts. The drop emerges as a complete ‘stand still’: No contact for a while at all. In the boys-dyad (Fig. 4), it can be seen that it took a while for the boys to start collaborating. A first remarkable dip around 661 seconds refers to boy 1 playing on his own just pushing animals around while boy 2 is building a tower.

Figure 3 illustrates that the total play session is organized around a small amount of sub-sessions in which the children have contact in a more intensive way. That is, the variation between episodes of intensive joint activity (the peaks) and episodes in which more solitary activity occurs is in no way the result of a random alternation of contact moments. In order to check whether this finding (that contact-episodes were not a random alternation of sequences, as found in these two representative cases; dyad 1 and dyad 2) were supported by all the other dyads, we calculated the number of peaks over all dyads. Our null hypothesis was that contact moments are randomly divided over the course of the play sessions, and that the observed alternation of periods of more and less play contact is just a coincidence. Based on a simulation of the bandwidth of peaks in randomly occurring contact moments, the null hypothesis was rejected[2]. The bandwidth of peaks in the real empirical data was significantly broader than that in the random data.

In sum, the model prediction that the contact-episodes would fluctuate in the form of relatively few peaks of intensified joint activity (hypothesis 2) is confirmed.

Gender-specific differences

Before discussing the difference between boy- and girl-dyads on each of the variables (hypothesis 3 -5), we calculated the probability that the totality of observed differences might just as well have resulted from the statistical null hypothesis model. To this end, we used the sum of squared effect sizes (differences between boy and girl dyads divided by the standard deviation of the variable) as a chi-square value that we could compare with the distribution of chi-square values generated by the statistical null hypothesis model. Five thousand simulations of the null hypothesis model resulted in a *p*-value of .024, which implies that we can reject the null hypothesis (the 95% confidence interval of the squared effect sizes for the null hypothesis model ranges from 0.45 to 6.05; the sum of squares effect sizes for the data is 6.14).

The duration and number of contact-episodes

The model predicted that boys-dyads will have, on average, fewer long contact-episodes than girls (hypothesis 3). To test this hypothesis, we first visually inspected the graphs of the representative girls- and boys-dyads. Figure 4 illustrates that the girls-dyad showed a number of contact-episodes that lasted for a considerable number of seconds, or even minutes. The boys-dyad does not have that kind of lasting contact-episodes. In addition, the boys start relatively late with their first, short, contact-episode.

Insert Figure 4 about here

Insert Table 4 about here

First, we confirmed that the *total* play- and contact-time of boys-and girls-dyads was roughly the same ($p = .29$, and .25 respectively). This means that we can safely compare the boys- and girls-dyads with regard to their duration and number of contact-episodes.

With regard to the *duration* of contact-episodes, the model predicted that boys-dyads would have *shorter* contact-episodes than girls-dyads. The results show (see Table 4) that the *average* duration of contact episodes per session is indeed shorter for the boys-dyads than for the girls-dyads (d = 0.96), as well as the *maximum* duration of contact episodes per session (d = 0.94).

With regard to the *number* of contact-episodes, the results did not confirm the model prediction that boys show *fewer* contact episodes than girls (*average number* boys-dyads = 18.8, girls-dyads = 15.8). Instead, boys-dyads were found to have more contact-episodes than girls-dyads (d = 0.66).

Peaks in the data (Hypothesis 4)

The model predicts that girls-dyads and boys-dyads will on average have the same number of 'peaks of activity' during their play session (hypothesis 4; boys m =3.2 peaks, girls m = 3.1. peaks; d = 0.03).

The results show (see Table 4) that there is a small effect size with regard to the difference in the *average* number of peaks per session of boys-dyads and girls-dyads (boys m =3.3 peaks, girls m = 2.9. peaks; d = 0.29). The data thus don not confirm the model predictions.

Changes over six play sessions

The model predicted that over the course of six sessions, the a*verage* duration of contact-episodes will increase both in boys-dyads and girls-dyads; the *maximum* duration of contact-episodes will also increase in boys-dyads and in girls-dyads as well; and the *number* of contact-episodes will stay the same for boys-dyads and girls-dyads. The strength of the in- or decrease, as expressed in effect size, is for average duration of episodes equal for boys-dyads and girls-dyads, and for maximum duration bigger. For the number of contact-episodes the model predicts that the number of contact-episodes will decrease for boys-dyads and increase for girls-dyads, but with negligible strength. With regard to the fluctuations in number and duration of the contact-episodes over the six sessions, both boys-dyads and girls-dyads increase in the duration and number of fluctuations, but boys-dyads show a smaller increase than girls-dyads (hypothesis 5).

First, we visually inspected an illustrative example of the data shown in Fig. 5. The change in the average mean duration of contact-episodes is represented over all the boys-dyads and girls-dyads. In addition, and for comparison reasons, the model's output with regard to this variable is also included. As can be seen, both in the boys-dyad and in the girls-dyad the average duration of contact-episodes increases over the six sessions, and this increase is also predicted by the model. The increase in the girls-dyads is stronger than the increase in the boys-dyads (the average duration of contact-episodes in sixth session for girls-dyads is 120 seconds, whereas the average duration for boys dyads is 60 seconds).

Insert Figure 5 about Here

In line with the abovementioned visual inspection, the results about the change over six sessions as expressed in the slope, show (see Table 4) that both boys-dyads and girls-dyads indeed increase in the *average mean* duration of contact-episodes. However, the strength of the increase in the average duration is much bigger in the girls-dyads than in the boys-dyads (slope boys = 3.85; slope girls = 17.99, *d* slope = 1.13). With regard to the change in the *maximum* duration of contact-episodes, the results show (see Table 4) that both boys- and girls-dyads show, as predicted, an increase in the duration of the longest contact-episodes (slope boys = 11.5, slope girls = 47. 75, *d* =1.11), and again the maximum length of the contact-episodes in the girls-dyad will increases more. With regard to the *number* of contact-episodes, the model predicted that the number of contact-episodes would decrease for boys-dyads and increase for girls-dyads, but with negligible strength, over six play sessions. The results show (see Table 4) that both boys-dyads and girls-dyads' number of contact-episodes decrease over the six sessions, (slope boys = -1.4; slope girls = -1.28). Instead of the expected difference in direction of change in number of contact-episodes, the data show that the contact-episodes of the boys simply tend to decrease more than those of the girls, but the probability that this is a chance effect is considerable (d = 0.08). With regard to the *amount of fluctuations* in the average duration of contact, boys-dyads differ -with girls-dyads over the six sessions (average number of fluctuations boys = 22; average number of fluctuations girls = 47, d = 1.05), in that the boys-dyads show less fluctuations. This could be an indicator that girls-dyads are more flexible in their contact versus non-contact alternations. With regard to the amount of fluctuations in the maximum duration of contact-episodes, boys-dyads again show fewer fluctuations than girls -dyads, with an average effect size (boys-dyads 93.28, girls-dyads 127.83, d = 0.45). Boys tend to fluctuate more across the six sessions in their *number* of contact episodes (boys = 6.02, girls = 5.53, d = 0.34). In summary, the great majority of the model predictions regarding changes over the course of six sessions in average duration, maximum duration, and number of contact episodes were consistent with the data (Table 4).

With regard to the difference in the *strength* of the change, the predictions of the model were partly confirmed. That is, the predicted higher amount of fluctuations in boys dyads in the maximum duration of contact episodes, and the predicted same slope for the number of contact episodes were confirmed. However, the model predictions were not confirmed with regard to the differences in strength of the slopes of the average and maximum duration of contact episodes, and the difference in the amount of mean fluctuations in the average duration and the number of contact episodes.

The data are based on a rather unbalanced sample of subjects, namely 12 boy dyads and six girl dyads, whereas the model simulations resulted in 500 boy and 500 girl dyads. Thus, in order to compare the quantitative predictions (regarding nine variables) of the model with the data, we suggest a simple comparison between the data and the 90% prediction intervals based on random sampling of 18 cases from the dynamic model simulations, each consisting of 12 boy and six girl dyads in order to arrive at a comparable sample composition with the empirical data. Figure 6 shows that seven out of nine observed variables are within or slightly higher than the upper boundary of the interval. One variable, relating to the fluctuation in the number of episodes, is considerably smaller than predicted by the model. In order to provide a very rough indication of whether a comparable fit with the simulation model might be achieved by the statistical null hypothesis model of the data, we calculated the probability that the *total effect size* predicted by the simulation model, the value of which is quite close to the total effect size of the data, could have resulted from the null hypothesis model ($p < .001$).

Insert Figure 6 about Here

CONCLUSION AND DISCUSSION

This study describes a new way of explaining the temporal structure of collaborative play, including gender differences, and of investigating the empirical predictions made by a theoretical model about collaborative play. This model was based on nonlinear dynamic systems principles, and simulated the interactions between two children, their basic concerns, emotional evaluations and behaviors, and generated sequences of actions covering the duration of a real collaborative play session. These simulated play sequences were compared with the temporal structure of observed (empirical) play sequences in girls and boys dyads, and predictions made on the basis of this model were tested against the time serial data of individual dyads. The model predictions focused in particular on gender related differences in the distribution of the main variable ‘contact-episodes’, more precisely the duration and number of contact-episodes, and change therein over six sessions (course of the intervention) of play. The predictions form a coherent pattern in that they all result from an underlying model of the short (micro-), meso- and long-term (macro-) dynamics of dyadic play interaction. In general, the data support this pattern of predictions.

The majority of the empirical findings were consistent with the model predictions. Especially the finding (in section 3.1) that the distribution of contact-episodes is typically heavy-tailed was consistent with the prediction made by that dynamic model. Heavy-tailed distributions are typically related to processes characteristic of complex and self-organizing dynamic systems (Mitzenmacher, 2004). The distributions are typically associated with power law distributions, which are scale-free. However, according to Stumpf and Porter (2012) the focus should be on the heavy-tailed nature of the distributions, rather than on the power law distribution per se. Heavy-tailed distributions have been demonstrated to result from processes of preferential attachment, that is to say processes of adding more weight to components that already have much weight (e.g. the tendency to visit websites that already have many visitors). Preferential attachment has been demonstrated in network structures (Barabasi, 2009; Baronchelli et al., 2013; Muchnik et al. 2013) and in the distribution of excellence in a variety of fields of human performance (Van Geert, Steenbeek, den Hartigh & Van Dijk, submitted). In the case of play, preferential attachment means that children tend to continue activities that have already been continued for a relatively long time within the current play session. In the dynamic play model, this tendency automatically evolves out of the interplay between intertwined components of activity, such as concern satisfaction, and the tension between continuing one’s own activity versus the tendency to connect with the activity of the interaction partner. Methodologically speaking, changes in such a system must be studied using methods that do justice to the system as a whole. Appropriate methods are micro-genetic and time-serial: In this study individual play sessions are coded in terms of successions of contact and noncontact episodes, and every dyad is studied in the form of a sequence of six such sessions over the long term. In this approach, it is the dyad and not the individual children that form the unit of analysis (Granott & Fischer, 1998). Our data describe the structure and dynamics of the system as a whole, i.e., as a system of two children as they locally and temporarily construct their own interaction context. The overarching aim of this kind of studies is to try to understand the underlying dynamics of the play process by building a simulation model of these dynamics (Hedström & Ylikoski, 2010; Machamer, Darden, & Craver, 2000). This type of modeling is entirely different from statistical models that try to understand the association between independent variables such as gender and dependent variables such as contact duration. The latter type of models cannot be used to simulate real-time successions of activities. With regard to the presumed gender differences in play, our findings point to differences in contact-episodes between boys- and girls-dyads, in the sense that boys have more contact-episodes which tend to last less long, both in the average duration and in the length of the longest contact-episodes. Over the course of an intervention, the form and number of contact-episodes change for both boys and girls-dyads.

A first limitation of this study is that the depth of the play was not incorporated in the analyses. Van der Aalsvoort and Van der Leeden (2010) showed that less frequent episodes coincided with deeper play involvement. The finding that the girls-dyad revealed a decrease of episodes refers to this phenomenon. Future studies should thus also incorporate the content of the play episodes. Second, the experimental condition consisted of 18 dyads which is a limited number that should be extended in future studies. Nevertheless, we obtained more insight into how long-term differences are built from small differences in individual play sessions. This information adds to findings from other play studies in that intra-individual variability and other important system properties play a fundamental role in grasping the typical flow that characterizes play. In sum, our approach of play as a nonlinear dynamic system adds to knowledge about children's development in play, in that we provided a technique that clarifies how gender-specific play activity patterns emerge in real-time (micro-genetic time scale). This approach reveals how principles that govern the emergence of such patterns may change over the course of repeated play sessions on the long-term time scale, and finally how these patterns differ between boys and girls on the basis of characteristic differences in activity parameters, such as the level of empathy. This helps to build up our knowledge base about play, the role of gender differences, and possibly relations to the (development of) psychosocial health of children (Berkhout, 2012).

ENDNOTES

[1] It is important to note that the parameter values have no intrinsic meaning. What matters is the relative difference between the parameters in relation to their function and the model. Hence, the difference in the symmetry parameter should not be interpreted as "girls are twice as empathic as boys", it simply means that girls have a stronger tendency than boys to mimic each other's behaviors

[2] The statistical simulation consisted of randomly reshuffling contact moments over a number of arbitrarily selected play sessions, and by checking whether the peaks resulting from these randomly ordered series showed a bandwidth that was as great as or greater than that of the peaks in the data. This was never the case.

REFERENCES

Auyeung, B., Wheelwright, S., Allison, C., Atkinson, M., Samarawickrema, N., & Baron-Cohen, S. (2009). The children's Empathy Quotient and Systemizing Quotient: Sex differences in typical development and in autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 39, 1509-1521. doi:10.1007/s10803-009-0772-x.

Barabási, A.-L. (2009). Scale-free networks: a decade and beyond. *Science*, *325*, 412–3. doi:10.1126/science.1173299

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in cognitive science. *Trends in Cognitive Sciences*, *17*, 348–360. doi:10.1016/j.tics.2013.04.010

Berk, L. (2013). *Child development* (9th edition). Boston: Allyn and Bacon. Berkhout, L. (2012). *Play and psycho-social health of boys and girls aged four to six*. Unpublised Doctoral Dissertation. University of Groningen- The Netherlands.

Broadhead, P. (2006). Developing an understanding of young children's learning through play: The place of observation, interaction and reflection. *British Educational Research Journal, 32*, 191-207. doi: 10.1080/01411920600568976

Chinn, C. A. (2006). The microgenetic method: Current work and extensions to classroom research. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Complementary Methods in Education Research* (pp. 439-457). Washington DC: AERA.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (second ed.). Lawrence Erlbaum Associates.

Coplan, R.J., Bowker, A., & Cooper, S.M. 92003). Parenting daily hassles, child temperament, and social adjustment in preschool. *Early Childhood Research Quarterly, 18,* 376-395. doi: 10.1016/S0885-2006(03)00045-0

Deci, E.L. & Ryan, R.M. (2002). Self-determination research: Reflections and future directions. In E.L. Deci, & R.M. Ryan (Eds.), *Handbook of Self-determination Research.* (pp. 431-441). Rochester, NY, US: University of Rochester Press.

DeLisi, R., & Golbeck, S. L. (1999). Implications of Piagetian theory for Peer Learning. In A. M. O'Donnell (Ed.), *Cognitive Perspectives on Peer Learning* (pp. 3-37). Mahwah NJ: Erlbaum Ass.

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin, 94*, 100-131. doi:10.1037//0033-2909.94.1.100

Elsenbroich, C. (2012). Explanation in Agent-Based Modelling: Functions, Causality or Mechanisms? *Journal of Artificial Societies and Social Simulation 15*, <http://jasss.soc.surrey.ac.uk/15/3/1.html>

Fabes, R.A., Lynn Martin, C., & Hanish, L.D. (2003). Young Children's' Play Qualities in Same-, Other-, and Mixed-Sex Peer Groups. *Child Development, 74,* 921-932. doi:10.1111/1467-8624.00576

Fantuzzo, J., Sutton-Smith, B., Coolahan, K.C., Manz, P. H., Canning, S., & Debnam, D. (1995). Assessment of Preschool Play Interaction Behaviors in Young Low-Income Children: Penn Interactive Peer Play Scale. *Early Childhood Research Quarterly, 10,* 105-120. doi: 10.1016/0885-2006(95)90028-4

Gliner, J. A., Leech, N. L., & Morgan, G. A. (2002). Problems With Null Hypothesis Significance Testing (NHST): What Do the Textbooks Say? *The Journal of Experimental Education, 71,* 83–92. doi:10.1080/00220970209602058

Granott, N., & Parziale, J. (2002). *Microdevelopment. Transition Processes in Development and Learning*. New York, NY: Cambridge University Press.

Haun, D. B. M., & Tomasello, M. (2011). Conformity to peer pressure in preschool children. *Child Development*, *82*, 1759-1767. doi:10.1111/j.1467-8624.2011.01666.x

Hedström, P. & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology, 36*, 49-67. doi:10.1146/annurev.soc.012809.102632

Howe, M. & Lewis, M.D. (2005). The importance of dynamic systems approaches for understanding development. *Developmental Review, 25*, 247-251. doi:10.1016/j.dr.2005.09.002

Jordan, B. (2009). Scaffolding learning and co-constructing understandings. In A. Anning, J. Cullen, & M. Fleer (Eds.), *Early Childhood Education* (pp. 39-52). Los Angeles: Sage.

Kello, C. T., Brown, G. D. A, Ferrer-I-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. C. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Sciences, 14,* 223-232. doi:10.1016/j.tics.2010.02.005.

Lutz, M.N., Fantuzzo, J., & McDermott, P. (2002). Multidimensional assessment of emotional and behavioral adjustment problems of low-income preschool children: Development and initial validation. *Early Childhood Research Quarterly, 17,* 338-355. doi:10.1016/S0885-2006(02)00168-0

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67,* 1-25. doi: http://www.jstor.org/stable/188611

Muchnik, L., Pei, S., Parra, L. C., Reis, S. D. S., Andrade, J. S., Havlin, S., & Makse, H. a. (2013). Origins of power-law degree distribution in the heterogeneity of human activity in social networks. *Scientific reports,* 3, 1783. doi:10.1038/srep01783

Robinson, C. C., Anderson, G. T., Porter, C. L., Hart, C. H., & Wouden-Miller, M. (2003). Sequential transitions patterns of preschoolers' social interactions during child-initiated play: Is parallel-aware play a bidirectional bridge to other play states? *Early Childhood Research Quarterly, 18,* 3-21. doi:10.1016/S0885-2006(03)00003-6

Romer, N., Ravitch, N. K., Tom, K., Merrell, K. W., & Wesley, K. L. (2011). Gender differences in positive social–emotional functioning. *Psychology in the Schools, 48,* 958-970. doi: 10.1002/pits.20604

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in preer relationship processes: potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin, 132*, 98-131.doi:10.1037/0033-2909.132.1.98

Rueckert, L., & Naybar, N. (2008). Gender differences in empathy: the role of the right hemisphere. *Brain and Cognition, 67,* 162-167.doi:10.1016/j.bandc.2008.01.002

Scheer, E. (2014). Introduction to Kernel Density Estimation. January 31, 2o014, http://www.ehow.com/info_11369017_introduction-kernel-density-estimation.html).

Scheer, E. (2014). Introduction to Kernel Density Estimation. Retrieved from http://www.ehow.com/info_11369017_introduction-kernel-density-estimation.html).

Steenbeek, H. (2006). *Modeling dyadic child-peer interactions; Sociometric status, emotional expressions, and instrumental actions during play*. Unpublished doctoral dissertation, University of Groningen, the Netherlands.

Steenbeek, H., & van Geert, P. (2005). A dynamic systems model of dyadic interaction during play of two children. *European Journal of Developmental Psychology, 2,* 105-145. doi:10.1080/17405620544000020

Steenbeek, H., Jansen, L., & Van Geert, P. (2012).Scaffolding dynamics and the emergence of problematic learning trajectories. *Learning and Individual Differences, 22.* 64-75. doi:10.1016/j.lindif.2011.11.014

Steenbeek, H.W., & van Geert, P. (2006). A Theory and Dynamic Model of Dyadic Interaction: Concerns, Appraisals and Contagiousness in a Developmental Context. *Developmental Review, 27,* 1- 40. doi:10.1016/j.dr.2006.06.002

Steenbeek, H.W., & van Geert, P.L.C (2007). 'Do you still like to play with him?; variability and the dynamic nature of children's sociometric ratings". *Netherlands Journal of Psychology, 3*, 86-110.

Steenbeek, H.W., & van Geert, P.L.C (2008). An Empirical Validation Of A Dynamic Systems Model Of Interaction: Do Children Of Different Sociometric Statuses Differ In Their Dyadic Play? *Developmental Science, 11*, 253-281. doi:10.1111/j.1467-7687.2007.00655.x

Steenbeek, H.W., & van Geert, P.L.C (2013). The emergence of Learning-Teaching Trajectories in Education: a Complex Dynamic systems Approach. *Nonlinear Dynamics, Psychology, and Life* Sciences. *17*: 233-269.

Todman, J. B., & Dugard, P. (2001*). Single-case and small-n Experimental Designs: A Practical Guide to Randomization Tests*. Mahwah (NJ): Erlbaum

Van der Aalsvoort, G. M., & Van der Leeden, R. (2010). The microgenetic emergence of collaboration play in 6-year-olds Developing At-risk. *International Journal of Educational Research, 48*, 274-285. doi:10.1016/j.ijer.2010.01.001

Van der Aalsvoort, G. M., Van Tol, A. M., & Karemaker, A. M. (2004). Social Play of Young Children At-Risk of Learning Difficulties: a situated performance? *International Journal of Disability*, *Development and Education*, *51*, 151-171. doi:10.1080/10349120410001687373

Van Dijk, M. & Van Geert, P. (2007). Wobbles, humps and sudden jumps: A case study of Continuity, discontinuity, and variability in early language development. *Infant and Child Development, 16*, 7-33. doi*:*10.1002/icd.506

Van Geert, P. (1994). *Dynamic Systems of Development; Change between Complexity and Chaos*. New York: Harvester Wheatsheaf.

Van Geert, P. (2003). Dynamic systems approaches and modeling of developmental processes. In J. Valsiner, & K. J. Conolly (Eds.), *Handbook of Developmental Psychology* (pp. 640-672). London: Sage.

Van Geert, P. (2008). Complex dynamic systems of development. In R. A. Meyers (Ed.), *Encyclopedia of Complexity and System Science. Vol. 2. Applications of Physics and Mathematics to social science* (pp. 1872-1916). New York: Springer.

Van Geert, P. (2009). Nonlinear complex dynamical systems in developmental psychology. In S.J. Guastello, M. Koopmans, & D. Pincus (Eds.), *Chaos and Complexity in Psychology; the theory of nonlinear dynamical systems* (pp. 242-281). Cambridge: Cambridge University Press.

Van Geert, P., & Steenbeek, H. (2005). Explaining after by before; basic aspects of a dynamic systems approach to the study of development. *Developmental Review, 25*, 408-422. doi:10.1016/j.dr.2005.10.003

Van Geert, P., & Steenbeek, H. (2006). The dynamics of scaffolding. *New Ideas in Psychology*, 23, 115-128. doi:10.1016/j.newideapsych.2006.05.003

Van Geert, P., den Hartigh, R., Steenbeek, H & van Dijk, M. (2013). A dynamic systems model of the emergence of talent and excellence across the life span. Manuscript submitted for publication.

Verba, M. (1994). The beginnings of collaboration in peer interaction. *Human Development, 37*, 125-139. doi:10.1159/000278249

Verba, M. (1998). Tutoring interactions between young children: How symmetry can modify symmetrical interactions. *International Journal of Behavioral Development, 22*, 195-216. doi:10.1080/016502598384577

Weisstein, E.W. (1999). *CRC Concise Encyclopedia of Mathematics*. Boca Raton-New York: CRC Press.

Wentzel, K. R., & Looney, L. (2007). Socialization in school settings. In J. E. Grusec, & P. D. Hastings (Eds.), *Handbook of socialization* (pp. 382-403). New York: The Guilford Press.

Table 1

The model predictions, based on the parameter values for boys-dyads and girls-dyads as described in the introduction section, and expressed in effect size

	micro-level			**meso-level**	**macro-level**					
	average duration of contact episodes	maximum duration of contact episodes	number of contact episodes	average number of peaks	average duration of contact episodes: *Slope*	average duration of contact episodes: *fluctuation*	maximum duration of contact episodes: *Slope*	maximum duration of contact episodes: *fluctuation*	number of contact episodes: *slope*	number of contact episodes: *fluctuation*
Model										
B	42.7	198.6	19.5	3.2	7.9	47.7	14.5	213.6	-0.1	9.5
G	86.1	297.3	22.2	3.1	13.7	109.1	11.7	272.8	0.2	18.7
effect size	0.71	0.84	0.5	0.03	0.17	0.69	-0.04	0.41	0.08	1.49
difference	B<G	B<G	B<G	B>G	B&G ↑	B<G	B&G ↑	B<G	B ↓ G ↑	B<G
effect size	middle	big	middle	nm	small	middle	nm	middle	nm	big

Table note: (*d <0.1= no meaningful difference(nm); 0. 1<d<0.3= small; 0.3<0.8 =middle; d>0.8 = big)*

Table 2

Means, standard deviations and F and p-values of the selection variables, ratings at the Early Childhood Education Rating Scales-Revised (ECERS-R), socioeconomic status (SES) and Free play time of the boys- and girls-dyad

	E-condition	C-condition	Boys-dyad	Girls-dyad
N	36	20		
	M (SD)	M (SD)		
Age in months	68.1 (5.6)	70. (4.72)	67	70
Raven CPM score	4.2 (0.95)	4.1 0.96	3	5
Language score	3.3 (1.55)	3.5 (1.28)	4	5
ECERS-R	40.9 (3.43)	39.2 (2.95)	39	45
SES	10.9 (3.38)	10.2 (3.56)	7	5
Free play time	8.7 (2.84)	10.1 (2.01)	18	8

Table 3

Mean, median, maximum value and standard deviation of contact episodes in girls- and boys-dyads

contact episodes		
	girls-dyad (1)	boys-dyad (2)
Mean	51.8	31.9
Median	24.5	15.5
Maximum	527	276
Standard deviation	76.8	43

Table 4

The empirical findings with regard to differences between boys-dyads and girls-dyads, expressed in direction of difference and effect sizes, compared with qualitative model predictions (see also Table 1)

	micro-level			**meso-level**	**macro-level**					
	average duration of contact episodes	maximum duration of contact episodes	number of contact episodes	average number of peaks	average duration of contact episodes: *Slope*	average duration of contact episodes: *fluctuation*	maximum duration of contact episodes: *Slope*	maximum duration of contact episodes: *fluctuation*	number of contact episodes: *slope*	number of contact episodes: *fluctuation*
Model										
difference	B<G	B<G	B<G	B>G	B&G ↑	B<G	B&G ↑	B<G	B ↓ G ↑	B<G
effect size	middle	big	middle	nm	small	middle	nm	middle	nm	big
Data										
B	47.51	173.89	18.81	3.3	3.8	21.9	11.5	93.3	-1.4	6.0
G	72.66	255.56	15.78	2.9	18.0	47.1	47.8	127.8	-1.3	5.5
effect size	0.96	0.94	0.66	0.29	1.13	1.05	1.11	0.45	0.06	-0.27
difference	B<G	B<G	B>G	B>G	B&G ↑	B<G	B&G ↑	B<G	B&G↓	B>G
confirmation	*yes*	*yes*	*no*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes/no*	*no*
effect size	big	big	middle	small	big	big	big	middle	nm	middle
confirmation	*no*	*yes*	*yes*	*no*	*no*	*no*	*no*	*yes*	*yes*	*no*

Table note: (*d <0.1= no meaningful difference (nm); 0. 1<d<0.3= small; 0.3<0.8 =middle; d>0.8 = big)*

FIGURE CAPTIONS

Figure 1. The conceptual basis of the simulation model of dyadic play, consisting of the dynamic interplay between the children's concern for Autonomy and the concern for relatedness, appraisals, behaviors, and empathy, at a particular moment during play.

Figure 2. Distribution of the durations of all contact episodes during six consecutive play sessions in a representative girls-dyad (above) and a representative boys-dyad (below), in comparison with a lognormal distribution model.

Figure 3. The smoothed data of the girls dyad (dyad 1; above) and the boys-dyad (dyad 2; below) of the first play session. 1 = dyad played together; 0 = dyad played alone.

Figure 4. The first play sessions of the girls-dyad (above) and the boys-dyad (below) (raw data) showing whether the children had contact (1), or not (0) over time.

Figure 5. Average mean duration of contact-episodes, in seconds as a function of the number of sessions for all girls-dyads and boys-dyads (series 1) and model output (series 2).

Figure 6. A comparison between the data and the 90% prediction intervals based on random sampling of 18 cases from the dynamic model simulations, each consisting of 12 boy and six girl dyads in order to arrive at a comparable sample composition with the empirical data

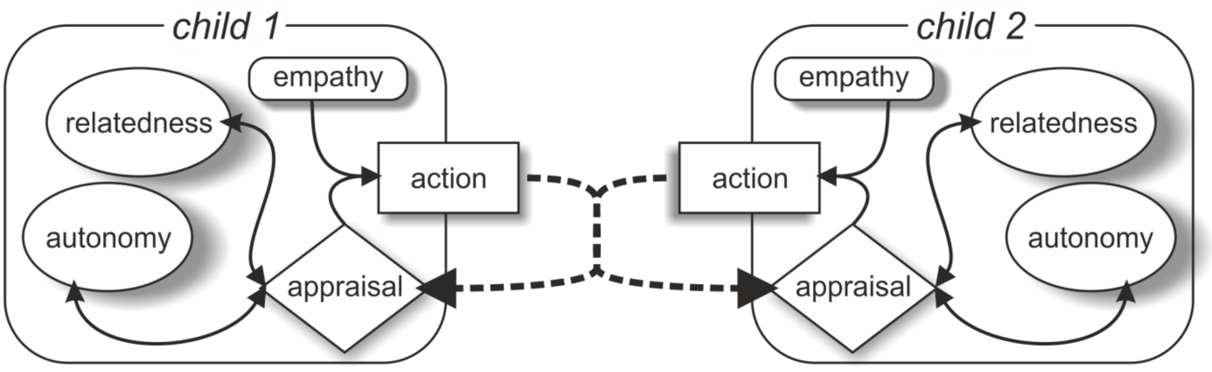

child 1
empathy
relatedness
action
autonomy
appraisal
child 2
empathy
relatedness
action
autonomy
appraisal

FIGURE 1

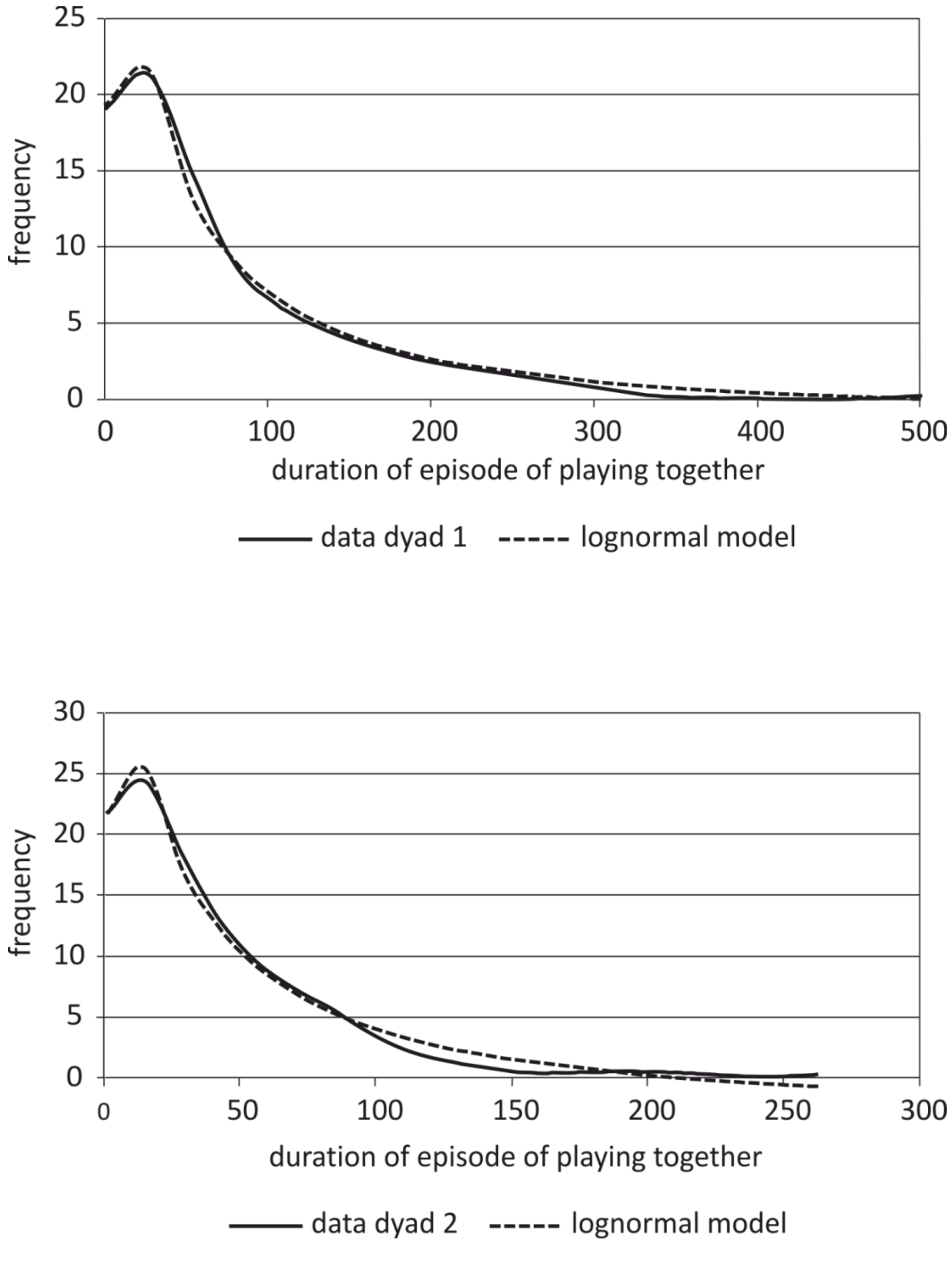
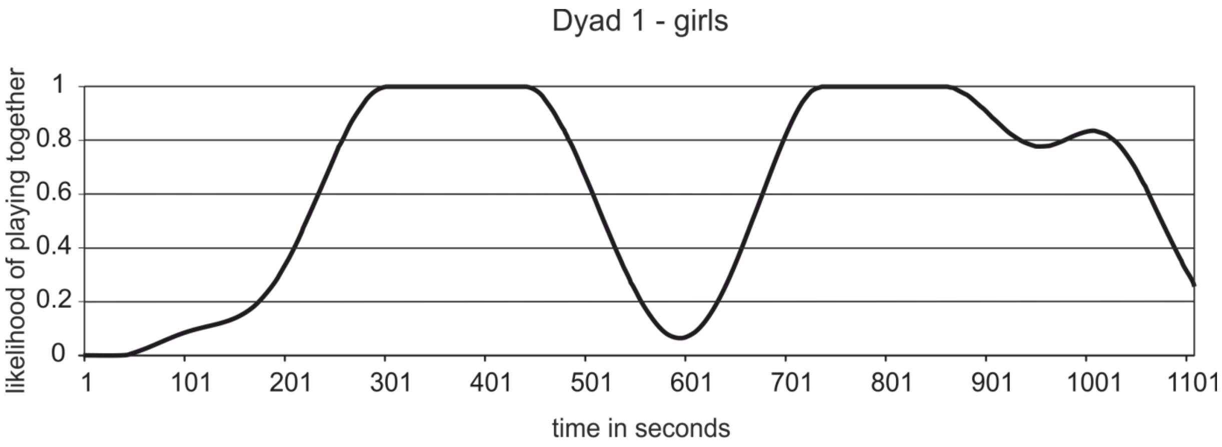

25
20
15
10
5
0
frequency
0
100
200
300
400
500
duration of episode of playing together
data dyad 1
lognormal model
30
25
20
15
10
5
0
frequency
0
50
100
150
200
250
300
duration of episode of playing together
data dyad 2
lognormal model

FIGURE 2


Dyad 1 - girls
likelihood of playing together
1
0.8
0.6
0.4
0.2
0
1
101
201
301
401
501
601
701
801
901
1001
1101
time in seconds

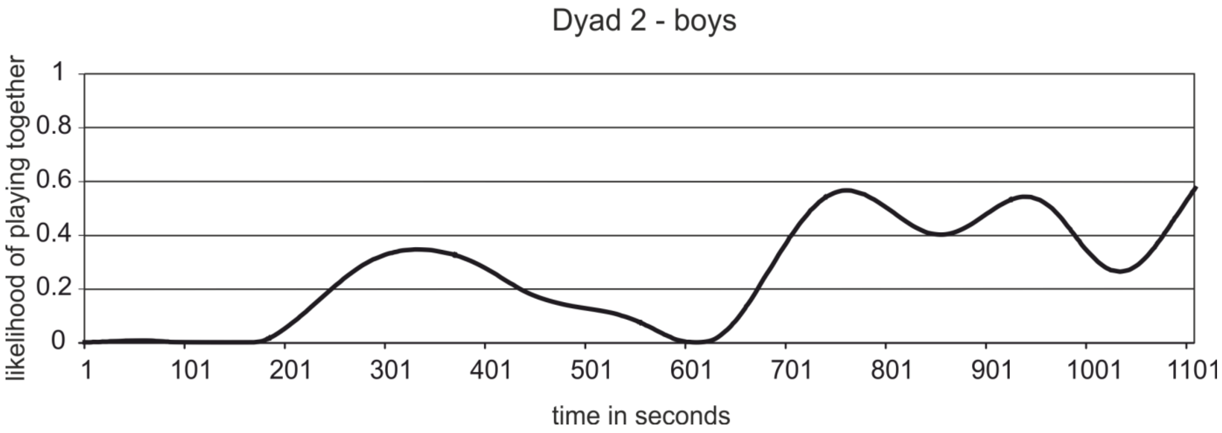

Dyad 2 - boys
likelihood of playing together
1
0.8
0.6
0.4
0.2
0
1
101
201
301
401
501
601
701
801
901
1001
1101
time in seconds

FIGURE 3

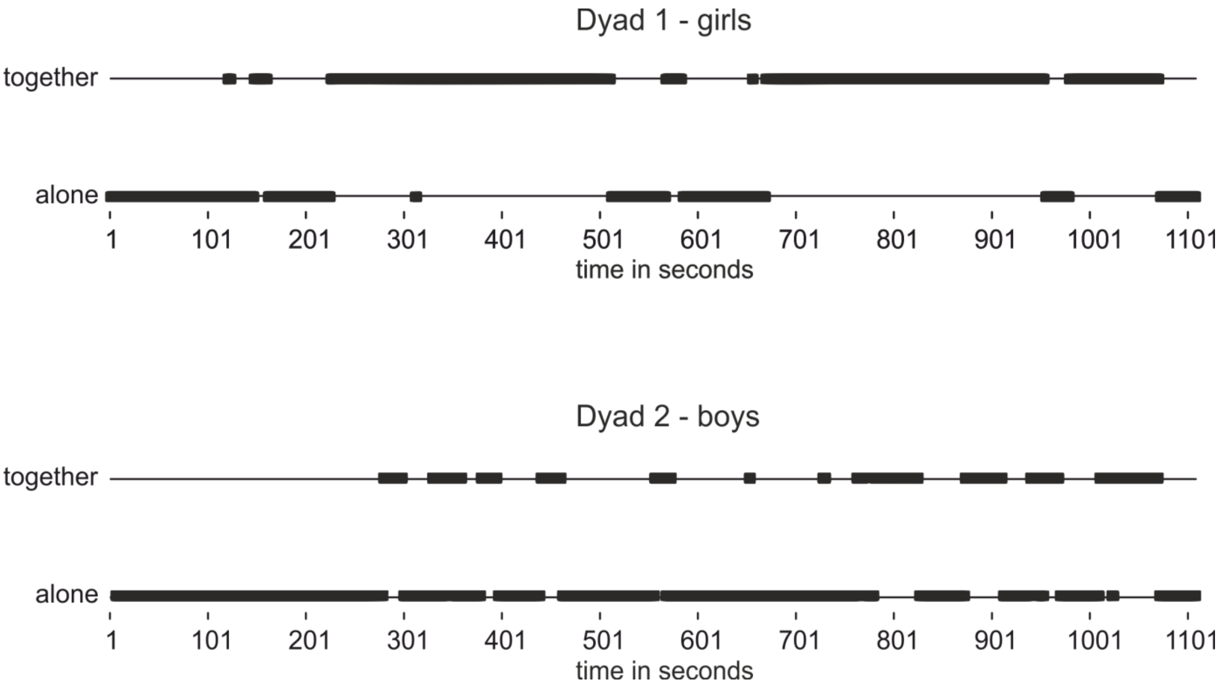

Dyad 1 - girls
together
alone
1
101
201
301
401
501
601
701
801
901
1001
1101
time in seconds
Dyad 2 - boys
together
alone
1
101
201
301
401
501
601
701
801
901
1001
1101
time in seconds

FIGURE 4

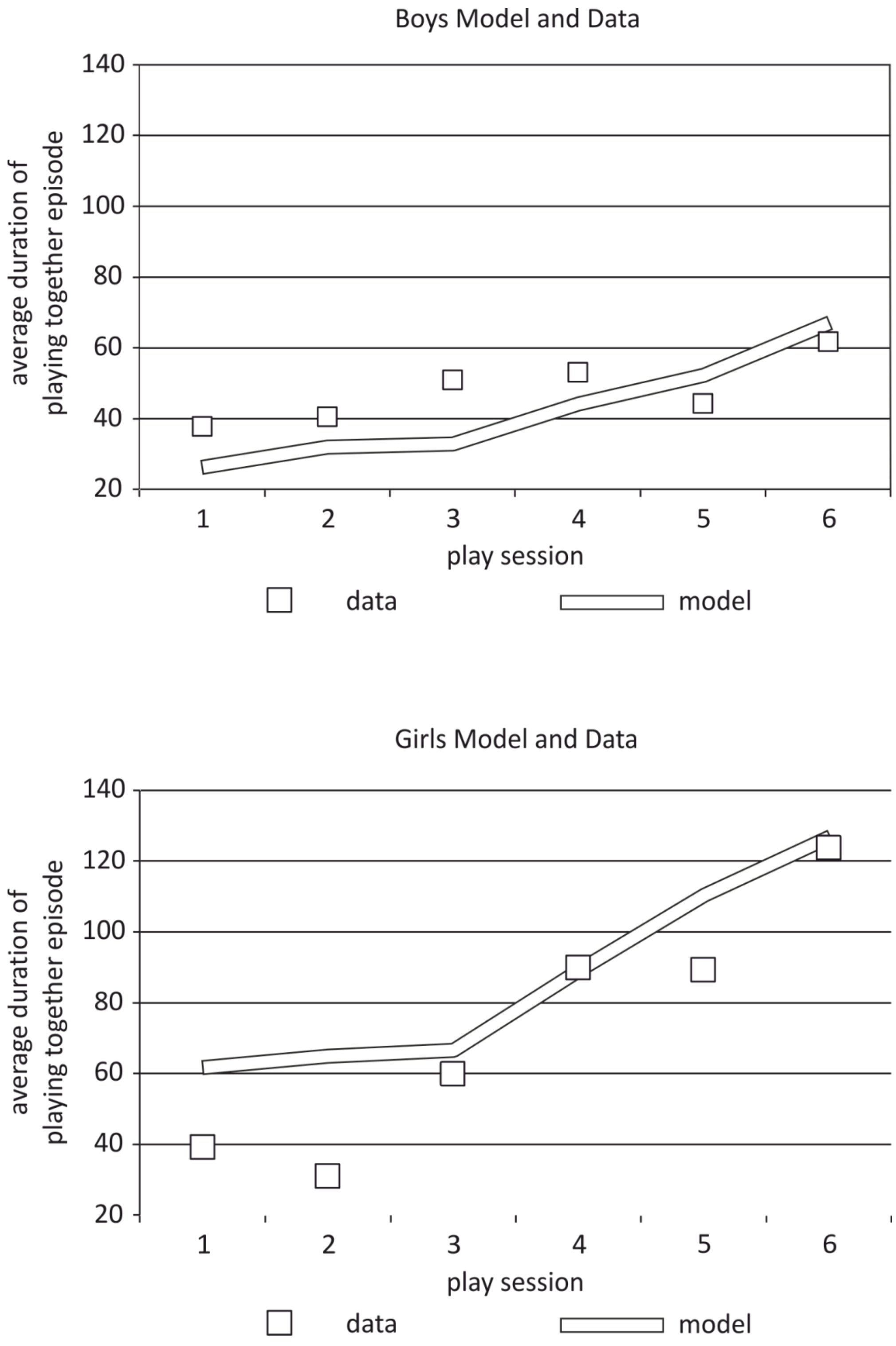

Boys Model and Data
average duration of playing together episode
140
120
100
80
60
40
20
1
2
3
4
5
6
play session
data
model
Girls Model and Data
average duration of playing together episode
140
120
100
80
60
40
20
1
2
3
4
5
6
play session
data
model

FIGURE 5

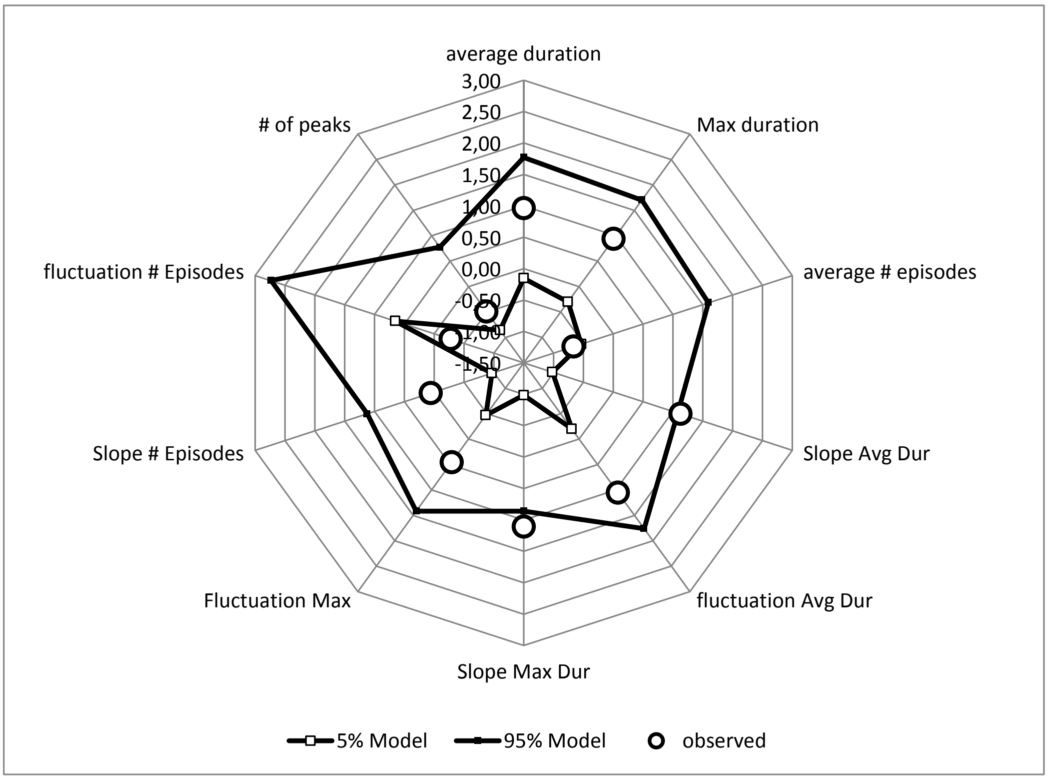

average duration
3,00
2,50
2,00
1,50
1,00
0,50
0,00
-0,50
-1,00
-1,50
Max duration
average # episodes
Slope Avg Dur
fluctuation Avg Dur
Slope Max Dur
Fluctuation Max
Slope # Episodes
fluctuation # Episodes
# of peaks
5% Model
95% Model
observed

FIGURE 6