Developmental Science

Developmental Science 11:2 (2008), pp 253–281 DOI: 10.1111/j.1467-7687.2007.00655.x

PAPER

An empirical validation of a dynamic systems model of interaction: do children of different sociometric statuses differ in their dyadic play?

Henderien Steenbeek and Paul van Geert

Department of Psychology, University of Groningen, The Netherlands

Abstract

Studying short-term dynamic processes and change mechanisms in interaction yields important knowledge that contributes to understanding long-term social development of children. In order to get a grip on this short-term dynamics of interaction processes, the authors made a dynamic systems model of dyadic interaction of children during one play session. The control parameters of the model relate to children's goal-directedness, concerns, emotional appraisals, social power, and social competence. Three groups of dyads of different sociometric statuses are represented by specific control parameter values. The model's order parameters consist of children's emotional expressions and other- versus self-directed actions. This article describes the empirical validation of the model and the methods needed for such validation. It focuses on the model's predictions of averages and distributions of the major variables, of the occurrence of attractors and power law distributions, and on the model's sensitivity. Overall, the model fits the empirical data well. In the discussion, we reflect on the developmental and methodological implications for explaining social interaction on the short-term as well as on the long-term time scale. In addition, implications for intervention and assessment are presented, in particular relating to the problem of rejection.

1. Introduction

Children develop a major part of their social behavioral repertoire in their interactions with peers (Rubin, Bukowski & Parker, 1998; Hartup & Laursen, 1999; Kindermann, 2003; Parke, Simpkins & McDowell, 2002; Ladd, 1999). Accordingly, the subject of peer interaction is thoroughly studied within the field of developmental psychology, including the link between parent–child and child–peer interaction, the emergence of friendship, differences between boys and girls, and the development of specific interaction skills, such as emotion regulation (Kupersmidt & Dodge, 2004; Contrerars, Kerns, Weimer, Gentzler & Tomich, 2000; Ladd & Le Sieur, 1995; Collins, Maccoby, Steinberg, Hetherington & Bornstein, 2000; Rubin *et al.*, 1998; Morales, Mundy & Crowson, 2005; Cicchetti, Ganiban & Barnett, 1991; Lopes, Salovey & Côté, 2005). Peer interaction implies change, both over the short term of a concrete interaction and over the long term, covering the development across childhood and adolescence. A main question for developmental researchers, how this short- and long-term change can be explained, will be answered by means of a dynamic systems model. The current article focuses on the description and empirical validation of this model.

1.1. Core aspects of the dynamics of interaction processes

A first important component in learning to establish and maintain satisfactory relations with peers consists of the child's social skills, which define the level of social competence, i.e. the child's effectiveness[1] in social interaction (Rose-Krasnor, 1997; Rubin *et al.*, 1998). Social effectiveness is defined as 'the ability to achieve personal goals in social interaction while simultaneously maintaining positive relationships with others over time and across situations' (Rubin *et al.*, 1998; Bierman, 2004). In general, social competence increases with age and individual differences become more apparent (Black & Logan, 1995; Asher & Parker, 1989).

A second important and intensively studied component in peer interaction is the child's social power (Reis, Collins & Berscheid, 2000; Forsyth, 1990). Social power is 'the possibility of inducing force on someone else' (see

[1] Social competence and social effectiveness are interchangeable terms. In the context of short-term interaction, we prefer to use the term effectiveness rather than competence, and reserve competence as a term for long-term characteristics.

Address for correspondence: Henderien Steenbeek, Department of Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands; e-mail: h.w.steenbeek@rug.nl

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

Lewin, in Bruins, 1999; French & Raven, 1959). One type of power is referent power, which is related to being the best liked member of the group, i.e. to having a popular status (Raven, 1992). Social power and power differences increase with age (Hawley, 2002), and so does the children's preference for interaction with high-power children, which are usually the children with a high sociometric status in the group (Parker & Gottman, 1989; Adler & Adler, 1998; Eder, 1985; Cillessen & Rose, 2005). Low-power individuals, children as well as adults, are particularly motivated to get along with the other person (Copeland, 1994; Dépret & Fiske, 1999). Lower-status children will thus tend to show more instrumental actions and more positive expressions when interacting with a higher-power child (Snyder & Kiviniemi, 2001).

Both social power and social competence are reflected by the sociometric status of a child within the group (Asher & Dodge, 1986; Asher, Markell & Hymel, 1981; Coie, Dodge & Coppotelli, 1982; Van Lier & Hoeben, 1991; Cillessen & Mayeux, 2004; Cillessen & ten Brink, 1991). Children with a popular status are attributed the highest social competence and power, and rejected children have the least social competence and power (Coie, Dodge & Kupersmidt, 1990; Newcomb, Bukowski & Pattee, 1993; Simeo-Munson, 2000; Eisenberg & Fabes, 1995; Hazen & Black, 1984, 1990; Hubbard, 2001; Asarnow & Callan, 1985; Coie, Dodge & Kupersmidt, 1990; Black & Logan, 1995; Rubin *et al.*, 1998; Black & Logan, 1995; Eisenberg, Fabes, Guthrie & Reiser, 2001; Denham, McKinley, Couchoud & Holt, 1990). Studies also report an association between adequate and emotionally positive interaction patterns with high social status on the one hand and less or inadequate and emotionally negative interaction patterns with rejected status on the other hand (Black & Logan, 1995; Asher, 1983; Asher & Parker, 1989; Rose-Krasnor, 1997; Gnepp, 1989; Edwards, Manstead & MacDonald, 1984; Vosk, Forehand & Figueroa, 1983; Gottman, Gonso & Rasmussen, 1975; Krantz, 1982; Eisenberg, Fabes, Bernzweig, Karbon, Poulin & Hanish, 1993; Eisenberg & Fabes, 1995; Cirino & Beck, 1991; Miller & Olson, 2000; Hubbard, 2001; Denham *et al.*, 1990).

How can the relation between competence, power, status, and interaction patterns be explained, both on the level of concrete interaction, and on the level of developmental time? They can be studied in the form of connections between variables over groups, but also by looking at the dynamics of interaction processes, which is the approach taken in the present article. Cillessen argues that the 'dynamic systems perspective is an important direction for future peer research' (2006, p. 49; see also Dishion, Bullock & Granic, 2002; Granic & Hollenstein, 2003). A dynamic approach can help solve two limitations in the present developmental research on competence, power, status, and social interaction in general.

1.2. Limitations of present research and potential solutions

A first limitation is that (peer-)context is either neglected, or not entirely adequately incorporated into current research methodology. An illustration of the relative neglect is the implicit assumption that children are primarily characterized by their individual child properties, e.g. their social skills that apply in every context. A statement like 'a child with a rejected status shows more disruptive behavior' (Newcomb *et al.*, 1993) may elicit the impression that the rejected child will always show this behavior more than any child with whom he or she is compared give or take some random variation. Although the role of (peer-)context is now increasingly acknowledged (Kupersmidt & Dodge, 2004; Parke & Ladd, 1992; Cillessen & Mayeux, 2004; Sandstrom & Zakriski, 2004), the current literature uses 'context' primarily in terms of 'distal', not as 'proximal' context. The distal context comprises aspects such as a child's relationships with key adults or the child's average position in the group over a longer time period. These aspects are literally distal in terms of time and space, whereas the proximal context is literally spatially and temporally close. In this study, we focus on the proximal context, which is a time-dependent and variable aspect of the dynamics of the interaction process itself. That is, the context emerges and evolves in the course of the interaction process. Any distal variable can be coupled to a concrete proximal context. For instance, sociometric status of a child can be treated as a general characteristic of this child, i.e. as a distal property, but also as a series of concrete and immediate determinants of the way a concrete interaction with another child unfolds, i.e. as a proximal property. Hence, context is a dynamic, not a static, characteristic of an interaction (see Gottman, Guralnick, Wilson, Swanson & Murray, 1997; for a distinction between static and dynamic models, see Howe & Lewis, 2005; Van Geert & Steenbeek, 2005). In this article we will argue that a dynamic context explains the occurrence of behaviors that differ from those expected on the basis of a static, i.e. distal, context.

A second limitation refers to the time scales covered by existing models. These models primarily focus on the possible predictive relationship between what children learn in their actual interactions with peers and their long-term social functioning (Parker & Asher, 1987; Kupersmidt, Coie & Dodge, 1990; Bagwell, Schmidt, Newcomb & Bukowski, 2001; Prinstein & La Greca, 2004; Bierman, 2004; Kupersmidt & Dodge, 2004). In our view, a comprehensive theory of interaction must, first, present an explicit account of the actual, short-term co-regulated process of interaction (Fogel, 1993). Second, it must explain what it is that changes over long-term developmental time, and how this change happens. Third, the theory must be able to specify the relationship between both time scales. The underlying idea is that all levels of the developing system interact with each other

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

and consist of nested processes that unfold over many time scales, from milliseconds to years (Thelen & Smith, 1994; Lewis, 2002). Crick and Dodge's (1994) widely used social information processing (SIP) model describes the short-term processing of actually given information through cognitive schemas leading to specific interactions. They implicitly link this with long-term development, for instance via changes in the children's memory and schemas, thus showing 'how children's social cognitions may change as a function of their social experience' (Kupersmidt & Dodge, 2004, p. 63). Kupersmidt and deRosier link a short-term model, which is highly similar to the SIP model, to Coie's conceptual long-term model of rejection (Asher & Coie, 1990). They add a wide array of mediational factors between peer experiences and adjustment (Kupersmidt & Dodge, 2004, p. 123), such as social experiences, cognitive mediators (cognitions about the self and cognitions about others), and contextual factors (e.g. interpersonal factors). Kupersmidt and deRosier plead for an integrative approach that 'attempts to understand not only how each process impacts on outcomes, but also how the processes work together to either promote or deter future negative patterns' (Kupersmidt & Dodge, 2004, p. 133). In our view, a limitation of these otherwise very interesting models is that only by means of a recursive, dynamic application of the mechanisms described can it be tested whether they indeed offer a good explanation of the observed behavioral and affective short-term outcomes, and of the long-term consequences of these outcomes (see Reis *et al.*, 2000, p. 852, for a comparable view).

In order to overcome these limitations, the current article will present an explicit recursive application of a model of short-term dynamics of interaction processes. The theoretical foundation of the model is presented in Steenbeek and van Geert (2007). The goal of the model is to explain the emergence of different patterns of social interaction on the level of the concrete time course. This goal can only be accomplished if the model proves to be empirically valid. In our view, a highly suitable context for validation is interaction among children who differ in social power and social competence, and, more precisely, who are likely to differ in sociometric status.

Two research questions follow from this aim; namely, first, what are the tools and steps needed for empirically validating this dynamic model? Second, how good is the model in representing the interaction process in reality? The answer to these two questions is conditional on answering the following question: how can the theoretical principles behind the model be applied to dyadic play of children of different sociometric status, and in addition, how can the principles be applied to long-term development? With regard to the latter, we will focus on the question of how rejection transpires. The structure of the article is as follows. This introductory section will be continued with a general description of the theory and the model, based on dynamic systems theory, and describe how the theoretical aspects are implemented in the form of model parameters, aiming to explain differences in social interaction among children with different sociometric status. We will then proceed with a discussion of method (empirical data collection, model building and simulation) and results, which for reasons of clarity are presented together. In the discussion, we will reflect on the developmental and methodological implications for explaining social interaction on the short-term as well as on the long-term time scale. In addition, implications for intervention and assessment are presented, in particular relating to the problem of rejection.

1.3. A dynamic systems approach to dyadic interaction

1.3.1. Four principles of a general theory of social interaction

The theory that we propose is based on four general principles. The first principle is that social behavior is intentional from an early age, i.e. it is aimed at the realization or pursuit of goals or intentions (Austin & Vancouver, 1996; Carver & Scheier, 1990; Powers, 2005). Many goals or intentions are largely unconscious and emerge under control of the context (Bargh & Chartrand, 1999, p. 468; see also Austin & Vancouver, 1996; Bargh & Ferguson, 2000). Children develop an understanding of the goal-directedness of their own and other people's actions during the first year of life (Beilin & Fireman, 2000; Flavell, 1999; Piaget, 1936; Tomasello, 1995; Tomasello & Rakoczy, 2003; Woodward, Sommerville & Guajardo, 2001). The pursuit of goals is closely related to the structure of the child's peer group and in particular to emerging differences in dominance of the group members.

The second principle is that goals represent interests or concerns, as Frijda (1986) calls them. The concern aspect implies that organisms automatically evaluate situations in function of their goals. That is, they evaluate whether the situation is good or bad (Scherer, 1999). Emotions play the role of immediate evaluations of the value of a situation with regard to the person's goals (Arnold, 1961). Pleasure and joy are important evaluative emotions, in particular in approach-directed situations such as play (Cabanac, 1992; Johnston, 2003; Panksepp, 2000; Roseman & Evdokas, 2004; Roseman, Wiest & Swartz, 1994). Appraisal of the value of contexts is a biologically fundamental system, already functional at a very early age (Camras, Meng & Ujiie, 2002; Messinger, 2002). Also from an early age, emotional appraisal is a socially transparent signal of one's evaluation of the situation in function of one's goals (Grolnick, Bridges & Connell, 1996; Murphy & Eisenberg, 2002).

The third principle views social interaction as a goal in itself (Austin & Vancouver, 1996). Children will prefer some people over others. This difference in preferences

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

for certain people emerges already at an early age, as the attachment literature shows (Ainsworth, 1979; Bowlby, 1980, etc.). By the age of 4 to 5, children have developed a pattern of differential proximity and interaction concerns, which can be related to sociometric status (Martin, Fabes, Hanish & Hollenstein, 2005; Rubin *et al.*, 1998). The preference for social play as a particular form of pleasurable social interaction increases with age and is well consolidated in children between 5 and 8 years of age (Lyytinen, 1991; Goncu, 1993).

The fourth principle entails that behavior, including emotion, is also deeply affected by a non-intentional component, namely the tendency of people in social interaction to automatically copy or mimic the behavior and emotions of the other person, relatively regardless of their own goals (Levy & Nail, 1993; Nail, McDonald & Levy, 2000; Wheeler, 1966; Neumann & Strack, 2000; Rizzolatti & Craighero, 2004). Contagiousness can be highly functional in that it contributes to effective behavior coordination between members of a group (Chartrand & Bargh, 1999). There is ample evidence that imitation is a fundamental biological tendency, already shown by infants and young children (Preston & de Waal, 2002; Thompson & Russell, 2004; Gergely, Bekkering & Király, 2002). Literature on social learning and modeling (Bandura & Walters, 1977) has shown that the tendency to imitate or mimic another child becomes increasingly coupled to the social power and status of the imitated child. These four principles form the foundation of our dynamic systems model of interaction.

1.3.2. Properties of a dynamic systems approach

A dynamic system is a means of describing how one state develops into another state over the course of time (Weisstein, 1999) and often produces complex, nonlinear behavior over time (Thelen & Smith, 1998). Dynamic systems must be studied as processes over time, not as associations between distributions of variables over populations. A dynamic system can be described in the form of a characteristic equation, namely

$$Y_{t+1} = f(y_t) \tag{1}$$

i.e. the value of y at time $t + 1$ is a function 'f' of the value of y at time t. The change in the value is a function of the variable's current value. Thus, a dynamic systems model of a social interaction process of children is an explicit prescription (f in the equation) of how the current state of the interaction process evolves into another state, at the next moment in time (for more details, see van Geert & Steenbeek, 2005). The basic equation is recursive (or with a synonym, iterative), i.e. it transforms y_t into y_{t+1}, y_{t+1} into y_{t+2}, and so on. The series of successive y's forms the explicit description of a process. The recursiveness of the equation provides a natural interpretation for the notion of time scales. That is, a sequence of iterative steps produces the time trajectory of a particular type of process, e.g. a play session. This process can then be used as a single step in a different dynamic model that iterates many such processes to explain the developmental changes in play, for instance. The first dynamic model thus describes the short-term dynamics, the second describes the long-term dynamics.

In many cases, the change in one variable will be related to the change in another variable, and vice versa, which is characteristic of coupled systems. For instance, in addition to affecting one's own next action, the action of one child (represented by y) also affects the reaction of the other, and the reaction of the other (represented by x) affects the consequent reaction of the first, represented mathematically as

$$y_{t+1} = f(y_t, x_t); \quad x_{t+1} = g(x_t, y_t) \tag{2}$$

Coupled systems provide a natural definition of context, more precisely dynamic context, because the dynamic context of y is x, and the context of x is y.

The actions of both children separately form an example of intertwining factors. Such intertwining processes are often characterized by nonlinear change, self-organization, and the existence of attractor states, in principle also under the influence of chance or stochastic factors (Thelen & Smith, 1994; van Geert, 1994). Note that the intertwining of multiple factors in such a process can only be understood by modeling it in the form of a dynamic, iterative process (Christiansen & Kirby, 2003).

In order to obtain a grip on the characteristic functioning of the system, it is important to distinguish the major control and order parameters. Control parameters are parameters 'to which the collective behavior of the system is sensitive and that moves the system through different collective states' (Thelen & Smith, 1994, p. 62). Control parameters are thus properties of the functions f and g in equation 2. Collective behaviors are forms of coordination of all the elements of the system. Order parameters are ways of describing and distinguishing different forms of collective states, i.e. 'dominant modes' (Haken, 1977; Thelen & Smith, 1994, p. 55). For instance, the order parameter 'emotional expression' allows us to make a distinction between smiles and anger in the human facial system. A smile, or anger for that matter, is an example of a collective state or 'dominant mode' of the facial system. It entails many subordinate variables, such as the movements of many muscles in the (smiling) face. Several factors, such as the level of certain hormones or events in the environment, can function as a control parameter of emotional expression. That is, they affect the ease with which emotions are expressed (Camras, 2000; Lewis & Granic, 2001). The variables y and x in equation 2 are in fact sets of order parameters describing the behavior of the system.

Figure 1 gives an overview of order parameters and control parameters on both short term and long term,

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

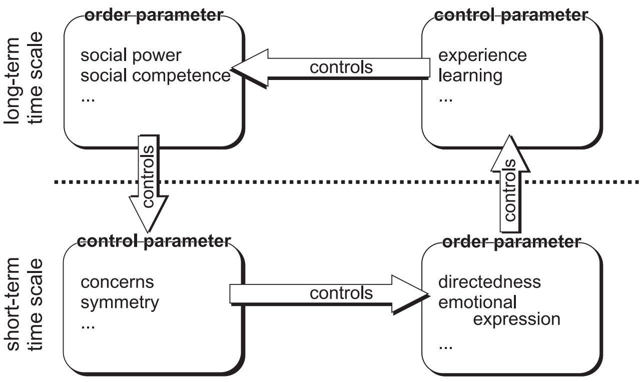

Figure 1 *An overview of order parameters and control parameters on both short term and long term, as distinguished by our model of interaction.*

as distinguished by our model of interaction. Notice that this model of social interaction departs from the perspective of an individual child. The basis is constituted by two order parameters, namely 'directed actions' and 'intensity of emotional expressions' of individual children. Changes in these short-term order parameters, e.g. a succession of different emotional expressions, are regulated by short-term control parameters. An example of a short-term control parameter is the 'social concern' (or 'goal') of a child in a particular situation, which co-determines the child's social actions and emotional expressions. We have also seen that concerns are determined by the sociometric status (and thus the social power) of the play partner. Social power is an example of a long-term order parameter, i.e. a property that changes over developmental time. To put it differently, a long-term order parameter, for instance 'social power', controls a short-term control parameter, for instance a 'child's social concerns or goals'. The question is, what controls the change in the long-term order parameters, such as 'social power'? We assume that such change takes place as a consequence of experiences with peers. This implies that experiences with peers as they accumulate over time are long-term control parameters of the long-term order parameter power. And finally, what controls the experiences? The experiences are dependent on the concrete interactions with others, such as the amount of action directed towards the other child during play sessions. Directedness is a short-term order parameter of interaction. Thus, short-term order parameters control long-term control parameters.

Note that the model in Figure 1 confines itself to psychological processes within individuals, individual behavior, and the resulting dyadic interaction between individuals. In accordance with Hinde's model of social complexity (1997), these elements can be conceived of as the basis of all kinds of interaction processes.

1.3.3. Testing a dynamic systems model

We have chosen to validate the model with empirical data, in which sociometric status of children is used as a dividing criterion. The theoretical justification for this choice is, first, that children of different sociometric status are likely to differ with regard to important fundamental long-term order parameters, namely competence and power. Second, in sociometric status, which is a property that characteristically changes over the long term, two aspects of short-term social interaction are incorporated, namely social preference and frequency of interaction. These aspects play a central role in the real-time interaction process: how pleasant is the interaction with the other child (relating to preference), and how much is the child inclined to interact with the other child (relating to frequency)? A concept like status has important meaning in an empirical and applied context. For instance, by using the model of short-term interaction, we hope to make a contribution to answering an important long-term question, namely how rejection transpires.

Validating the model requires, first, that the theoretical principles are transformed into a dynamic model. Second, empirical predictions are generated by using the model to simulate many short-term interactions with model parameters relating to sociometric status. Third, these predictions are tested against empirical data.

What are the criteria for a good model and a good empirical fit? An important general criterion for a good model is that it 'provides convincing answers to the questions we put to it' (Casti, 1997, p. 25) and that it generates new, testable questions (Gottman, Murray, Swanson, Tyson & Swanson, 2002). Put more concretely, it means that, first, the model must be based on valid theory(-ies), with valid definitions of behavior. Second, the model must technically 'behave well', for instance not showing chaotic patterns, if in reality the represented system does not either. Finally, the model must show sufficient similarity in its output compared to results of empirical observations of the process (Balci, 1997; Van Dijkum, DeTombe & van Kuijk, 1999; Gilbert & Troitzsch, 1999). The question is of course: what is 'sufficient similarity' and 'comparable output'? This question will be answered in the method section.

1.4. The dynamic model of dyadic interactions

The simulation model is based on the theoretical principles described earlier. The mathematical elaboration of the model and program code are explained in more detail at the website www.gmw.rug.nl/~model. The simulation model is a combination of a dynamic systems model and an agent model in which two agents (two children) interact with each other. Examples of other agent models are described in Gilbert and Troitzsch (1999), Jager (2000), Jager, Popping and van de Sande (2001), Kohler and Gumerman (2000), Staller and Petta (2001), Conte and

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Table 1 *An overview of the groups of control parameters, the specific parameters that are distinguished, and what can be adjusted*

Input parameter groups	Which parameters are distinguished	What can be adjusted?
1 Concerns group	Involvement Autonomy	Strength of the concerns in relation to each other
2 Realizability group (influence of behavior on realization of concerns)	Influence of 'playing together' Influence of 'playing alone'	Strength of the influence of behavior on the realization of a concern
3 Expressiveness group (relation between emotional appraisal and emotional expression)	Positive expression Neutral expression Negative expression	Ease with which emotional appraisal is translated into an emotional expression
4 Preference group (influence of emotional expression on preference of concerns)	Influence of positive expression Influence of negative expression	Strength of the influence of emotional expression on the preference of a concern
5 Non-intentional behavior group	Continuity Symmetry	Strength of the non-intentional principles in relation to each other

Paolucci (2001), and Conte, Hegselmann and Terna (1997). For examples of other dynamic systems models of dyadic and group interaction, we refer to Buder (1991), Gottman *et al.* (1997), Felmlee and Greenberg (1999), Felmlee (2007), Warner (1992), Olthof and Kunnen (2000), van Geert (1994), and Steenbeek and van Geert (2005a, 2007). In addition, for more general information about the use of dynamic systems models in social developmental psychology in the broad sense of the word, we refer to van Geert (1994, 2003), Valsiner (1998), Granic and Hollenstein (2003), Fogel (1993, 2001), Lewis (1995, 2000a, 2002), Lewis and Granic (2000) and Dishion *et al.* (2002).

The model uses a number of control parameters to specify the operation of dynamic rules.

1.4.1. Control parameter groups

The theoretical assumptions described above are reflected by a number of short-term control parameters. These control parameters comprised five parameter groups, presented in Table 1.

The first and most important parameter group relates to *concerns*, more specifically one that refers to a specific, maximally pleasurable balance between the strength of a concern named 'Involvement', which is the tendency to direct one's behavior towards the other person, and the strength of a concern named 'Autonomy', which is the tendency to perform a solitary action.[2]

The second group, or *realizability* group, represents the influence of behavior on the realization of concerns. This group determines the appraisal of events, i.e. to what extent does a particular event contribute to realizing the child's concern for either Involvement or Autonomy?

The third parameter group, or *expressiveness* group, represents the strength of the relation between emotional appraisal and emotional expression, for instance the ease with which a child shows a positive emotional expression, if the appraisal is positive.

The fourth parameter group, or *preference* group, represents the influence of emotional expression of oneself or the play partner on one's Involvement–Autonomy balance (the concern aspect). For instance, the parameter specifies to what extent a positive expression that accompanies an event of 'playing together' makes playing together (the Involvement aspect) more desirable or pleasurable for this child.

The fifth – and last – parameter group refers to *non-intentional* basic principles of behavior. It represents the preferred balance between the tendency to continue one's own behavior (continuity), and the opposite tendency, namely to do what the other person is doing (symmetry, which depends on the contagiousness of the behavior of the other person).

During each model run, the parameter values are stochastically varied within preset limits.

1.4.2. Order parameter groups and dynamic evolution rules

The values of the control parameter groups moderate the way in which the model's evolution rules determine the short-term change in order parameters. The first order parameter concerns the child's other- versus self-directed actions. The second order parameter concerns the intensity of emotional expressions, which range from -4 (very negative) to 5 (very positive). The evolution rules comprise a number of steps as represented in Figure 2.

The series of process steps starts with the concerns of the participants, i.e. the children's preferences for directedness towards the other child ('involvement') or directedness towards oneself ('autonomy'). At each moment in time, the difference between a child's preferred and realized value of each concern results in a certain strength of drives to enact either the involvement or the autonomy behavior. The concern with the highest drive will determine the actual behavior. Via an appraisal function, the average of the drives will result in an emotional expression. The likelihood of positive expressions

[2] Note that by focusing on these two concerns in the concerns parameter group, we aim to concentrate on fundamental parameters in the interaction process. It does not mean that we deny the existence of other concerns in the process.

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

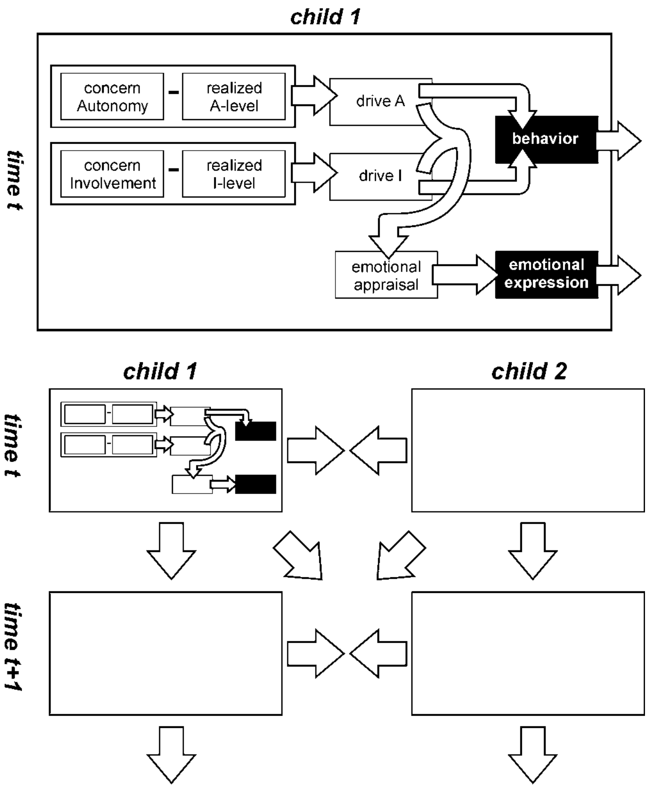

Figure 2 *The evolution rules of the two order parameters; the child's other- versus self-directed actions and the intensity of emotional expressions.*

increases as the concerns come closer to their preferred values (and the other way round for the negative emotional expressions).

The connection between two child-specific processes at two successive moments takes place via a dyadic process that runs as follows. First, the directedness at time t of a child has an influence on the realized level of the child's and the play partner's accompanying concern at time $t + 1$. Second, the emotional expression at time t of a child has an influence on the preferred level of the child's and the play partner's accompanying concern at time $t + 1$. For instance, assume that at time t the child is directed towards the other child, enjoys it and shows it with a big smile. This implies that, for time $t + 1$, the realized level of the concern Involvement is increased (through the action itself) and the preferred level of the concern Involvement is increased (through the accompanying positive expression). The updated preferred and realized levels of the concerns will determine the likelihood of directedness and emotional expression of both children at time $t + 1$.

These child-specific and dyadic processes take place during and between each successive moment, i.e. each simulation step. If a step corresponds with 1 second, a 7-minute play session is simulated by means of 420 iterative steps. For a detailed description of these processes, we refer to Steenbeek and van Geert (2005a).

In order to be able to compare the model's output with empirical data, we used 22 operational variables that describe all potentially relevant properties of the time evolution of the order parameters directedness and emotional expressions (see Table 2).

Table 2 *Operational variables, such as derived from the model output and derived from the empirical data*

Variables	Description
Child or play partner	
Directedness	proportion of directed actions ('playing together') over the total of all actions (both 'playing together' and 'playing alone') of the child
Proportion shared directedness	proportion of directed actions of this child, accompanied by a directed action of the play partner
Positive expressions[a]	proportion (percentage of time) of positive expressions over the total number of expressions (neutral, negative or positive expressions)
Intensity positive time[a]	proportion of total average intensity of positive expressions
Intensity positive number	proportion of intensity of positive expressions divided by the amount of positive expressions
Proportion shared positive	proportion of positive expressions of this child, accompanied by a positive expression of the play partner
Negative expressions	proportion (percentage of time) of negative expressions over the total number of expressions (neutral, negative or positive expressions)
Intensity negative time	proportion of total average intensity of negative expressions
Intensity negative number	proportion of intensity of negative expressions divided by the amount of negative expressions
Dyad	
Coherence dyad	proportion of time that both children show directed actions ('playing together') of the total time of the play session
Shared positive expressions[b]	proportion of shared positive expressions over the total number of expressions. This variable can be read as a measure for 'coherence of positive expressions'
Shared negative expressions	proportion of shared negative expressions over the total number of expressions. This variable can be read as a measure for 'coherence of negative expressions'
Contrast dyad	proportion of contrast in intensity of expressions of both children over the total time. The time that both children express a neutral expression is not included (coded as zero)

Notes: [a] Concerning variables 'positive expressions' (1) and 'intensity positive time' (2): in (1) the intensity is not calculated, in (2) the proportion in relation to intensity.
[b] Concerning variable 'shared positive expression' (of the dyad): a high level of shared positive expression does not necessarily imply a high level of positive expressions *per se*.

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

The operational variables refer to the duration of directed actions (five variables) and the duration or intensity of expressions (17 variables) of each child separately and of the dyad. An example of an action-related variable is the variable 'directedness'; i.e. the proportion of actions directed towards the other child, over the total of all actions. If at time *t* both children show directedness simultaneously, *t* is coded as 'coherent'. An example of an expression-related variable is 'positive expression', which is the child's proportion of positive expressions over the total of all expressions (positive, neutral, and negative; see Table 2, third row). Subsequently, these operational variables can be averaged over a number of model runs of a particular dyad group. These averaged operational variables can be compared with our empirical data, since in the latter we also coded the child's emotional expressions and directed actions.

1.4.3. Research questions: How does the model explain differences and similarities in social interaction of children of different sociometric status?

A child's sociometric status covaries with the long-term order parameters social competence and social power. These long-term order parameters determine the values of the five control parameter groups. A child with a specific sociometric status, in the context of playing with another child with its own particular properties, can thus be represented by means of specific settings of the five control parameter groups.

Table 3 shows the control parameter settings for the participants of three types of dyads.[3] They consist either of a child with a rejected status playing with an average play partner (the 'rejected' dyad); a child with a popular status playing with an average play partner (the 'popular' dyad), or a child with an average status playing with a play partner with an average status (the 'average' dyad). The default settings, referring to average social competence and average social power, are those that apply to the average child, playing with an average play partner. These default values have been obtained by estimating the best possible set of model parameters, based on a qualitative comparison with the empirical data from 'average' dyads. The default setting for the concerns-parameter group is that Involvement is stronger than Autonomy. In addition, actions directed towards the other child and actions directed towards oneself have an average influence on the realization value of the corresponding concern (realizability group of control parameters).

Note that for reasons of simplicity these quantitative parameter settings are specified in qualitative terms (such

Table 3 *Settings of input parameter groups of children of different sociometric statuses in the context of playing with a play partner*

Type of dyad	Rejected dyad		Average dyad		Popular dyad	
Status of child	Rejected		Average		Popular	
Status of play partner		Average		Average		Average
1 Concerns group	I *much stronger* than A	I *little bit stronger* than A	I *stronger than* A	I *stronger than* A	I *little bit stronger* than A	I *much stronger* than A
2 Realizability group	I = average	I = *high*	I = average	I = average	I = *high*	I = average
	A = *high*	A = average	A = average	A = average	A = average	A = *high*
3 Expressiveness group	positive = moderate	positive = moderate	positive = moderate	positive = moderate	positive = moderate	positive = moderate
	negative = *difficult*	negative = moderate	negative = moderate	negative = moderate	negative = moderate	negative = *difficult*
4 Preference group	positive = *big*	positive = average	positive = average	positive = average	positive = average	positive = *big*
	negative = *big*	negative = average	negative = average	negative = average	negative = average	negative = *big*
5 Non-intentional behavior group	continuity = average	continuity = average	continuity = average	continuity = average	continuity = average	continuity = average
	symmetry = *high*	symmetry = *low*	symmetry = average	symmetry = average	symmetry = *low*	symmetry = *high*

Notes: *a*. The table can be read as follows: a child with a rejected status that plays with a play partner with an average status (upper left of the table) has as setting for the first input parameter group Concerns that his Concern Involvement is much stronger than his Concern Autonomy, which is expressed as 'I *much stronger* than A'.
b. I = Concern Involvement, A = Concern Autonomy. The average dyad has default settings (average, moderate). Everything that differs from these default settings is printed in *italics* (*high, low, difficult, big*).
c. Information about the corresponding values of these settings of input parameter groups for children of different statuses in the context of playing with a play partner with a different status, can be found at the website www.gmw.rug.nl/~model

[3] This is in accordance with the design of the broader study that we conducted in collaboration with the University of Utrecht. The main goal of this collaborative project was to study differences in interaction of popular and rejected children in a standardized play context, with dyads composed as described above (see for instance Gerrits, 2004).

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

as 'average', 'stronger than', etc.). The accompanying numerical values can be found in the parameter worksheet of the simulation model which is available in the form of an excel-file at website http://www.gmw.rug.nl/~model (see also Steenbeek & van Geert, 2005a, 2007).

The rejected child in the 'rejected' dyad and the average child in the 'popular' dyad have similar parameter settings (for a mathematical justification, see the web materials at http://www.gmw.rug.nl/~model). That is, their concern Involvement is much stronger than their concern Autonomy (first control parameter group), the realizabilty of Autonomy is high (second group), the likelihood of negative expressions is low (third group), the preference effect of both positive and negative expressions is big (fourth group) and the tendency to symmetry is high (fifth group). This means for the rejected child that he has a high concern for actions directed towards the other child (Involvement). In addition, he is easily satisfied with 'actions directed towards itself', meaning that solitary play is easily satisfied (because it is not particularly rewarding in comparison to playing with the higher-power play partner) and tends to suppress negative emotional expressions. In addition, the play partner's positive expressions during actions directed towards the rejected child, which will usually refer to playing together, will strongly increase the child's preference for other-directed actions. Finally, his tendency to imitate the behaviors of the play partner is strong.

The average child in the 'rejected' dyad and the popular child in the 'popular' dyad also have similar parameter settings. Their concern for Involvement is a little bit stronger than the concern Autonomy, the realizabilty of Involvement is high, and the tendency to symmetry is low.

Recapitulating, our research questions are first, what are the tools and steps needed for empirically validating this dynamic model, and, second, how good is the model in representing the interaction process in reality? The answers are conditional on the following question: how can the theoretical principles behind the model be applied, first, to the short-term dynamics of dyadic play of children of different sociometric statuses, and, second, to long-term development of status and interaction?

2. Method and results

2.1. Empirical data

In this section, we will present the main lines of our empirical study (for a more elaborate discussion, see Steenbeek and van Geert, 2007).

2.1.1. Participants

Twenty-four dyads of grade 1 pupils with mean age of 6.5 years participated. They were selected on the basis of their sociometric status. Each dyad consisted of two same-sex children (12 female and 12 male). Three types of dyads were formed: the first consisted of a child with a rejected status, coupled with a neutral play partner with an average status ('rejected' dyads); the second consisted of a child with a popular status, and an average play partner ('popular' dyads); and the third of a child with an average status, and an average play partner ('average' dyads).

2.1.2. Procedure

First, the sociometric status of the participants was determined on the basis of repeated measures of a rating test (Asher, Singleton, Tinsley & Hymel, 1979), which was analyzed with the computer program SS-rat (Maassen, Akkermans & van der Linden, 1996; the stability of the procedure is discussed in Maassen, Steenbeek & van Geert, 2004; Steenbeek & van Geert, 2005b). Second, the dyads were videotaped during a 10-minute play session in a separate room in the school. The only instruction was to play together with four groups of toys that were placed on the table. After giving the instruction, the researcher left the room, leaving the children alone with the toys and the camera.

The 24 dyads were videotaped three times, with intervals of approximately one and a half months. In principle, the second and third round were selected for coding. Due to practical limitations, 17 dyads were coded twice and 7 dyads were coded once. This resulted in a total of 41 coded interactions ('rejected' dyads; n = 13, female/male ratio 11/2; 'average' dyads; n = 14, f–m ratio 6/8; 'popular' dyads; n = 14, f–m ratio 5/9; the small sample sizes are explained by the fact that popular and rejected children form a small minority of the total number of children in the class). The sample does not satisfy the assumption of independent measures, i.e. independent subjects or dyads. The question is whether this interferes with making statistically reliable statements about relevant group differences. We haven chosen to bypass the problem by using statistical tests of group differences that do not depend on the assumption of independent measures, namely random permutation tests (Manley, 1997; Good, 1999). In addition, we checked whether the differences found with this sample are similar to differences found in a sample of independent measures, consisting of the first available observation of each dyad. We found the same differences, i.e. differences in the same direction and with comparable p-values. We chose to use the original sample of 41 dyads, because this greater number of cases has advantages with regard to testing group differences by means of distributions and counting characteristic dynamic features of the interaction patterns.

The female–male ratios above show that there is an overrepresentation of girls in the rejected group of dyads. However, there is no statistically significant gender difference in the two most important variables (the p-values are .86 and .58 for the expression measures of

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

child and partner, and .94 and .25 for the action measures, respectively). We concluded that for the present testing of the model, gender does not need to be taken into account as an additional explanatory variable.

2.1.3. Coding and variables

The recordings were coded with the computerized system Observer 4.0 pro (Noldus Information Technology, 1999). The inter-observer reliability between the observers was determined with a nonparametric permutation test (see section 'Random permutation analysis' for more information about permutation techniques; see also van Geert & van Dijk, 2003). The reliability was determined in advance and can be considered good in terms of percentage agreement (.8 for coherence, p = .01, .81 for expressions, p = .01).

Two order variables were coded, namely emotional expressions and instrumental actions of each child separately, with a precision of a second. The variable emotional expression was coded on a −4 to +5 scale. Categories −4 to −2 represented negative expressions, −1 to +1 neutral expressions, and +2 to +5 positive expressions. The variable action was coded with the help of three overt variables, namely verbal turn, nonverbal turn and focus. On the basis of these partial variables, a child's current behavior is coded as action directed towards the other child or directed towards oneself. If the child displays neither a verbal turn, nor a nonverbal turn, nor a focus (towards the play partner or the mutual play activity), the child is supposed to display self-directed actions ('playing alone'). Otherwise the child is coded as displaying other-directed actions. If both the child and the play partner show mutually responsive directed actions, the behavior is coded as 'coherence in dyad', which is the only action variable on the dyad level.

The coded order variables were transformed into operational variables as described in Table 2.

2.1.4. Results[4]

What follows is a short summary of the findings presented in Steenbeek and van Geert (2007). First of all, the total number of 22 variables was reduced to a subset of 10 core variables. Over the total pattern of this selection, we found significant differences between 'rejected' dyads and 'popular' dyads (X^2 = 8.6, p = .001).

Second, by examining each of the 22 variables separately, we found statistically significant differences between 'rejected' dyads and 'popular' dyads in seven variables. In six of these variables, the 'rejected' dyads scored significantly higher than the 'popular' dyads, namely in the child's (other-)'directedness', 'positive expressions', and 'intensity positive time'; in the play partner's 'proportion shared directedness' and 'intensity negative number'; and finally in the dyad's 'coherence'. Note that the findings concerning positive expressions and other-directedness run against the expectation based on the literature that popular children show more positive expressions and other-directedness than rejected children (for expectations based on the literature, see the Introduction section).

In only one variable of this set of seven, namely the child's 'proportion shared positive', did the 'popular' dyads scored significantly higher than the 'rejected' dyads. This finding is consistent with the literature, which points to the existence of more 'mutuality' in interactions in which a popular child is involved.

We did not find significant differences between 'rejected' dyads and 'popular' dyads in ten variables, namely the child's 'proportion shared directedness', 'intensity positive number', and 'intensity negative number'; the play partner's 'directedness', 'positive expressions', 'intensity positive time', 'intensity positive number' and 'proportion shared positive'; and the dyad's 'shared negative expressions', and 'contrast child–play partner'. This means that contrary to what has been reported in the literature about comparable variables, we did not find differences for 'rejected' and 'popular' dyads. For instance, we did not find that the play partner of a popular child directs his actions more often to the popular child than the play partner of a rejected child directs his actions to the rejected child.

Finally, five variables showed a trend[5] (with p-values between .1 and .3). These variables were the child's 'negative expression' and 'intensity negative time'; the play partner's 'negative expression' and 'intensity negative time'; and the dyad's 'shared positive expressions'. This is again not consistent with the literature. The trend found in our data is that popular children show more 'negative expressions' than rejected children; and popular dyads show less 'shared positive expressions' than rejected dyads, whereas the literature reports more negative expressions in rejected children, and more positive expressions in popular children. In addition, the trend found in our data is that play partners of popular children show less 'negative expressions' than play partners of rejected children.

The discrepancy between our findings and the literature is likely to be explained as follows. First, we used a process model to make predictions, which can easily differ from those that depart from child-specific factors found by calculating averages over many interactions.

[4] Results over the 'average' dyads are not reported because analyses reveal that in all variables but one, the mean scores of the 'average' dyads lie in-between the mean scores of the 'popular' dyads and the 'rejected' dyads. The one exception is the variable 'intensity negative number' of the child, in which the difference is not statistically significant.

[5] The reason why p-values up to .3 are included is that the hypotheses are in fact not independent. That is, they originate from a single underlying model. Consequently, the significance of p-values should be treated in the context of other p-values, comparable to what happens in a meta-analysis in which p-values are combined.

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Second, we used a specific play situation in our research setting, namely one in which the children were explicitly asked to play with the other child, and where adults monitored them. Note that this situation is different from a free play situation in the daily life of children. In short, the conflict between our findings and the literature is to a certain extent only apparent. The fact that we found rejected children to behave in this way in this particular situation does not exclude another fact, namely that their overall daily experience might be one of predominantly negative emotions and little interaction. A challenge for our process model is to expand our predictions to other interaction situations, and try to explain why there are infrequent and often negatively loaded interactions for rejected children.

2.2. Procedure

2.2.1. Generation of model output

The output of the model is obtained by running the model 5000 times for each set of control parameters that represents a specific type of dyad. All averages and distributions of the operational variables resulting from the model runs of different types of dyads are significantly different.

2.2.2. Random permutation analysis

To solve the problem of the major difference in size between empirical and model samples (13, 14, and 14 empirical dyads versus 5000 runs for each modeled dyad), we use random permutation analysis. This technique is highly flexible, is convenient for small and unbalanced datasets and can test explicitly formulated null hypotheses (Good, 1999; Manley, 1997; Todman & Dugard, 2001). It estimates the probability that an observed result is caused by chance alone by drawing a very large number (5000 in this case) of accidental samples (e.g. accidental mixtures of scores of popular and rejected children), and then counting the number of times that the observed phenomenon (or an even 'stronger' one) occurs in the accidental samples.

A broader theoretical justification of the procedures used in the current article is as follows. In the present kind of research, basically two pictures of reality are generated. One is a picture based on observations (the empirical study), the other is based on the simulation model. The question is: to what extent do the two pictures resemble each other? A good resemblance does not imply that the pictures look alike to the level of the smallest details, but there should be sufficient similarity so that an unbiased viewer can recognize one picture as a representation of the other. The hypothesis is that this similarity is meaningful. The nul hypothesis is that the model is a 'chance-machine' and that this similarity is accidental. The simplest form of chance consists of breaking the association that the model has made between outcome values of operational values and various types of dyads by randomly shuffling the 'output' of the model. This procedure will be followed with the pattern analysis of averages and distributions (see section 'Fitting averages' and 'Fitting distributions').

A second form of chance, which is related to a stricter form of null hypothesis testing, is represented by randomizing the input (control parameter values) of the model. The null hypothesis thus implies that the way the model describes statuses through its control parameter values has the same veracity as an arbitrary chance combination. This type of testing will be applied with the distribution analysis over distinct variables, mainly because in this particular case the simpler form of null hypothesis testing is not applicable (see Uebersax, 2005; van Geert & van Dijk, 2003).

A consequence of using this form of hypothesis testing through randomization tests is that our goodness-of-fit statistics (the measure that says how similar the data and the model are) will mostly consist of simple distance measures that are appropriate for the type of comparison made and for which the randomization estimates the *p*-value.

2.2.3. Operationalization of the research questions – an overview of the fitting methods

We answer the question by means of which tools and steps the empirical data can be fitted to the model by comparing the operational variables in four ways. First, we compare the average of each operational variable from the model runs with the empirical averages of 13, 14, and 14 dyads, respectively. Second, we compare the distributions of modeled variables with the empirical distributions. These first two tests address static properties of the empirical and model samples. The third focuses on dynamic properties, namely the emergence and temporal distribution of attractor states and the distribution of emotional intensities. Fourth, a sensitivity analysis is conducted, which is a check on how sensitive the model is to changes in parameter values.

In short, in addition to using averages we also fit distributions, which provide information about the differences between dyads from a particular dyad group. We also examine whether the fit of the static properties is supported by a fit on dynamic properties. Finally, by performing sensitivity analysis, we attempt to find ranges of adequate values of input parameter groups, which are likely to provide a more valid representation of particular dyad groups than single central values.

With the help of these four fitting methods, we answer the second question, namely how well the model fits the data. In answering the first aspect of this second question, namely to what extent do the chosen parameter values fit the data, we use our comparison of averages and distributions of each variable and the two dynamic properties tested. The second aspect is whether the

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

chosen parameter settings are the best possible choice of parameter settings. This aspect can be decided by means of sensitivity analysis.

Given the complexity of the method, the following section combines the presentation of the method and results.

2.3. Fitting averages

The question whether the model gives a good representation of the group averages is divided in two specific questions. The first relates to a comparison between model and data for the pattern of all variables together over the three types of dyads. The second involves a comparison between the simulated and the empirical averages of each variable separately for the distinct types of dyads.

2.3.1. The pattern of all variables over all status groups

The specific question is: does the model give a good representation of the pattern of empirical averages? The null hypothesis is that the pattern generated by the model does not give a better representation of the pattern of empirical averages than can be expected from random combinations of values of operational variables with particular status groups.

First, in order to compare values of distinct variables, we rescaled all empirical and model values by dividing them by their maximum value. Second, the distance, i.e. the absolute difference, between the rescaled simulated and the corresponding empirical value are calculated for each separate variable, for each separate status group. Third, these distances are summed over all variables and over all status groups. The resulting sum is compared with a chance model by means of a random permutation method.[6] Finally, we calculated the probability that a random model yields a sum of distances that is as small as or smaller than the actual simulation model does.

The simulation model yields a model–data distance that is significantly smaller than can be expected on the basis of chance ($p = .01$).

2.3.2. Per variable and per status group

The specific question is: does the model give a good representation of the empirical averages, for each variable separately? This question can be operationalized in two ways. First: does the average of a variable generated by the model fall within the range of the empirical confidence interval? Second: is the order of ranking of the simulated averages of the three groups the same as the empirical one? Checking the order of ranking makes sense only for variables that show empirical differences between 'popular' dyads and 'rejected' dyads (see section 'Empirical data').

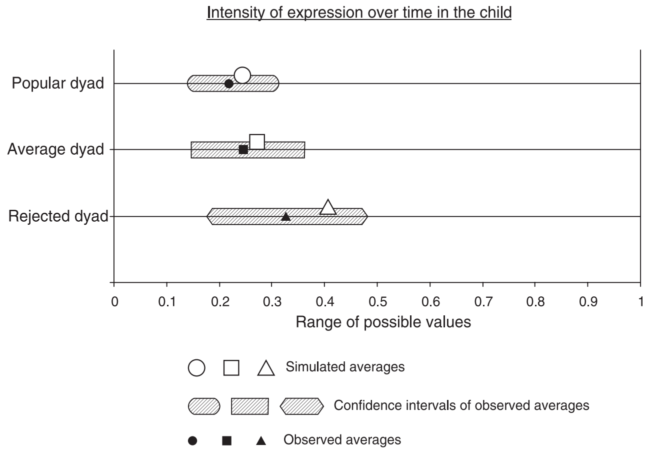

Figure 3 *An example of the position of the simulated averages compared to those of the empirical confidence range, for the variable 'intensity positive expressions child'.*

First, the 95% confidence interval of the empirical values is determined via a bootstrap procedure (Efron, 1988). Second, the position of the averages produced by the model is compared to that of the empirical confidence range.

Figure 3 gives an example for the variable 'intensity positive expressions child' and shows that the averages of the three types of dyad lie within the confidence intervals of the empirical averages. The order of ranking of the simulated averages for the three types of dyad is the same as the empirical ranking.

Table 4 shows the fit quality of all variables for the three types of dyad.

First, we will discuss the goodness-of-fit as expressed in the confidence interval. The goodness-of-fit refers to the distance between the variable generated by the model and the confidence interval of the empirical variable. It is defined by a position measure calculated as follows: $1 - (\mathrm{ABS}(V - L) + \mathrm{ABS}(V - U) - \mathrm{ABS}(U - L))$, when V is the value produced by the model, L the lower boundary, and U the upper boundary of the empirical confidence interval. A goodness-of-fit of 1 means that the simulated average falls within the empirical confidence interval; smaller values reflect increasing distances.

For example, the goodness-of-fit for the variable child's 'directedness' is 0.92, 1, and 0.84, for the three dyad groups, with a p-value of .04. Overall, the goodness-of-fit per variable is excellent for 10 variables ($p < .001$), good for six variables ($p < .05$), moderate for four variables ($0.05 < p < .1$), and poor for two variables ($p > .1$).

The goodness-of-fit per type of dyad is specified in the last two rows of the table. The three p-values are smaller than .01, which implies that they are significantly better than chance. The rejected and average dyads fit equally well (goodness-of-fit of 0.95 and 0.97). The popular dyad fits less well (0.85), and this difference in goodness-of-fit is statistically significant ($p < .01$).

[6] We also calculated an alternative calculation of distance, namely chi-square, and a stricter rescaling version, in which x – min x/(max x – min x). These alternative versions also produce significant results; in the chi-square ($p = .05$), and in the stricter version of rescaling ($p < .01$).

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Table 4 *Results of fitting procedure over averages: per variable, per status group*

	Confidence interval					Order of ranking			
	Rejected	Average	Popular	Total	p-values	Rejected	Average	Popular	Total
Child variables									
Directedness	0.92	1	0.84	0.92	0.04	1	1	1	1
Proportion shared directedness	0.84	0.96	0.75	0.85	0.07	0	0	0	0
Positive expressions	1	1	1	1	0	1	1	1	1
Intensity positive time	1	1	1	1	0.01	1	1	1	1
Intensity positive number	1	0.65	0.61	0.75	0.04	0	0	0	0
Proportion shared positive	0.87	1	0.73	0.87	0.1	0	0	0	0
Negative expressions	1	1	1	1	0	1	0	0	0.33
Intensity negative time	1	1	1	1	0	1	0	0	0.33
Intensity negative number	1	1	0.59	0.86	0	0	0	0	0
Play partner variables									
Directedness	0.7	1	0.86	0.85	0.08	0	0	0	0
Proportion shared directedness	0.74	0.96	0.85	0.85	0.07	0	0	1	0.33
Positive expressions	0.94	1	1	0.98	0	0	0	0	0
Intensity positive time	0.95	1	1	0.98	0.02	0	0	0	0
Intensity positive number	1	0.68	−0.07	0.54	0.21	0	0	0	0
Proportion shared positive	1	1	0.92	0.97	0.03	0	0	0	0
Negative expressions	1	1	0.97	0.99	0	0	1	0	0.33
Intensity negative time	1	1	0.92	0.97	0	0	0	0	0
Intensity negative number	1	1	0.64	0.88	0	1	1	1	1
Dyad variables									
Coherence dyad	1	1	1	1	0.01	0	0	1	0.33
Shared positive expressions	1	0.99	1	1	0	1	1	1	1
Shared negative expressions	1	1	0.99	1	0	0	0	0	0
Contrast dyad	1	1	1	1	0.3	0	0	0	0
Average fit quality over status groups	0.95	0.97	0.85	0.92	0	0.58	0.5	0.58	0.56
p-values	0	0	0	0	0	0.06	0.17	0.06	0.01

Notes: *a.* The fit quality can be interpreted as follows: For the confidence interval, a fit quality of 1 means that the simulated average falls within the empirical confidence interval; smaller values reflect increasing distances. For the order of ranking, a fit quality of 1 means that the rank order of the simulated average is the same as the rank order of the empirical average; 0 means that these rank orders are not the same.
b. The seven variables that yielded significant differences between status groups in the empirical study are printed in *italics*.

Second, we will discuss the goodness-of-fit as expressed in the order of ranking of the values of the three types of dyad. Note that we report ranking only for the 12 variables that are empirically different across dyad types. Seven of these 12 variables had a p-value smaller than .05. Note further that since calculating p-values for separate variables makes little sense (the number of possible permutations is too small), we report only the p-values for the three types of dyad. For the 'rejected' dyads and the 'popular' dyads, the fit is moderate ($p = .06$). For the average dyads, the fit is poor ($p = .17$). Over all dyads, the fit is excellent ($p < .01$).

2.4. Fitting distributions

The question is: do the simulated model distributions give a good representation of the observed (empirical) distributions of all variables? In accordance with the discussion of the averages, we make a distinction between the fit over the total pattern of distributions and the fit over the variables separately.

2.4.1. The pattern of all variables over all status groups

Does the model give a good representation of the total pattern of distributions of all variables over the three status groups? This question is answered with the help of two criteria, namely the position of the coordinates of the peak in the distribution and the histogram of the distribution itself. Just as in the analyses of the averages, we assume that the model yields a specific pattern of association between histograms and types of dyad. The null hypothesis is that there exists no systematic association between a particular histogram and a particular variable.

The first criterion consists of the coordinates of the peak. The hypothesis is operationalized as follows: do the coordinates of the peak of the simulated histograms resemble the empirical coordinates more than can be expected on the basis of chance? Coordinates of the peak are both the position on the y-axis (the maximal frequency) and the position on the x-axis (the value that occurs most frequently).

First, histograms are determined for the empirical data, the model data, and the null hypothesis data which are random couplings of distributions to variables. Second, for each variable (in the model data, the empirical data, and the random model data), the coordinates of the peak are determined. Third, for each variable we calculate the distance between the coordinates of the peak in the model distribution and the coordinates of the peak in the empirical distribution and the distance between the coordinates of the peak in the random model and those in the empirical distribution. Finally,

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

we average the distance over all variables and coordinates and compare it with the average obtained on the basis of the null hypothesis or random model.

The result of the random permutation test is that the distance between the average of the coordinates of the peak in the model distribution and the average of the coordinates of the peak in the empirical distribution is significantly smaller ($p < .001$) than can be expected on the basis of chance.

The second criterion consists of the histograms. The question is whether the simulated histograms of the distribution resemble the histograms of the empirical distribution more than can be expected on the basis of chance. The null hypothesis is that a random association between histograms and variables produces an equally good fit with the histograms of the empirical distribution.

The statistical method can be compared with the method described for the coordinates of the peak, but we now focus on values for each distinct bin in the histogram.

The result of the test is that the distance between the histograms of the model distribution and those of the empirical distribution, for all variables for all status groups, is significantly smaller ($p < .001$) than the corresponding distance between the histograms of the random model distribution and those of the empirical distribution.

2.4.2. Testing a model with an arbitrary parameter range for separate variables and status groups

The specific question is: does the model give a good representation of the distributions of variables separately over distinct status groups? The hypothesis is that the model histograms resemble the observed histograms and that this resemblance is meaningful. By 'meaningful' we understand that the resemblance is the specific result of the parameter values chosen. According to the null hypothesis, the resemblance is accidental. This implies, in this particular case, that any arbitrarily chosen set of parameter values could have produced a comparable resemblance. Remember that the model uses parameter sets that are considered specific for the three types of dyad. The random model employs randomly chosen parameter values. They were taken from an interval twice the range of the parameter values featuring in the real (i.e. specific) model[7] in order to avoid unrealistic values. Note that in this case the model parameters for the three types of dyad are nested within this broader range.

First, a measure of fit is calculated, which is the sum of the distances (absolute difference) between the histogram produced by the model and the corresponding empirical histogram. The same is done for the random model. The difference between these two measures is an indicator for the goodness-of-fit of the model. The null hypothesis assumes that the values of the random parameter model and the values of the real model come from the same underlying distribution. This assumption is tested with the help of a random permutation procedure (1000 runs).

Table 5 shows the p-values resulting from the random permutation procedure. To simplify the interpretation of this table, the number of times that the p-value is smaller than .01 has been counted for each variable (represented in the rows) and for each type of dyad (represented in the columns).

Overall, in 10 of the 22 variables, the real model resembles the empirical distributions better than the random model for all status groups ($p < .01$); in the remaining 12 variables this is the case for two of the three status groups. We evaluate the goodness-of-fit as 'excellent' for the 'average' dyads (22 variables out of 22 are significantly better); for the 'rejected' group the fit is considered 'very good' (20 variables significant). Finally, we consider the fit for the popular group 'moderate', because of the combination of a low number of significant p-values (14) with a considerable number (7) of high p-values that refer to a better fit of the random model.

For a discussion of the fits resulting from a random model that is not limited to realistic values, we refer to the Endnote.

2.5. Fitting dynamic properties

The fitting of dynamic properties is confined to two indicators of the dynamics of a process, namely the emergence and temporal distribution of attractor states and the distribution of emotional intensities.

2.5.1. Attractor states: episodes of dominance of self- versus other-directed actions

The test of attractor states will focus on the time distribution of the order variable other- versus self-directedness. We define an attractor as a dominant behavioral mode of the system that has a sufficient macroscopic duration (see also the state space grid literature; Hollenstein & Lewis, 2006). Applied to the dyadic level there are three possible states: one in which both children are directed towards each other (real interaction), one in which both children are directed towards their own activity (real solitary play) and one in which one child is directed towards the other and the other is directed towards itself. Coordinated other-directedness and coordinated self-directedness are the most likely attractors because they are self-sustaining. A situation in which a child is other-directed and the peer is self-directed is

[7] For instance, in the first parameter group (strength of concerns) in the model, the rejected child in the 'rejected' dyad has the highest value, namely 0.8, for the concern Involvement (CI), the average child in the 'average' dyad has a value of 0.7 and the popular child in the 'popular' dyad has the lowest value, namely 0.6. The range between the highest and lowest value equals 0.8–0.6 = 0.2. The random model is based on an arbitrary choice of parameter values within twice that range, i.e. between 0.5 and 0.9. Comparable ranges for random parameter selection are of course determined for all other parameters.

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Table 5 *Results of fitting procedure over distributions; per variable, per status group, as expressed in* p*-values*

Type of dyads	Rejected dyad	Average dyad	Popular dyad	Marginal total ($p < .01$) per variable
Child variables				
Directedness	0	0	0.28	2
Proportion shared directedness	0	0	1	2
Positive expressions	0	0	0	3
Intensity positive time	0.04	0	0	2
Intensity positive number	0.67	0	0	2
Proportion shared positive	0	0	1	2
Negative expressions	0	0	1	2
Intensity negative time	0	0	0	3
Intensity negative number	0	0	1	2
Play partner variables				
Directedness	0	0	0	3
Proportion shared directedness	0	0	0	3
Positive expressions	0	0	0.03	2
Intensity positive time	0	0	1	2
Intensity positive number	0	0	1	2
Proportion shared positive	0.74	0	0	2
Negative expressions	0	0	0	3
Intensity negative time	0	0	0	3
Intensity negative number	0	0	0	3
Dyad variables				
Coherence dyad	0	0	0	3
Shared positive expressions	0	0	1	2
Shared negative expressions	0	0	0	3
Contrast dyad	0	0	0	3
Marginal total ($p < .01$) per status group	20	22	14	

Note: *a*. The fit quality is expressed in *p*-values: e.g. a *p*-value of 0 ($p \leq .001$) for the variable 'directedness child' for the 'rejected' group of dyads means that the real model resembles the empirical distributions better than the random model, for this variable, for this group of dyads.
b. The marginal sums per variable expresses the number of times that $p < .01$ for this variable; the maximum value is 3 and the minimum value is 0. The marginal sums per status group expresses the number of times that $p < .01$ for this status group; the maximum value is 22 and the minimum value is 0.

likely to be a transitory state, resulting either in the peer picking up the invitation for interaction from the child, or the child giving up the attempt to establish a real interaction. The question is: How many such coordinated states are present during a real play session and does the model adequately predict that frequency?

The statistical method is as follows. The empirical data and the model output specify sequences of 1's representing actions directed towards the other person and 0's representing self-directed actions for the child and peer separately. An attractor state is operationally defined as an interval in which one behavioral mode (other-directedness or self-directedness) dominates, e.g. occurs 80% or more of the time. What matters is not the exact percentage of the behaviors, but the difference among such patterns over time, which should be big enough to correspond with switches in dominant modes.

In order to clearly reveal eventual patterns of such episodes, the 1–0 patterns need to be statistically smoothed. The smoothing was carried out by means of a Savitzky-Golay smoothing procedure, with a smoothing window of one-tenth of the number of time points, based on a polynomial of order six and three consecutive smoothing passes (the software used is TableCurve 2D, a smoothing and curve estimation program). It goes without saying that all trajectories – observed, or based on any of the models – were smoothed with identical smoothing parameters.

Attractors in the sense of coordinated patterns of either real interaction or solitary play require that the behavioral modes of child and peer are about similar. In order to calculate such similarity as an indicator of attractor states, we first perform a standardization on the smoothed data series of child and peer separately. The standardization is of the form $2(\times-0.5)$ and makes the data vary between -1 (self-directedness) and 1 (other-directedness). The similarity at any point in time can be calculated by multiplying the standardized values, which is mathematically similar to calculating a covariance over time. If the covariance approaches 1, similarity is near maximum, and the episode is clearly an attractor of either interaction or solitary activity in both children. By rescaling the multiplied data back to the range between 0 and 1, a similarity trajectory results in which the peaks express the occurrence of attractor states over time. Peaks are defined as points with a difference of at least two standard deviations (0.4) from a valley, i.e. a point of minimal similarity. By means of a Visual Basic macro running under Microsoft Excel, the number, position and height of the major peaks and valleys, characterized by the two-standard deviation-difference criterion were calculated for the smoothed trajectories based on the data and the dynamic model simulation, respectively.

Results are given in Figure 4 and Table 6. On average, the number of peaks generated by the dynamic model is

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Table 6 *The number of peaks in coherent, i.e. similar, behavior among participants*

	Data		
	Average number of peaks	Modal number of peaks	Average value of difference between peaks and values
Popular	2.86	3	0.64
Average	2.86	3	0.63
Rejected	2.17	1	0.61
Total	2.65	3	0.63
	Model		
	Average number of peaks	Modal number of peaks	Average value of difference between peaks and values
Popular	2.23	2	0.61
Average	2.01 (2.68)	0 (1)	0.63
Rejected	1.99	2	0.6
Total	2.1 (2.3)	2 (2)	0.61

Note: If number is 0, the play session shows a single pattern of coherence (either only coordinated other- or coordinated self-directed behavior; numbers between brackets are based on calculations where 0-numbers have been left out).

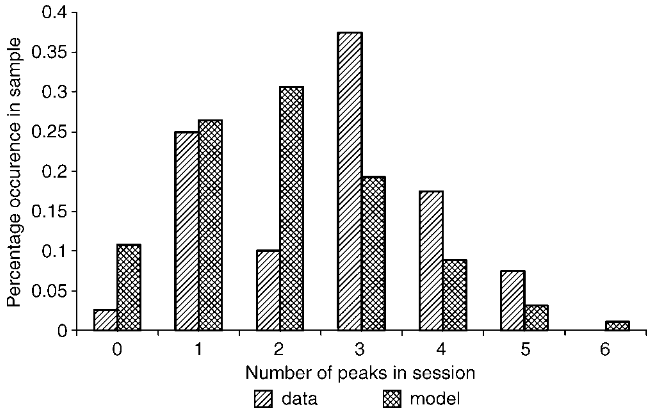

Figure 4 *Frequency distribution in the number of action attractors per session, in the data and the model.*

less (about 0.5 to 1) than the number of peaks in the data. The height of the peaks, which is a simple indicator of the strength of the attractors, is on average 0.61 and practically identical for the model and the data. In summary, although there are differences in the average number of peaks, the qualitative similarity between data and model is convincing.

The average number of peaks in the empirical data of rejected dyads is lower than in the popular and average dyads (p = .08, random permutation test). The model correctly predicts that the rejected dyad has less peaks (p = .04), but it differs only with the popular dyads, and the difference is smaller.

One might object that the similarity in number of peaks between the model and the data is accidental, in that a model based on a random dispersion of 1's and 0's according to the averages obtained through the data might just as well give a comparable number of peaks, or comparably small differences with the data. In order to test this possibility, we generated 200 sets of randomly dispersed 1's and 0's over 420 steps, based on the observed averages, which were then statistically processed in the same way as the data and model. The results are clearly different from the dynamic model and do not account for the data. The random model predicts an average 0.1 peak, versus 2.1 and 2.65 for the dynamic model and data.

2.5.2. Distributions governed by a power law

Emotional expressions are likely to be a self-organizing critical phenomenon. They occur with a certain intensity and frequency and consist of sudden shifts from one expression to another, with a majority of expressions of low intensity (neutral expressions; see Izard, Ackerman, Schoff & Fine, 2000; Lewis, 2000b; Camras, 2000; Mascolo, Harkins & Harakal, 2000). Self-organized, critical phenomena show scale invariance. That is, the distribution of these phenomena at various scales of intensity is invariant. Scale invariance is typically represented by a power law distribution (examples are discussed in Bak, 1996, and Schroeder, 1991). More precisely, the number of times a phenomenon of magnitude m occurs (expressed as $N(m)$) is equal to a certain power p of m, and is expressed by the equation $N(m) = a*m^{-p}$.

Does scale invariance, i.e. a power law distribution, also occur in the case of emotional expressions? Are the model and data similar or comparable in this respect?

The statistical method is as follows. Recall that emotional expressions were coded in terms of intensity ranging from −4 to +5. We will assume that these coded intensities are our best possible approximation of 'objective' quantifications of emotional intensities. We thus expect that the frequencies will behave in accordance with the power law distribution.

A power law distribution amounts to a straight, descending line if the intensities and frequencies are represented by their log-values. Hence, a simple way to check the distribution is by visual inspection of the linearity of the log-log graphs of the intensities and frequencies. The observed frequencies for the three dyads

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Table 7 *Exponents and adjusted R^2 values of the power law distribution of emotional expressions of various intensities*

	Popular	Average	Rejected
Data			
Exponent	2.66	3.22	2.51
R^2 adj	0.95	0.77	0.94
Model			
Exponent	1.91 (2.3)	1.64 (1.79)	1.85 (2.82)
R^2 adj	0.85 (0.98)	0.89 (0.9)	0.82 (0.98)

and the child and peer separately show a good fit with the expected power law distribution. There is, however, one typical aberration in all the data sets, namely a striking underrepresentation of the emotional expressions of intensity 2 ('a little bit positive or negative'). It seems as if the raters had difficulty discerning this category from the neutral expression category. If the intensity-2 category is omitted from the data, the log-log representation indeed corresponds quite well with a linear descending slope characteristic of a power law distribution.

Visual inspection of the log-log graphs of the frequencies of emotional categories produced by the model showed a clear linear relationship, except for the emotions of highest intensity, which were clearly overrepresented. This is due to a certain tendency of the model to run into extreme situations in terms of the level of realization of concerns, which is a weakness of the current dynamic model.

Results are represented in Table 7 which shows the exponents and adjusted R^2 values for the emotional intensities counted over the dyads (the frequencies for the child and peer separately can be found under the web materials at the first author's website). The smaller the value of the exponent, the higher the number of higher-intensity expressions. The main conclusion is that emotional intensities occurring over the course of a play session follow the predicted power law distribution quite well (R^2 values are high). Differences between the types of dyads in terms of distribution exponents are not particularly clear and are left for further study.

Given the small number of intensity-2 expressions and the possibility that the raters worked with an implicit intensity scale with less than the six categories they had to score, we tried several aggregations of the observed emotional intensities and checked whether they provided better fits with the power law model. The best fitting model with at least four emotional expression categories is the model based on the aggregation of intensities 1 and 2, 3 and 4, 5 and 6 (adjusted $R^2 = 0.97$).

2.6. Sensitivity analysis

To what extent does the model react to changes in values of control parameter groups? This question concerns the sensitivity of the model. Figure 5 shows examples of three degrees of fit sensitivity, based on imaginary cases, and one empirical example. An optimal fit occurs if a relatively small fraction of the control parameter interval corresponds with the empirical confidence interval of a particular order parameter (Figure 5a).

2.6.1. Exploratory qualitative analysis of sensitivity

The question is whether the model is either too sensitive to small or too insensitive to greater changes in control parameter values.

A model can be too sensitive, in which case a small change in a control parameter value causes a major change in the goodness-of-fit of the model, with no apparent reason for why this should be so (see Figure 5b). For instance, it is highly unlikely that a very small increase in the Involvement concern parameter value would lead to a major difference in the trajectory of one or more of the order variables, thus leading to a sudden deterioration of the goodness-of-fit, especially if a bigger change in the increase again suddenly leads to an optimal fit. Particularly in models with stochastic control parameters, as in the present dyadic interaction model, the expectation is that no over-sensitivity will occur. It is expected that the model will show s-shaped or otherwise monotonic responses of its order parameters to changes in the control parameters.

We have checked whether the model did not show too many fluctuations in its order parameters, if control parameters are varied. Since it is virtually impossible to calculate all possible combinations of control parameter values, we confine ourselves to manipulating one control parameter at a time. Charts were generated with on the x-axis a control parameter that varied from a minimum value to a maximum value; and on the y-axis a particular variable of an order parameter. Charts were visually inspected with a focus on discontinuities and fluctuations that qualitatively resemble Figure 5b. The control parameters chosen for inspection were 'Involvement of the child', 'Involvement of the play partner' (both belong to the Concerns parameter group), and the 'Contribution of the behavior to the realization of the concerns' (Realizability parameter group). The variables describing properties of the major order parameters were the dyad's 'coherence', the child's 'directedness', 'positive expressions', 'intensity positive time' and 'negative expressions', and the play partner's 'directedness'. There occurred no discontinuities, instabilities, and fluctuations in the figures that could point to oversensitivity of the model. Figure 5d presents an illustration of this finding, namely in the variable 'coherence dyad' with the parameter 'involvement of the child', for the rejected group of dyads. It shows that an s-shaped curve appears, with a small dip on the right side.

A model like the current one should not be too sensitive, but it should not be too insensitive either. We must therefore check if there are no control parameter values for which a major change in a parameter value causes no noticeable changes in a particular control parameter,

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

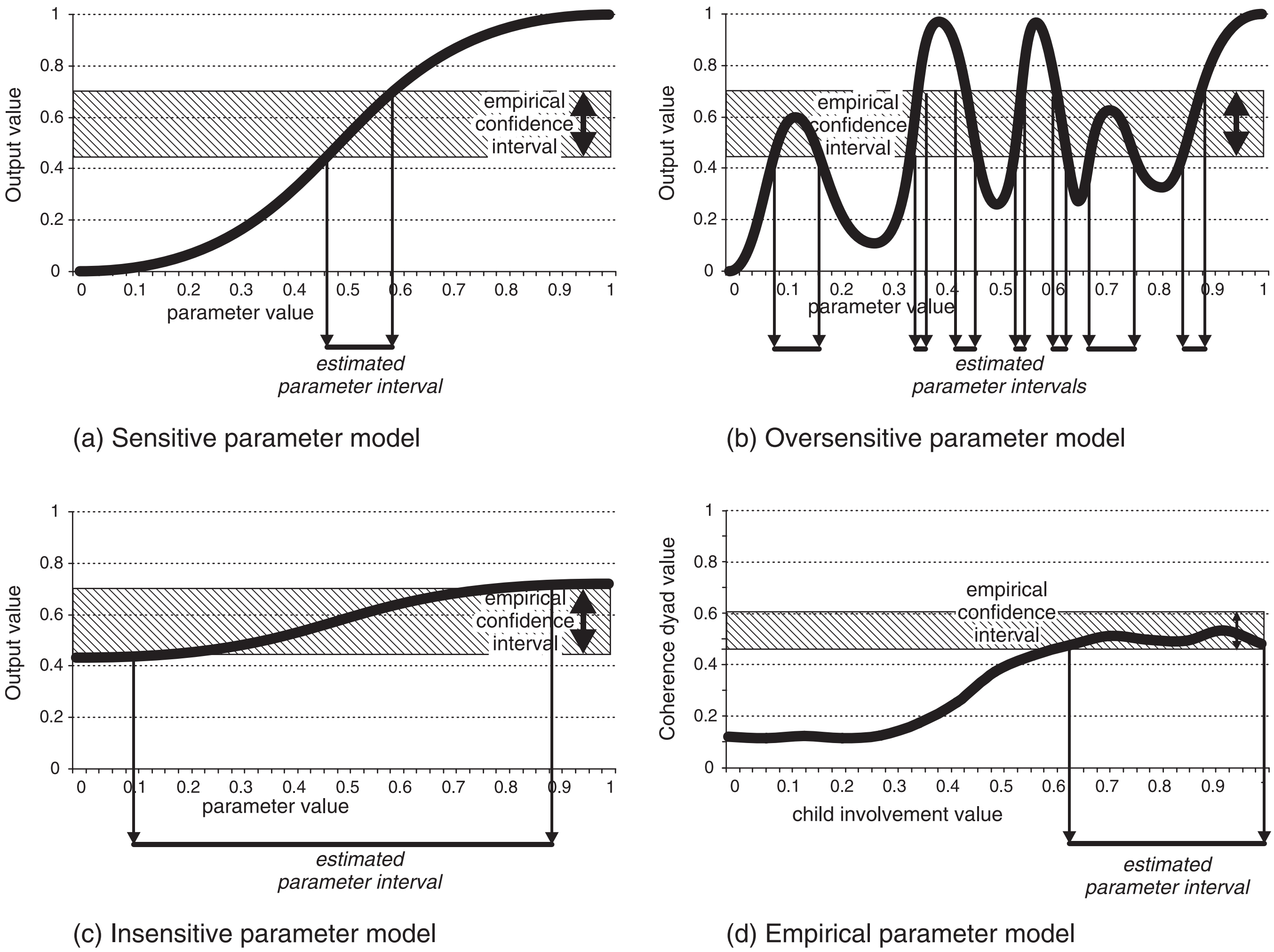

Figure 5 *Examples of three degrees of fit sensitivity, based on imaginary cases and one empirical example: a. Sensitivity parameter model; b. Oversensitive parameter model; c. Insensitive parameter model; d. Empirical parameter model.*

and thus no noticeable changes in the goodness-of-fit (see Figure 5c for an imaginary insensitive model). Insensitivity of the model can be checked in the same way as oversensitivity. The only difference is that we now check whether an order parameter variable remains within the limits of the empirical confidence interval when the underlying parameter is varied to an arbitrarily large extent. We visually inspected the same sample of charts as described for the oversensitivity check and found that in none of the variables the range of control parameter values that yields a good fit is too broad; i.e. the model is not insensitive.

2.6.2. Using sensitivity analysis to determine the quality of fit

Sensitivity analysis can also be used to improve the goodness-of-fit of the model. To that end, we first checked whether the simulated average of a certain order parameter variable falls within the limits of the empirical confidence interval. We did this for discrete subranges (bins) of the control parameter in question. If a control parameter bin corresponds with an order parameter value that falls within the empirical confidence interval, it is granted one point. We did this for six important variables in total. All points for each bin are summed, and the total number represents the goodness-of-fit of the model for this parameter, in this particular bin (see Figure 6).

Figure 6 illustrates this principle with the control parameter 'child_involvement', and six control parameter variables (three variables referring to the order parameter of self- versus other-directed action, three variables about emotional expression), of the 'popular' group of dyads.

We can see that if the parameter 'child_involvement' has a value between 0 and 0.43, only 1 order parameter variable ('directedness play partner') falls within the limits of the empirical confidence interval. Between 0.43 and 0.5, none of the variables gives a good fit. However, between 0.5 and 0.57, five variables fall within the empirical confidence interval, which means that the fit is good, but still not optimal (fit quality 5). After this, the fit quality decreases, to 3, 2, and finally 1. Notice that the

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

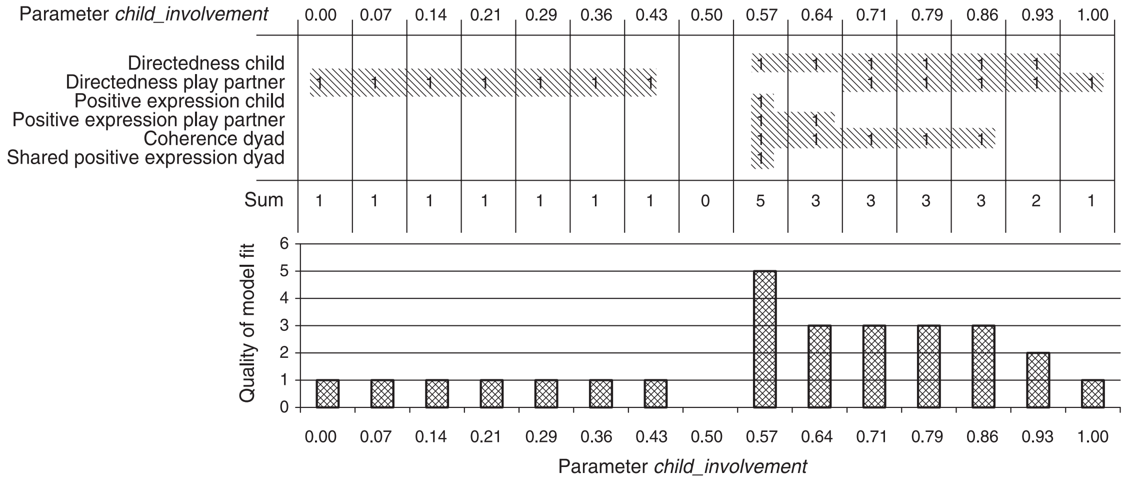

Figure 6 *Sensitivity analyses of the parameter 'child_involvement' and six variables (three action-based variables and three expression-based variables) in the 'popular' group of dyads.*

good fit we found for control parameter values between 0.5 and 0.57 corresponds with the parameter value in the model for the popular child in the 'popular' dyad, which was set to 0.57 and which was based on a fairly intuitive process of estimation of values. This means that the sensitivity analysis eventually provides a good underpinning of the estimation of control parameter values.

3. Discussion

3.1. Conclusion

The goal of the present article was to report on the empirical validation of a recursive application of an action-based model of short-term dynamics of interaction processes. The empirical validation concerned data from dyads consisting of children of different sociometric statuses. The tools and steps needed for empirically validating this dynamic model included static properties (averages and distributions), dynamic properties (attractor states and distribution of emotional intensities), and sensitivity analyses. The results show that overall the model showed a good empirical fit, although the goodness-of-fit for the popular dyads was only moderate. A potential explanation for this lack of fit derives from additional research (see Steenbeek & van Geert, 2007) suggesting that a child's social competence plays a major role in realizing concerns. A high level of social competence is particularly characteristic of popular children. Social competence appeared to be insufficiently implemented in the parameters in the current model, and this might explain the lesser fit for the 'popular' dyads (see the more extensive description on the website www.gmw.rug.nl/~model).

As the preliminary issues regarding the goodness of fit now are satisfactorily dealt with, we can focus on the question what the model teaches us about differences in social interaction of 6- and 7-year-old children of different sociometric status. In order to answer this question, first we will present implications of our findings for explaining social interaction of children of different sociometric status on both the short-term time scale of interaction and on the long-term time scale of the development of interaction and status. Second, we will discuss a number of methodological implications that follow from our dynamic model-oriented approach. Finally, we will summarize a number of general implications that we see as the 'take-home message' of our approach to social interaction and social status.

3.2. Implications for explaining social interaction of children of different sociometric status: the short-term time scale

We consider the overall good fit of the model as support for the model's theoretical assumptions about interaction in the short term. These assumptions concerned the goal-directedness of interactions, in which context-specific concerns are the guiding force behind a child's actions and emotional expressions. Additional assumptions concerned the extent to which events contribute to the realization of the child's concerns, the role of emotional appraisals and pleasure in this and non-intentional aspects of behavior, such as contagiousness in an interaction process. Finally, sociometric status expresses differences in social power and social competence, which are viewed as long-term order parameters that differentially affect the properties of social interaction between children.

3.2.1. Using context as a dynamic factor

One of the points of criticism we had of existing research is that it tends to define context as an independent variable. In this study, we see context as a time-dependent and variable aspect of the dynamics of the interaction

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

process itself. Hence, context is a dynamic and constructed, not a static and given characteristic of an interaction. To the extent that sociometric status of a child or play partner is an element of the context of the play situation, it is treated as a series of concrete and immediate determinants of the way a concrete interaction with another child unfolds, i.e. as a proximal property. With this approach to context, we can obtain a better understanding of the dynamics of real-time interaction, which is a crucial starting point for understanding change in interactions over developmental time (Mascolo, 2005).

The first implication of this approach to context regards the necessity to distinguish various situations and contexts in research of interaction studies to a greater extent than is already done (Pettitt, Brown & Mize, 1998; Hubbard, Dodge, Cillessen, Coie & Schwartz, 2001; Pickard, 2004). In addition, it is crucial to use context–person interaction as the starting point for predictions (Thelen & Smith, 1994; Gottman *et al.*, 1997; Fogel, 1993). Finally, our findings strongly support the importance of using context-specific goal-directedness of behavior in peer interaction research, emphasizing the role of concerns as directing forces in behavior (Packer & Scott, 1992; Renshaw & Asher, 1983; Crick & Dodge, 1994; Mize & Ladd, 1990; Erdley & Asher, 1996; Rabiner & Gordon, 1992; Brown, Odom & Holcombe, 1996; Want & Gattis, 2005).

The second implication of a dynamic view of context is that it helps to better understand empirical findings in specific contexts, and thus potentially contributes to better prediction. In the current context of an adult-initiated play situation, specific differences and similarities have been found between popular, average and rejected children in dyads. One such finding is that in this situation rejected children show more positive expressions than popular children. Our model helps to understand this finding by referring to the major role of concerns. It describes the specific process in which these children use their emotional expressions in trying to establish satisfactory interactions with others and the specific way in which other children react to these expressions. This leads to the verified prediction of rejected children showing an 'overflow' of positive expressions. A comparable finding is that the amount of positive expressions that are reciprocated by a positive expression of the play partner (as expressed in the variable 'proportion shared positive') is higher for popular children than for rejected children. The dynamic model predicts this finding by showing how the child's social effectiveness and social contagiousness act as mutually amplifying forces in the case of high-power individuals. Another example of a predicted empirical finding is the occurrence of less alternations among attractor states in the rejected dyad interaction pattern. It appears that rejected dyads tend to stay somewhat longer in the attractor state of either playing together or either playing alone than average dyads or dyads containing a popular child. This lesser switching from one state to another might be related to lesser flexibility in the interaction with a rejected child. This suspicion is based on the assumption that average and popular play partners succeed in higher reciprocity in the interaction and that their somewhat higher switching frequency is a sign of that reciprocity (see also Granic, Hollenstein & Dishion, 2003, on the occurrence of lesser interaction flexibility in deviant youths).

These predictions could probably not have been made on the basis of existing models that focus primarily on static relationships, which are considerably less context-sensitive and not process-based. Certainly not all predictions of the model have been verified by the data and thus further refinements are needed, but the important point is that a recursive process model like ours allows us to make such predictions on theoretical grounds.

3.2.2. Applied implications

A better theoretical understanding of empirical phenomena leads to better intervention and assessment. Based on the model, we can make the following suggestions. First, in addition to focusing on child-specific factors, assessment and intervention should also explicitly account for context-specific factors in a process-oriented framework. Desired changes in behavior can be realized by varying the context, which means that the teacher can initiate the child into contexts that he himself is unlikely to initiate spontaneously. For instance, the teacher can let a rejected child cooperate with a more popular child in the class. This must be done in a particular action context, in which the rejected child can practice new skills, and experience positive emotions in doing so. In addition, other children can be taught to react positively to rejected or less socially preferred children (Tanta, Deitz, White & Billingsley, 2005; Schuele, Rice & Wilcox, 1995; Haring & Lovinger, 1989; Hendrickson, Strain, Tremblay & Shores, 1982). While initiating such contexts, the teacher must take the child's and interaction partner's goals and concerns into account and reckon with the children's perceptions of how and to what extent specific actions and events contribute to the realization of their concerns. The question is of course how the teacher gets to know those concerns, intentions and perceptions. One way to achieve this is to have a preliminary assessment of the concerns and appraisals of a child. For instance, before involving a child in social skills training, the child can be interviewed about his concerns or goals that pertain to his actions and the actions of others, including socially undesired behaviors, such as aggression. The child's own interpretations can be highly informative, if assessed in a proper way (Murphy & Eisenberg, 2002; Singer, Doornenbal & Okma, 2004; Visser, Singer, van Geert & Kunnen, 2006). Another way is to observe children's actions and emotions, with the explicit goal of finding out about their concerns, which can be done, in particular, in problematic or conflict situations (Troop-Gordon & Asher, 2005;

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Murphy & Eisenberg, 2002; Delveaux & Daniels, 2000; Rose & Asher, 1999).

Finally, the general principles underlying the dynamic systems model of interaction can be applied to the process of intervention itself. That is, interventions are part of an ongoing iterative process, in which each step in the process helps to determine the next intervention step, governed by concerns, appraisals, emotional expressions and so forth. For the teacher, for instance, it can be of great importance to think about what the teacher's own concerns are in the context of an intervention and how he or she perceives the emotional and drive value of differences between the preferred and the realized level of concerns (which now pertain to goals such as reducing the adverse effects on a child of rejected status). The action theory behind our model (Steenbeek & van Geert, 2007) is general, intuitive and fairly straightforward and can act as a broad frame-of-reference for practitioners in the class (van Geert & Steenbeek, 2006). It can function as an overarching framework for more specific principles of intervention that teachers might apply. Let us take behavior modification, which is widely applied in classes to solve various behavioral problems. For instance, while applying reinforcement, it is necessary to know what the child finds reinforcing, which requires insight into the child's concerns and intentions. Reinforcement is also often based on rewarding children for doing something they do not prefer, such as showing more on-task behavior, by letting this non-preferred behavior be followed by things they do prefer, such as playing a computer game. This principle, which is widely applied and which capitalizes on the child's repertoire of concerns, is known in behavior modification as the Premack principle (Brown, Spencer & Swift, 2002; Klatt & Morris, 2001).

3.3. Implications for explaining social interaction of children of different sociometric status: the long-term time scale

3.3.1 Long-term order parameters and selection of interaction partners

In this section, a first exploration of linking the models of short-term social interaction with a model of social development is described (see also van Geert & Steenbeek, 2005; Steenbeek, 2006). A play session governed by a particular short-term dynamics can be conceived of as a single step in a series of such interactions over a long term, for instance a school year, or a developmental period such as childhood. Development concerns the long-term change of order parameters such as power and competence, which determine the control parameter values that govern the short-term interaction process.

In order to build a model of long-term change, a number of additional parameters or variables must be specified, which are left implicit in the short-term model. An important example of such a variable concerns selection, i.e. the child's variably successful attempts to initiate an interaction (Matsui, Muto & Kadoyama, 2001; Pepler & Craig, 1998; Hauck, Fein, Waterhouse & Feinstein, 1995). Selection thus requires the availability of a range of potential partners (e.g. the child's class mates, children from the neighborhood, etc.). The long-term model must allow for some sort of representation in the child of their network in terms of past experiences of how pleasurable the interaction with a particular member of the child's network actually is. The child will tend to select those interaction partners with whom a pleasurable interaction can be established. It is likely that such continuous selection and resulting interactions lead to a differentiation in the social network, in terms of friends (with whom the child attempts to engage frequently) and others, but also in terms of sociometric status (children who are selected by many peers are popular, and those selected by few peers, if any, are rejected). The success of interaction initiations is likely to be moderated by the social competence, i.e. effectiveness, of the children in question. It is this dynamics of selecting and being selected by interaction partners that leads to changes in the long-term order parameters of power, popularity and social competence and, consequently, to changes in the short-term control parameters associated with these order parameters.

In order to check this long-term scenario of status differentiation we developed a simple dynamic model of long-term selection in a group of children. The model specifies that Initiating an interaction with another child is based on preference for that other child, that preference is partly based on sociometric status, and that status is defined by the frequency of being selected by others. Each step in this model amounts to a short-term interaction event, such as playing or working together. The model shows that sociometric differentiation is a self-organizing process, i.e. the group naturally divides into a distribution of statuses, i.e. frequencies of choice by others.

3.3.2 The emergence of long-term rejected status, or: how does rejection transpire?

A question of theoretical but also of considerable applied importance relates to the emergence and consolidation of rejection. According to the assumptions from the short-term model, rejected children are characterized by the fact that other children have a low preference for them. That is, the concern of other children for playing with a rejected child is low, in comparison with the concern for playing with popular children, for instance. The question is: where does this low concern of others come from and why does it last? The simplest answer, given from a static point of view, is that rejection is primarily an expression of socially repulsive properties of the child, such as aggressiveness, egocentrism, lack of social competence and so forth. From a dynamic point of view,

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

however, this need not be the case. The long-term model dynamic that we briefly mentioned in the preceding section shows that rejection as well as popularity are to a considerable extent the result of self-organizing processes. The occurrence of this process depends on three conditions. The first is that the pleasure of playing together with another child is likely to depend on one's preference for that child. The second is that this preference is (among other things) based on the popularity of the play partner. The third is that the popularity of the play partner is a function of how often the potential play partner is chosen or approached by other children. If these three conditions are satisfied, a dynamic emerges that inevitably results in a differentiation among children. This differentiation is basically driven by accidental events and thus also occurs if the children are basically equal at the start in terms of likeability or social competence. Our long-term dynamic model shows that if the frequency of being chosen by other children drops below a certain threshold, rejection can become a relatively persistent phenomenon. The road towards rejection is probably to a considerable extent a matter of bad luck, not a matter of (antecedent) bad habits.

In a dynamic model, causes and effects can become inextricably intertwined. Thus, children who from the start show a little less pleasurable social behavior than others, or any other characteristic that has a slightly reducing effect on the concern of others for playing with them, are more likely to become trapped in the downward spiral. Meanwhile, the relative lack of pleasurable interactions with others may over time reduce the rejected child's own concern for playing with others, or it might force the child to impose itself ever more strongly onto other children, through demanding, awkward or even aggressive behavior, which has the effect of even further diminishing the willingness of other children to play with them. However, as our data and short-term dynamic model suggest, this need not be the only possible scenario. In many cases, rejected children appear to conserve the ability to establish positive interactions with others and they are likely to do their best to be good interaction partners, if given the opportunity. But since the lack of opportunities, i.e. the low interaction concerns of others, is the hallmark of rejection, it is likely that it will persist if nothing is done about it.

The assumption that rejection, as well as popularity, is to a considerable extent the result of a self-organizing dynamics, which can in principle be counteracted, is based on conclusions drawn from a simple long-term model of selection and preference, based on principles taken from the short-term model. Empirical research is needed to show whether or not rejection indeed emerges and persists in this way.

3.4. Methodological implications

A first implication concerns Murray's relativizing the importance of empirical fits. He states that 'fitting the data does not tell you that you have the right mechanism. . . . We have to free ourselves from the idea that goodness of fit is the sine qua non of science' (Gottman *et al.*, 2002, pp. 67–68). This statement supports the necessity of a good theoretical justification of a model. The empirical fit makes sense only against the background of the underlying theory. We have tried to conform to this requirement by, first, incorporating the process component in the model, in particular by using principles from dynamic systems, and second, by founding the model on adequate theoretical assumptions (for a full explanation of these principles, see Steenbeek & Van Geert, 2007).

A related issue concerns the number of parameters in the model. In model fitting in general, this number is primarily determined by a statistical criterion, in which the model is penalized for the number of parameters used (based on goodness-of-fit statistics such as Akaike's Information Criterion (AIC) or Bayesian Information Criterion (BIC)). In our case, the – considerable – number of parameters follows directly from the underlying theory. Parsimony becomes an issue if a single underlying theory leads to distinct models, in which case the most parsimonious model should be chosen.

In the present model, the control parameter values are based on status alone, and are thus similar for all dyads of the same status group. In reality, it is more likely that each type of dyad, which corresponds with a group of actual dyads, is represented by control parameters that consist of a range of values, instead of just one specific value. Testing the fit qualities of such ranges could be a next step in the further elaboration of the model.

A third remark concerns the use of the random component in the model. This random component is incorporated into the model in such a way that its influence is more or less constant, without dominating the influence of the values of the control parameter groups. However, this model random component can be no more than only an approximate representation of the empirical stochastic component, in which all kinds of variables exert an influence of varying strength. In further research, attempts should be made to estimate the magnitude of the influence of this random component.

A final remark concerns the empirical aspects. The sample on which the model fitting has been based does not claim to be representative for the population. The sample is relatively small, and is recruited from a single, semi-rural community. In addition, some dyads have been measured twice (see the section Participants). The sample has primarily been chosen with the intention of validating the dynamic model. In that case, representativeness is desirable but not strictly necessary. What is necessary, though, is that the sample is 'exemplary', in the sense that it must contain typical examples of popular, average and rejected children. We have no reasons to believe that the sample is flawed in this respect. A separate issue concerns the gender difference. On average, rejected girls, for instance, may be characterized by other

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

properties than rejected boys. So far the model does not take gender into account, and gender as such is not a control variable in itself. A question for further concern is how gender differences can be represented in terms of characteristic settings of the model's current control parameters.

3.5. General implications

The main message that we would like to convey is that the developmental study of social interaction must take the dynamics of social interaction as its starting point and that it must yield models and explanations of those dynamics. Whatever properties are collected from observing social interactions, those properties come about in a dynamic, iterative process of action–reaction that focuses on the dyad (or group, for that matter) as the target of research, the components of which are the individual agents. A model of the dynamics of a process requires that the model be recursively or iteratively applied. The strength of dynamic models does not particularly lie in the underlying content principles, because such principles are often taken from existing research and theory (see the action principles that underlie our model, for instance). The strength lies in the fact that such principles are transformed into a recursively applicable model of how things change from time step to time step. We have founded our model on a general theory of concern-driven action, but other perspectives are also possible. A model such as Crick and Dodge's (1994) social information processing model, for instance, describes principles for changing the values of the variables, which operate in the short term. To turn it into a dynamic model, the nature of the relationships between the variables leading to changes in those variables must be specified and applied in a recursive way.

Conceptualizing the process of interaction in the form of a recursive, dynamic model of how things change over time has a number of implications that are of importance to studying development in general and thus extend the scope of the current application to social interaction in particular.

The first implication concerns the relation between person and context, given the belief that context, i.e. the immediate environment, exerts a great influence on a person's development. The model defines the context as a dynamic phenomenon, co-constructed by the person's actions. The current dyadic interaction model describes a so-called coupled dynamics, with the actions of one child iteratively coupled to the actions of the other. In this regard, the context for the actions of one child is the actions of the other child, and vice versa. If such a notion of context is taken into account, it can help us put our empirical findings into a new perspective. For instance, our finding that in the current play context rejected children show more positive expressions and other-directed actions than higher-status children might in itself not be a spectacular finding, but it derives its significance from the fact that it illustrates the principles of interaction underlying the dynamic model. The general issue is that context and person are to a considerable extent dynamically interdependent, yet also independent enough to allow for flexibility and a variety of outcomes.

A second implication relates to the fact that a process dynamically unfolds over time, and thus requires that the processes' fundamental time-scales be distinguished and described. In the case of social interaction, the time scales concern the short-term time scale of real-time interaction and the long-term time scale of the development of social competence, power and status. A model must be formulated in such a way that the dynamic link between both time scales can be explicitly made. The distinction between dynamics at different time scales is a fundamental issue for developmental theory, which must be able to explain how the short-term process of action and interaction contributes to the long-term course of development, and also how the long-term course of development contributes to the particularities of action and interaction in real time. This linking of time scales is not only of theoretical, but certainly also of applied importance, given that interventions, for instance, amount to short-term actions carried out with a long-term goal in mind.

A third implication relates to the nature of the processes that occur in the short as well as in the long term. Actions, interactions and development are not the unfolding of internal properties in the individual. They are in the first place self-organizational processes that result from the dynamic intertwining of all factors involved. Self-organization has many appearances and does not cohere with an approach that looks for single solutions to how development really proceeds. Self-organization creates opportunities, enhancing the possibility of positive outcomes, but it can also impose constraints on processes, forcing them to change in ways that are not intended or desired. In order to understand how self-organization works, averaging over occasions and persons is not the adequate approach. It can only be understood if studied on its proper level and in its proper context, for instance the level of short-term dyadic interaction that constituted the topic of the current article.

The final implication relates to the nature and use of models. The function of a model is primarily to enhance our understanding of a phenomenon, and not in the first place to provide a best possible fit with data. This statement may seem a somewhat remarkable conclusion for an article that set out to empirically validate a model. A model that has no fit with the data is not very likely to enhance understanding, but fitting the data can occur in various ways. There can be a fit in the classical sense of direct correspondence with the data and making correct predictions, but there can also be a fit on the level of foreseeing qualitative patterns and properties, as was the case with the long-term model of rejection that we

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

briefly described in the discussion section. A great advantage of dynamic models of the sort we have presented in this article is that they take qualitative principles and insights of how things probably change, and put them together in a framework that yields an explicit description of a process, an occurrence of linked events over time. In this sense, they are clarifications of our theories and intuitions. Dynamic models can be used as convenient tools for (computationally) experimenting with various assumptions and observing the effects on the trajectories that the model produces. If based on general enough assumptions, dynamic models can be used as building blocks or starting points for applications to other fields. In a sense, the model of dyadic interaction presented and empirically validated in the current article presents a toy world. But like a good toy should, it is intended to capture the essence of the phenomenon in a way that our understanding can grasp and that preludes applications to a host of related phenomena.

Endnote

Fitting distributions per variable

Results are based on an alternative random model (with a non-overlapping parameter range). The question is: if qualitative (dis)similarity between model and random model is considered important, does it matter whether one uses specific parameter values, for instance based on theoretical assumptions, or will randomly chosen parameter values also suffice? Remember that in the random model that has just been tested, the range from which parameter values were randomly sampled shows a 50% overlap with the range covering the parameter sets based on a deliberate, theoretical choice (see the aforementioned explanation of taking twice the width of the theoretically justified range). We have just seen that this particular random range results in distributions that are qualitatively similar to those of the real parameter range, at least for a significant number of variables. If the values of the parameters do not matter, a comparable qualitative similarity should also result if we take parameters from an entirely different parameter range, for instance, one which is also twice as broad as the theoretically determined range, but which is moved to the left (i.e. covering the lower part of the possible parameter values). Note that the model of the three types of dyads is not nested in this alternative model.

Statistical method: is the same as described in the paragraph 'testing an arbitrary model'.

Results: All variables yield significant *p*-values ($p < .001$), except for one variable ('negative expressions' for the group of 'popular' dyads). This means that the real model distribution resembles the empirical distribution better than the random model for all variables but one and for the three types of dyad. This picture is confirmed by visual inspection of the data, which is illustrated in Steenbeek and van Geert (2005a). The difference between the real model and the alternative random model is much bigger than the difference resulting from the original random model. Thus, the question whether any randomly selected set of parameters can achieve a reasonable qualitative fit with the empirical distributions can be answered in the affirmative if the parameters are selected from a range of values that overlap with the theoretically funded range, and in the negative if the parameters come from a more peripheral, non-overlapping range.

References

Adler, P.A., & Adler, P. (1998). *Peer power: Preadolescent culture and identity*. New Brunswick, NJ: Rutgers University Press.

Ainsworth, M.D. (1979). Infant–mother attachment. *American Psychologist*, **34**, 932–937.

Arnold, M.B. (1961). *Emotion and personality. Vol. I: Psychological aspects*. London: Cassell & Co.; New York: Columbia University Press.

Asarnow, J.R., & Callan, J.W. (1985). Boys with peer adjustment problems: social cognitive processes. *Journal of Consulting and Clinical Psychology*, **53** (1), 80–87.

Asher, S.R. (1983). Social competence and peer status: recent advances and future directions. *Child Development*, **54**, 1427–1434.

Asher, S.R., & Coie, J.O. (Eds.) (1990). *Peer rejection in childhood*. Cambridge: Cambridge University Press.

Asher, S.R., & Dodge, K.A. (1986). Identifying children who are rejected by their peers. *Developmental Psychology*, **22** (4), 444–449.

Asher, S.R., Markell, R.A., & Hymel, S. (1981). Identifying children at risk in peer relations: a critique of the rate-of-interaction approach to assessment. *Child Development*, **52** (4), 1239–1245.

Asher, S.R., & Parker, J.G. (1989). The significance of peer relationship problems in childhood. In B.H. Schneider, G. Attili, J. Nade, & R.P. Weissberg (Eds.), *Social competence in developmental perspective* (pp. 5–23). Amsterdam: Kluwer Academic.

Asher, S.R., Singleton, L.C., Tinsley, B.R., & Hymel, S. (1979). A reliable sociometric measure for preschool children. *Developmental Psychology*, **15**, 443–444.

Austin, J.T., & Vancouver, J.B. (1996). Goal constructs in psychology: structure, process, and content. *Psychological Bulletin*, **120**, 338–375.

Bagwell, C.L., Schmidt, M.E., Newcomb, A.F., & Bukowski, W.M. (2001). Friendship and peer rejection as predictors of adult adjustment. In D.W. Nangle & C.A. Erdley (Eds.), *The role of friendship in psychological adjustment: New directions for child and adolescent development*, no. 91 (pp. 25–49). San Francisco, CA: Jossey-Bass.

Bak, P. (1996). *How nature works: The science of self-organized criticality*. New York: Springer-Verlag.

Balci, O. (1997). Principles of simulation model validation, verification, and testing. *Transactions of the Society for Computer Simulation International*, **14** (1), 3–12.

Bandura, A., & Walters, R. (1977). *Social learning and personality development*. London: Holt, Rinehart, and Winston.

Bargh, J.A., & Chartrand, T.L. (1999). The unbearable automaticity of being. *American Psychologist*, **54**, 462–479.

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Bargh, J.A., & Ferguson, M.J. (2000). Beyond behaviorism: on the automaticity of higher mental processes. *Psychological Bulletin*, **126**, 925–945.
Beilin, H., & Fireman, G. (2000). The foundation of Piaget's theories: mental and physical action. In H.W. Reese (Ed.), *Advances in child development and behavior*, Vol. 27 (pp. 221–246). San Diego, CA: Academic Press.
Bierman, K.L. (2004). *Peer rejection: Developmental processes and intervention strategies*. New York: Guilford Press.
Black, B., & Logan, A. (1995). Links between communication patterns in mother–child, father–child and child–peer interactions and children's social status. *Child Development*, **66**, 255–271.
Bowlby, J. (1980). *Attachment and loss. Vol. 3: Loss, sadness, and depression* (3 vols). New York: Basic Books.
Brown, J.F., Spencer, K., & Swift, S. (2002). A parent training programme for chronic food refusal: a case study. *British Journal of Learning Disabilities*, **30** (3), 118–121.
Brown, W.H., Odom, S.L., & Holcombe, A. (1996). Observational assessment of young children's social behavior with peers. *Early Childhood Research Quarterly*, **11**, 19–40.
Bruins, J. (1999). Social power and influence tactics: a theoretical introduction. *Journal of Social Issues, Special issue: Social influence and social power: Using theory for understanding social issues*, **55**, 7–14.
Buder, E.H. (1991). A nonlinear dynamic model of social interaction. *Communication Research*, **18** (2), 174–198.
Cabanac, M. (1992). Pleasure: the common currency. *Journal of Theoretical Biology*, **155**, 200.
Camras, L.A. (2000). Surprise! Facial expressions can be coordinative motor structures. In M.D. Lewis & I. Granic (Eds.), *Emotion, development, and self-organization: Dynamic systems approaches to emotional development* (pp. 100–124). Cambridge: Cambridge University Press.
Camras, L.A., Meng, Z., & Ujiie, T. (2002). Observing emotion in infants: facial expression, body behavior, and rater judgments of responses to an expectancy-violating event. *Emotion*, **2** (2), 179–193.
Carver, C.S., & Scheier, M.F. (1990). Origins and functions of positive and negative affect: a control-process view. *Psychological Review*, **97**, 19–35.
Casti, J.L. (1997). *Would-be-worlds: How simulation is changing the frontiers of science*. New York: John Wiley & Sons.
Chartrand, T.L., & Bargh, J.A. (1999). The chameleon effect: the perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, **76**, 893–910.
Christiansen, M.H., & Kirby, S. (2003). Language evolution: consensus and controversies. *Trends in Cognitive Sciences*, **7** (7), 300–307.
Cicchetti, D., Ganiban, J., & Barnett, D. (1991). Contributions from the study of high-risk populations to understanding the development of emotion regulation. In J. Garber & K.A. Dodge (Eds.), *The development of emotion regulation and dysregulation* (pp. 15–48). New York: Cambridge University Press.
Cillessen, A.H.N. (2006). Peer rejection: bridging theory and application. *Human Development*, **49** (1), 44–53.
Cillessen, A.H.N., & Rose, A.J. (2005). Understanding popularity in the peer system. *Current Directions in Psychological Science*, **14** (2), 102–105.
Cillessen, A.H.N., & ten Brink, P.W. (1991). Vaststelling van relaties met leeftijdgenoten [Assessment of children's peer relationships]. *Pedagogische Studiën*, **68** (1), 1–14.
Cillessen, A., & Mayeux, L. (2004). Sociometric status and peer group behavior: previous findings and current directions. In J.B. Kupersmidt & K.A. Dodge (Eds.), *Children's peer relations: From development to intervention* (pp. 3–20). Washington, DC: American Psychological Association.
Cirino, R.J., & Beck, S.J. (1991). Social information processing and the effects of reputational, situational, developmental, and gender factors among children's sociometric groups. *Merrill-Palmer Quarterly*, **37**, 561–582.
Coie, J.D., Dodge, K.A., & Coppotelli, H. (1982). Dimensions and types of social status: a cross-age perspective. *Developmental Psychology*, **18** (4), 557–570.
Coie, J.D., Dodge, K.A., & Kupersmidt, J.B. (1990). Peer group behavior and social status. In S.R. Asher & J.D. Coie (Eds.), *Peer rejection in childhood. Cambridge studies in social and emotional development* (pp. 17–59). New York: Cambridge University Press.
Collins, W.A., Maccoby, E.E., Steinberg, L., Hetherington, E.M., & Bornstein, M.H. (2000). Contemporary research on parenting: the case for nature and nurture. *American Psychologist*, **55** (2), 218–232.
Conte, R., Hegselmann, R., & Terna, P. (1997). *Simulating social phenomena*. Berlin: Springer-Verlag.
Conte, R., & Paolucci, M. (2001). Intelligent social learning. *Journal of Artificial Societies and Social Simulation*, **4** (1), from http://www.soc.surrey.ac.uk/JASSS/4/1/3.html.
Contreras, J.M., Kerns, K.A., Weimer, B.L., Gentzler, A.I., & Tomich, P.L. (2000). Emotion regulation as a mediator of associations between mother–child attachment and peer relationships in middle childhood. *Journal of Family Psychology*, **14** (1), 111–124.
Copeland, J.T. (1994). Prophecies of power: motivational implications of social power for behavioral confirmation. *Journal of Personality and Social Psychology*, **67**, 264–277.
Crick, N.R., & Dodge, K.A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, **115**, 74–101.
Christiansen, M.H., & Kirby, S. (2003). Language evolution: consensus and controversies. *Trends in Cognitive Sciences*, **7** (7), 300–307.
Delveaux, K.D., & Daniels, T. (2000). Children's social cognitions: physically and relationally aggressive strategies and children's goals in peer conflict situations. *Merrill-Palmer Quarterly*, **46**, 672–692.
Denham, S.A., McKinley, M., Couchoud, E.A., & Holt, R. (1990). Emotional and behavioral predictors of preschool peer ratings. *Child Development*, **61** (4), 1145–1153.
Dépret, E., & Fiske, S.T. (1999). Perceiving the powerful: intriguing individuals versus threatening groups. *Journal of Experimental Social Psychology*, **35**, 461–480.
Dishion, T.J., Bullock, B.M., & Granic, I. (2002). Pragmatism in modeling peer influence: dynamics, outcomes and change processes. *Development and Psychopathology*, **14** (4), 969–981.
Eder, D. (1985). The cycle of popularity: interpersonal relations among female adolescents. *Sociology of Education*, **58** (3), 154–165.
Edwards, R., Manstead, A.S., & MacDonald, C.J. (1984). The relationship between children's sociometric status and ability to recognize facial expressions of emotion. *European Journal of Social Psychology*, **14** (2), 235–238.
Efron, B. (1988). Bootstrap confidence intervals: good or bad? *Psychological Bulletin*, **104** (2), 293–296.

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Eisenberg, N., & Fabes, R.A. (1995). The relation of young children's vicarious emotional responding to social competence, regulation, and emotionality. *Cognition and Emotion*, **9** (2–3), 203–228.

Eisenberg, N., Fabes, R.A., Bernzweig, J., Karbon, M., Poulin, R., & Hanish, L. (1993). The relations of emotionality and regulation to preschoolers' social skills and sociometric status. *Child Development*, **64**, 1418–1438.

Eisenberg, N., Fabes, R.A., Guthrie, I.K., & Reiser, M. (2001). Dispositional emotionality and regulation: their role in predicting quality of social functioning. *Journal of Personality and Social Psychology*, **78** (1), 136–157.

Erdley, C.A., & Asher, S.R. (1996). Children's social goals and self-efficacy perceptions as influences on their responses to ambiguous provocation. *Child Development*, **67**, 1329–1344.

Felmlee, D.H. (2007). Application of dynamic systems analysis to dyadic interactions. In A.D. Ong & M. van Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 409–422). Oxford: Oxford University Press.

Felmlee, D.H., & Greenberg, D.H. (1999). A dynamic systems model of dyadic interaction. *Journal of Mathematical Sociology*, **23**, 155–180.

Flavell, J.H. (1999). Cognitive development: children's knowledge about the mind. *Annual Review of Psychology*, **50**, 21–45.

Fogel, A. (1993). *Developing through relationships: Origins of communication, self, and culture*. New York: Harvester Wheatsheaf.

Fogel, A. (2001). A relational perspective on the development of self and emotions. In H. Bosma & S. Kunnen (Eds.), *Identity and emotion; Development through self-organization* (pp. 93–115). Cambridge: Cambridge University Press.

Forsyth, D.R. (1990). *Group dynamics* (2nd edn.). Belmont, CA: Brooks/Cole Publishing.

French, J.R.P., & Raven, J.B. (1959). The bases of social power. In Dorwin Cartwright (Ed.), *Studies in social power* (pp.150–167). Oxford: University of Michigan Press.

Frijda, N.H. (1986). *The emotions: Studies in emotion and social interaction*. Cambridge: Cambridge University Press.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants, *Nature*, **415**, 755.

Gerrits, M.H. (2004). Perceived security, self-representation, and dyadic child–parent and child–peer play interactions: a longitudinal study. Unpublished doctoral dissertation, University of Utrecht, The Netherlands.

Gilbert, N., & Troitzsch, K.G. (1999). *Simulation for the social scientist*. Buckingham: Open University Press.

Gnepp, J. (1989). Personalized inferences of emotions and appraisals: component processes and correlates. *Developmental Psychology*, **25** (2), 277–288.

Goncu, A. (1993). Development of intersubjectivity in the dyadic play of preschoolers. *Early Childhood Research Quarterly*, **8**, 99–116.

Good, P.I. (1999). *Resampling methods: A practical guide to data analysis*. Boston, MA: Birkhauser.

Gottman, J., Gonso, J., & Rasmussen, B. (1975). Social interaction, social competence, and friendship in children. *Child Development*, **46**, 709–718.

Gottman, J.M., Guralnick, M.J., Wilson, B., Swanson, C.C., & Murray, J.D. (1997). What should be the focus of emotion regulation in children? A nonlinear dynamic mathematical model of children's peer interaction in groups. *Development and Psychopathology*, **9**, 421–452.

Gottman, J.M., Murray, J.D., Swanson, C.C., Tyson, R., & Swanson, K.R. (2002). *The mathematics of marriage: Dynamic nonlinear models*. Cambridge, MA: MIT Press.

Granic, I., Hollenstein, T., & Dishion, T.J. (2003). Longitudinal analysis of flexibility and reorganization in early adolescence: a dynamic systems study of family interactions. *Developmental Psychology*, **39** (3), 606–617.

Granic, I., & Hollenstein, T. (2003). Dynamic systems methods for models of developmental psychopathology. *Development and Psychopathology, Special issue: Conceptual, methodological, and statistical issues in developmental psychopathology: A special issue in honor of Paul E. Meehl*, **15** (3), 641–669.

Grolnick, W.S., Bridges, L.J., & Connell, J.P. (1996). Emotion regulation in two-year-olds: strategies and emotional expression in four contexts. *Child Development*, **67**, 928–941.

Haken, H. (1977). *Synergetics: An introduction*. Heidelberg: Springer-Verlag.

Haring, T.G., & Lovinger, L. (1989). Promoting social interaction through teaching generalized play initiation responses to preschool children with autism. *Journal of the Association for Persons with Severe Handicaps*, **14**, 58–67.

Hartup, W.W., & Laursen, B. (1999). Relationships as developmental contexts: retrospective themes and contemporary issues. In W.A. Collins & B. Laursen (Eds.), *Relationships as developmental contexts. The Minnesota symposia on child development*, vol. 30 (pp. 13–35). Mahwah, NJ: Lawrence Erlbaum Associaties.

Hauck, M., Fein, D., Waterhouse, L., & Feinstein, C. (1995). Social initiations by autistic children to adults and other children. *Journal of Autism and Developmental Disorders*, **25** (6), 579–595.

Hawley, P.H. (2002). Social dominance and prosocial and coercive strategies of resource control in preschoolers. *International Journal of Behavioral Development*, **26** (2), 167–176.

Hazen, N.L., & Black, B. (1984). Social acceptance: strategies children use and how teachers can help children learn them. *Young Children*, **39**, 26–36.

Hazen, N.L., & Black, B. (1990). Social status and patterns of communication in acquainted and unacquainted preschool children. *Developmental Psychology*, **26** (3), 379–387.

Hendrickson, J.M., Strain, P.S., Tremblay, A., & Shores, R.E. (1982). Interactions of behaviorally handicapped children: functional effects of peer social initiations. *Behavior Modification*, **6** (3), 323–353.

Hinde, R.A. (1997). *Relationships: A dialectical perspective*. Hove: Psychology Press.

Hollenstein, T., & Lewis, M.D. (2006). A state space analysis of emotion and flexibility in parent–child interactions. *Emotion*, **6** (4), 656–662.

Howe, M.L., & Lewis, M.D. (2005). The importance of dynamic systems approaches for understanding development. *Developmental Review*, **25** (3–4), 247–251.

Hubbard, J.A. (2001). Emotion expression processes in children's peer interaction: the role of peer rejection, aggression, and gender. *Child Development*, **72** (5), 1426–1438.

Hubbard, J.A., Dodge, K.A., Cillessen, A.H.N, Coie, J.D., & Swartz, D. (2001). The dyadic nature of social information processing in boys' reactive and proactive aggression. *Journal of Personality and Social Psychology*, **80** (2), 268–280.

Izard, C.E., Ackerman, B.P., Schoff, K.M., & Fine, S.E. (2000). Self-organization of discrete emotions, emotion patterns, and emotion–cognition relations. In M.D. Lewis & I. Granic (Eds.), *Emotion, development, and self-organization:*

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Dynamic systems approaches to emotional development (pp. 15–36). Cambridge: Cambridge University Press.

Jager, W. (2000). *Modelling consumer behaviour*. Doctoral dissertation, University of Groningen. Groningen: Universal Press.

Jager, W., Popping, R., & van de Sande, H. (2001). Clustering and fighting in two-party crowds: simulating the approach–avoidance conflict. *Journal of Artificial Societies and Social Simulation*, **4** (3), 1–18.

Johnston, V. (2003). The origins and function of pleasure. *Cognition and Emotion*, **17**, 167–179.

Kindermann, T. (2003). Children's relationships and development of person–context relations. In J. Valsiner & K. Connolly (Eds.), *Handbook of developmental psychology* (pp. 407–430). London: Sage Publications.

Klatt, K.P., & Morris, E.K. (2001). The Premack principle, response deprivation, and establishing operations. *Behavior Analyst*, **24** (2), 173–180.

Kohler, T.A., & Gumerman, G.J. (2000). *Dynamics in human and primate societies: Agent-based modeling of social and spatial processes*. New York: Oxford University Press.

Krantz, M. (1982). Sociometric awareness, social participation, and perceived popularity in preschool children. *Child Development*, **53** (2), 376–379.

Kupersmidt, J.B., Coie, J.D., & Dodge, K.A. (1990). The role of poor peer relationships in the development of disorder. In S. Asher & J.D. Coie (Eds.), *Peer rejection in childhood. Cambridge studies in social and emotional development* (pp. 274–305). New York: Cambridge University Press.

Kupersmidt, J.B., & Dodge, K.A. (2004). *Children's peer relations: From development to intervention*. Washington, DC: American Psychological Association.

Kupersmidt, J.B., & deRosier, M.E. (2004). How peer problems lead to negative outcomes: an integrative mediational model. In J.B. Kupersmidt & K.A. Dodge (Eds.), *Children's peer relations: From development to intervention* (pp. 119–138). Washington, DC: American Psychological Association.

Ladd, G.W. (1999). Peer relationships and social competence during early and middle childhood. *Annual Review of Psychology*, **50**, 333–359.

Ladd, G.W., & Le Sieur, K.D. (1995). Parents and children's peer relationships. In M.H. Bornstein (Ed.), *Handbook of parenting, Vol. 4: Applied and practical parenting* (pp. 377–409). Hillsdale, NJ: Lawrence Erlbaum Associates.

Levy, D.A., & Nail, P.R. (1993). Contagion: a theoretical and empirical review and reconceptualization. *Genetic, Social & General Psychology Monographs*, **119**, 233–284.

Lewis, M.D. (1995). Cognition–emotion feedback and the self-organization of developmental paths. *Human Development*, **38** (2), 71–102.

Lewis, M.D. (2000a). The promise of dynamic systems approaches for an integrated account of human development. *Child Development*, **71** (1), 36–44.

Lewis, M.D. (2000b). Emotional self-organization at three time scales. In M.D. Lewis & I. Granic (Eds.), *Emotion, development, and self-organization: Dynamic systems approaches to emotional development* (pp. 37–69). Cambridge: Cambridge University Press.

Lewis, M.D. (2002). Interacting time scales in personality (and cognitive) development: intentions, emotions, and emergent forms. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 183–212). New York: Cambridge University Press.

Lewis, M., & Granic, I. (Eds.) (2000). *Emotions, development, and self-organization: Dynamic systems approaches to emotional development*. Cambridge: Cambridge University Press.

Lopes, P.N., Salovey, P., & Côté, S. (2005). Emotion regulation abilities and the quality of social interaction. *Emotion*, **5** (1), 113–118.

Lyytinen, P. (1991). Peer interaction in children's dyadic play. *Early Child Development and Care*, **71**, 105–115.

Maassen, G.H., Akkermans, W., & Van der Linden, J.L. (1996). Two-dimensional sociometric status determination with rating scales. *Small Group Research*, **27**, 56–78.

Maassen, G.H., Steenbeek, H.W., & van Geert, P.L.C. (2004). Stability of three methods for two-dimensional sociometric status determination based on the procedure of Asher, Singleton, Tinsley and Hymel. *International Journal of Social Behavior and Personality*, **32** (6), 535–550.

Manley, B.F. (1997). *Randomization, bootstrap and Monte Carlo methods in biology* (2nd edn.). Boca Raton, FL: Chapman and Hall.

Martin, L.C., Fabes, R.A., Hanish, L.D., & Hollenstein, T. (2005). Social dynamics in the preschool. *Developmental Review*, **25**, 299–327.

Mascolo, M.F. (2005). Change processes in development: the concept of coactive scaffolding. *New Ideas in Psychology*, **23**, 185–196.

Mascolo, M.F., Harkins, D., & Harakal, T. (2000). The dynamic construction of emotion: varieties in anger. In M.D. Lewis & I. Granic (Eds.), *Emotion, development, and self-organization: Dynamic systems approaches to emotional development* (pp. 125–152). Cambridge: Cambridge University Press.

Matsui, M., Muto, T., & Kadoyama, M. (2001). The initiation of peer interaction: analysis of preschoolers' free play. *Japanese Journal of Developmental Psychology*, **12** (3), 195–205.

Messinger, D.S. (2002). Positive and negative infant facial expressions and emotions. *Current Directions in Psychological Science*, **11**, 1–6.

Miller, A.L., & Olson, S.L. (2000). Emotional expressiveness during peer conflicts: a predictor of social maladjuststment among high-risk preschoolers. *Journal of Abnormal Child Psychology*, **28**, 339–352.

Mize, J., & Ladd, G.W. (1990). Toward the development of successful social skills training for preschool children. In S.R. Asher & J.D. Coie (Eds.), *Peer rejection in childhood* (pp. 338–361). Cambridge: Cambridge University Press.

Morales, M., Mundy, P., & Crowson, M.M. (2005). Individual differences in infant attention skills, joint attention, and emotion regulation behaviour. *International Journal of Behavioral Development*, **29** (3), 259–263.

Murphy, B.C., & Eisenberg, N. (2002). An integrative examination of peer conflict: children's reported goals, emotions, and behaviors. *Social Development*, **11**, 534–557.

Nail, P.R., MacDonald, G., & Levy, D.A. (2000). Proposal of a four-dimensional model of social response. *Psychological Bulletin*, **126**, 454–470.

Neumann, R., & Strack, F. (2000). 'Mood contagion': the automatic transfer of mood between persons. *Journal of Personality and Social Psychology*, **79**, 211–223.

Newcomb, A.F., Bukowski, W.M., & Pattee, L. (1993). Children's peer relations: a meta-analytic review of popular, rejected, neglected, controversial, and average sociometric status. *Psychological Bulletin*, **113** (1), 99–128.

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

Noldus Information Technology (1999). *The Observer video pro Version 4.0 for windows*. Wageningen, The Netherlands.
Olthof, T., & Kunnen, S. (2000). Simulating mother–child interaction: exploring two varieties of a non-linear dynamic systems approach. *Infant and Child Development*, **9** (1), 33–60.
Packer, M.J., & Scott, B. (1992). The hermeneutic investigation of peer relations. In L.T. Winegar & J. Valsiner (Eds.), *Children's development within social context* (pp. 75–111). Hillsdale, NJ: Erlbaum.
Panksepp, J. (2000). The riddle of laughter: neural and psycho-evolutionary underpinnings of joy. *Current Directions in Psychological Science*, **9**, 183–186.
Parke, R.D., & Ladd, G.W. (1992). *Family–peer relationships: Modes of linkage*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Parke, R.D., Simpkins, S.D., & McDowell, D.J. (2002). Relative contributions of families and peers to children's social development. In P.K. Smith & C.H. Hart (Eds.), *Blackwell handbook of childhood social development* (pp. 156–177). Malden, MA: Blackwell Publishing.
Parker, J.G., & Asher, S.R. (1987). Peer relations and later personal adjustment: are low-accepted children at risk? *Psychological Bulletin*, **102** (3), 357–389.
Parker, J.G., & Gottman, J.M. (1989). Social and emotional development in a relational context: friendship interaction from early childhood to adolescence. In T.J. Berndt & G.W. Ladd (Eds.), *Peer relationships in child development* (pp. 95–131). Oxford: John Wiley & Sons.
Pepler, D.J., & Craig, W.M. (1998). Assessing children's peer relationships. *Child Psychology and Psychiatry Review*, **3** (4), 176–182.
Pettit, G.S., Brown, E.G., & Mize, J. (1998). Mothers' and fathers' socializing behaviors in three contexts: links with children's peer competence. *Merrill-Palmer Quarterly*, **44** (2), 173–193.
Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant*. Neuchâtel. Paris: Delachaux & Niestlé.
Pickard, J.K. (2004). Contextual variability in early adolescents' state masculinity, femininity and peer interaction goals. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, **65** (5-B), 2670.
Powers, W.T. (2005). *Behavior: The control of perception* (2nd edn. rev. & exp.). Los Altos, CA: Benchmark Press.
Preston, S.D., & de Waal, F.B.M. (2002). Empathy: its ultimate and proximate bases. *Behavioral and Brain Sciences*, **25**, 1–72.
Prinstein, M.J., & La Greca, A.M. (2004). Childhood peer rejection and aggression as predictors of adolescent girls' externalizing and health risk behaviors: a 6-year longitudinal study. *Journal of Consulting and Clinical Psychology*, **72** (1), 103–112.
Rabiner, D.L., & Gordon, L. (1992). The coordination of conflicting social goals: differences between rejected and nonrejected boys. *Child Development*, **63**, 1344–1350.
Raven, B.H. (1992). A power/interaction model of interpersonal influence: French and Raven thirty years later. *Journal of Social Behavior and Personality*, **7**, 217–244.
Reis, H.T., Collins, W.A., & Berscheid, E. (2000). The relationship context of human behavior and development. *Psychological Bulletin*, **126**, 844–872.
Renshaw, P.D., & Asher, S.R. (1983). Children's goals and strategies for social interaction. *Merril-Palmer Quarterly*, **29**, 353–374.
Rizzolatti, G., & Craighero, L. (2004). The mirror neuron system. *Annual Review of Neuroscience*, **27**, 169–192.
Rose, A.J., & Asher, S.R. (1999). Children's goals and strategies in response to conflicts within a friendship. *Developmental Psychology*, **35**, 69–79.
Rose-Krasnor, L. (1997). The nature of social competence: a theoretical review. *Social Development*, **6**, 111–135.
Roseman, I., & Evdokas, A. (2004). Appraisals cause experienced emotions: experimental evidence. *Cognition and Emotion*, **18**, 1–28.
Roseman, I.J., Wiest, C., & Swartz, T.S. (1994). Phenomenology, behaviors, and goals differentiate discrete emotions. *Journal of Personality and Social Psychology*, **67**, 206–221.
Rubin, K.H., Bukowski, W., & Parker, J.G. (1998). Peer interactions, relationships, and groups. In M. Damon & N. Eisenberg (Eds.), *Handbook of child psychology. Volume III. Social, emotional, and personality development* (pp. 619–700). New York: Wiley & Sons.
Saarni, C. (2006). Emotion regulation and personality development in childhood. In D.K. Mroczek & T.D. Little (Eds.), *Handbook of personality development* (pp. 245–262). Mahwah, NJ: Lawrence Erlbaum Associates.
Sandstrom, M.J., & Zakriski, A.L. (2004). Understanding the experience of peer rejection. In J.B. Kupersmidt & K.A. Dodge (Eds.), *Children's peer relations: From development to intervention* (pp. 101–118). Washington, DC: American Psychological Association.
Scherer, K.R. (1999). Appraisal theory. In T. Dalgleish & M.J. Power (Eds.), *Handbook of cognition and emotion* (pp. 637–663). New York: Wiley.
Schroeder, M. (1991). *Fractals, chaos, power laws*. New York: Freeman.
Schuele, C.M., Rice, M.L., & Wilcox, K.A. (1995). Redirects: a strategy to increase peer initiations. *Journal of Speech and Hearing Research*, **38**, 1319–1333.
Simeo-Munson, J.A. (2000). Differences among peer status groups on levels of loneliness, social competence, and empathy. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, **61** (2-B).
Singer, E., Doornenbal, J., & Okma, K. (2004). Why do children resist or obey their foster parents? The inner logic of children's behavior during discipline. *Child Welfare*, **83** (6), 581–610.
Snyder, M., & Kiviniemi, M.T. (2001). Getting what they came for: how power influences the dynamics and outcomes of interpersonal interaction. In A.Y. Lee-Chai & J.A. Bargh (Eds.), *Use and abuse of power: Multiple perspectives on the causes of corruption* (pp. 133–155). New York: Psychology Press.
Staller, A., & Petta, P. (2001). Introducing emotions into the computational study of social norms: a first evaluation. *Journal of Artificial Societies and Social Simulation*, **4** (1), retrieved from http://www.soc.surrey.ac.uk/JASSS/4/1/2.html.
Steenbeek, H. (2006). Modeling dyadic child–peer interactions: sociometric status, emotional expressions and instrumental actions during play. Doctoral Dissertation, University of Groningen, Groningen, The Netherlands.
Steenbeek, H.W., & van Geert, P.L.C. (2005a). A dynamic systems model of dyadic interaction during play of two children. *European Journal of Developmental Psychology*, **2** (2), 105–145.
Steenbeek, H.W., & van Geert, P.L.C. (2005b). 'Do you still

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.

like to play with him?' Variability and the dynamic nature of children's sociometric ratings. Manuscript submitted for publication. Groningen, The Netherlands; University of Groningen, Section of Clinical and Developmental Psychology.

Steenbeek, H.W., & van Geert, P.L.C. (2007). A theory and dynamic model of dyadic interaction: concerns, appraisals and contagiousness in a developmental context. *Developmental Review*, **27** (1), 1–40.

Tanta, K.J., Deitz, J.C., White, O., & Billingsley, F. (2005). The effects of peer-play level on initiations and responses of preschool children with delayed play skills. *American Journal of Occupational Therapy*, **59**, 437–445.

Thelen, E., & Smith, L.B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thelen, E., & Smith, L.B. (1998). Dynamic systems theories. In W. Damon & R.M. Lerner (Eds.), *Handbook of child psychology: Volume 1: Theoretical models of human development* (5th edn.; pp. 563–634). New York: John Wiley & Sons.

Thompson, D.E., & Russell, J. (2004). The ghost condition: imitation versus emulation in young children's observational learning. *Developmental Psychology*, **40**, 882–889.

Todman, J.B., & Dugard, P. (2001). *Single-case and small-n experimental designs: A practical guide to randomization tests*. Mahwah, NJ: Erlbaum.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P.J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind and Language*, **18**, 121–147.

Troop-Gordon, W., & Asher, S.R. (2005). Modifications in children's goals when encountering obstacles to conflict resolution. *Child Development*, **76**, 568–582.

Uebersax, J. (2005). Statistical methods for rater agreement. Retrieved from http://ourworld.compuserve.com/homepages/jsuebersax/agree.htm

Valsiner, J. (1998). The development of the concept of development: historical and epistemological perspectives. In W. Damon & R.M. Lerner (Eds.), *Handbook of child psychology: Volume 1: Theoretical models of human development* (5th edn.; pp. 189–232). New York: John Wiley & Sons.

Van Dijkum, C., DeTombe, D., & van Kuijk, E. (1999). *Validation of simulation models*. Amsterdam: SISWO (SISWO Publication 403).

Van Geert, P.L.C. (1994). *Dynamic systems of development: Change between complexity and chaos*. New York: Harvester Wheatsheaf.

Van Geert, P.L.C. (2003). Dynamic system approaches and modeling of developmental processes. In J. Valsiner & K. Connolly (Eds.), *Handbook of developmental psychology* (pp. 640–672). London: Sage Publications.

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** (3–4), 408–442.

Van Geert, P., & Steenbeek, H. (2006). The dynamics of scaffolding. *New Ideas in Psychology*, **23**, 115–128.

Van Geert, P.L.C., & van Dijk, M. (2003). The problem of inter-observer reliability in ambiguous observation data. *First Language*, **23** (3), 259–284.

van Lier, P.A., & Hoeben, S.M. (1991). Begeleiding op school van kinderen met relatieproblemen: het belang van afstemming en inbedding [Treatment of children with problems in peer relations in school]. *Pedagogische Studiën*, **68** (2), 86–100.

Visser, M., Singer, E., Van Geert, P., & Kunnen, S. (2006). What makes children behave aggressively? A study of the inner logic in narratives of Dutch children in special education. Manuscript submited for publication.

Vosk, B.N., Forehand, R., & Figueroa, R. (1983), Perception of emotions by accepted and rejected children. *Journal of Behavioral Assessment*, **5** (2), 151–160.

Want, S.C., & Gattis, M. (2005). Are 'late-signing' deaf children 'mindblind'? Understanding goal directedness in imitation. *Cognitive Development*, **20** (2), 159–172.

Warner, R. (1992). Cyclicity of vocal activity increases during conversations: support for a nonlinear systems model of dyadic social interaction. *Behavioral Science*, **37** (2), 128–138.

Weisstein, E.W. (1999). *CRC concise encyclopedia of mathematics*. Boca Raton, FL: CRC Press.

Wheeler, L. (1966). Toward a theory of behavioral contagion. *Psychological Review*, **73**, 179–192.

Woodward, A.L., Sommerville, J.A., & Guajardo, J.J. (2001). How infants make sense of intentional action. In B.F. Malle, L.J. Moses, & D.A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 149–169). Cambridge, MA: MIT Press.

Received: 7 October 2005
Accepted: 16 January 2007

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.