Uncertainty in developmental psychology

van der Maas, H.L.J.; Raijmakers, M.E.J.

Published in:
Chance and Uncertainty. Their roles in various disciplines

Citation for published version (APA):
Uncertainty in developmental psychology

Han L.J. van der Maas & Maartje E.J. Raijmakers
Developmental Psychology
UVA

Herziene versie 20-12-94

1. Introduction

The development of mind, as studied in developmental psychology, is as complex as the mind itself. The role of chance in change is large and has different sources and consequences. Predictability of the development of psychological properties, like intelligence and personality, is low. The traditional explanations of this lack of predictability are based on the idea that there are too many partially unknown or unmeasurable variables (both genetic and environmental), on the idea of free will and on the idea of inherent stochasticity in the brain.

Recently an additional framework of explanation has become available. This framework concerns non-linear dynamic system theory which consists of a number of partly overlapping mathematical theories like chaos theory, catastrophe theory, synergetics, non-equilibrium thermodynamics, bifurcation theory and dynamic system theory. In this mathematical research the properties of rather simple but non-linear difference and differential equations are studied. The new insights generated by this research have been successfully integrated in natural science and slowly enter the field of social science.

The studied remarkable dynamic properties are chaos in deterministic systems, sudden discontinuities and self-organisation. They are directly relevant for developmental psychology. In past years psychologists have started to apply models and techniques of non-linear dynamic system theory. Here an introduction will be given to these attempts and the typical problems we meet will be discussed.

It will be shown that the application of non-linear dynamic systems theory gives rise to new interesting research and new insights in the role of chance and uncertainty in development. The new explanation of why developmental psychologists fail to predict development is analogous to the explanation that weather men give nowadays. The explanation is that long-term prediction is excluded as a matter of principle. Even when the brain could be understood as a simple (low number of variables) deterministic (but non-linear) system, the accumulation of tiny measurement errors will exclude the possibility of long-term prediction. Though this aspect has got a lot of attention in popular publications there is another side. That is, non-linear dynamic system theory predicts not only divergence of processes, but also convergence of processes. In systems that are
characterized by a one point attractor, all developmental paths converge to this position whatever their initial situation.

A simple example of a point attractor is the pendulum losing energy until it reaches its final equilibrium state. The initial position and velocity are unimportant for this final state. Convergence is also found in the possibility of self-organisation. Under certain conditions systems first lose stability and then transform to higher structured forms (by using energy from outside). These diverging and converging tendencies have a complex relation in the development of living systems (Kauffman, 1993). How this occurs is the subject of a large body of very recent research in biology, mathematics, computer science and psychology. This research might result in a new metaphor for understanding the developing human mind, namely the metaphor of the weather.

2. Deterministic chaos

2.1 Van Geert model

Deterministic chaos can be explained by a simple example, the so-called Verhulst equation

\[ L_{t+1} = r L_t (1 - L_t / K) . \]

\( L_{t+1} \) is the level of the variable of interest (for instance number of flies at time \( t+1 \)) and is a function of \( L_t \) of the foregoing generation, \( r \) is the growth parameter and \( K \) is the upper level of \( L \). The first part, \( r L_t \), causes exponential growth, but is slowed down by the second part which gets important if \( L_t \) reaches \( K \). This equation is a difference equation which means that it uses discrete instead of continuous time. This choice can be based on the fact that propagation takes place in generations. For different values of \( r \) between 0 and 4 this non-linear difference equation shows all kinds of behaviour, varying from extinction, to logistic growth to a stable level, periodic behaviour and chaos. This is shown in table 1 in which the first 20 iterations of the Verhulst equation are displayed for \( K=1 \) and various values of \( r \). \( L_0 \) is the initial level of \( L \).

<table>
<thead>
<tr>
<th>( r )</th>
<th>( L_1 )</th>
<th>( L_2 )</th>
<th>( L_3 )</th>
<th>( L_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>3.0</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>3.9</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>4.0</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Much more can be said about this equation's behaviour, but for present purposes this demonstration of types of behaviour suffices. In terms of populations of organisms all these types of behaviour are interpretable and observed in real populations. The divergence of \( L \) for \( r = 4 \) caused by arbitrary small initial differences has become famous in weather
forecasting since computer models of the weather appeared to show deterministic chaos too.

An obvious way to introduce deterministic chaos in psychology is by using the Verhulst equation as a psychological model. This has been done by Van Geert (1991). Van Geert assumes that there exists an equivalence between the growth of populations of organisms and the development of cognitive strategies, words, symbols, ideas etc. For instance, L is the number of words in a child's vocabulary, r is its growth rate and K is the carrying capacity defined by constraining factors or limited resources.

This way all described behaviours of the Verhulst equation could occur. Van Geert introduces all kinds of modifications of the basic equation to model competitive, supportive and mixed relations between several growers (for instance types of words). By these modifications many more typical kinds of behaviour can be understood and simulated. The theoretical importance of Van Geert's work is that it demonstrates that even simple non-linear dynamic models give rise to very complex unpredictable behaviour. The practical limitation is that developmental psychologists lack the technical possibilities to observe such behaviour in real development. More generally, it is difficult to test Van Geert's models with empirical data. The estimation of parameters and the statistical comparison to other growth models are problematic issues.

Finally, we could criticize the model itself. Can we give hypotheses on the value of r, the growth rate, which influences the form of the predicted curve strongly? Why should a difference equation be preferred above a differential equation (there are no generations)? If the differential Verhulst equation (assuming continuous time) is used, deterministic chaos will not occur. Yet, Van Geert models show clearly that the non-linear dynamic system models can change our ideas on psychological development importantly.

2.2 Neural networks: Autoassociater

Another way to demonstrate the possible presence of deterministic chaos is to show its occurrence in existing models. Especially recent popular neural network models are appropriate for this goal. Neural network models apply to the behaviour of neurons or groups of neurons of which the brain is built. Neural network models do not follow the traditional computer metaphor but start with the brain itself. Traditional computer models in artificial intelligence, assume the presence of symbols and reasoning, whereas in neural networks such higher order phenomena should emerge. There are many different neural network models, from little to more realistic, from little to more effective and powerful, and from little to more complex. A general interesting aspect for developmental psychologists is that neural networks are all learning models. Other general characteristics are that learned knowledge is distributed over the connections and that information processing takes place in a parallel way.
A very simple neural network model is the autoassociator.

Three neurons are displayed. Input enters from the left side and is distributed by the connections. Each unit gives an output depending on the input and the activity given by the connections. Connections can be positive or negative, having an excitatory or an inhibitory influence, respectively. Connections can be changed according to a learning rule. If we present the network repeatedly with an input pattern of -1 and 1, the connections will change until this pattern is learned. If we then present the network a partial pattern with some missing information, the complete output pattern will be generated by the connections. If we train the network with three basic patterns which are presented with some random noise, the network learns these three basic patterns and classifies new patterns correctly.

The autoassociator is defined by two equations which determine the change of the connections and the change of the activities of the neurons.

**Change of connections (learning rule):**

\[ W_{ij}(k+1) = (1-D_w) W_{ij}(k) + \eta a_i(k) a_j(k). \]

where \( W_{ij} \) is the connection or weight constant modulating the activation of unit \( i \) by unit \( j \), \( a_i \) the activation of unit \( i \), \( I_i \) the external input of unit \( i \), and \( k \) indexes the recursion step. \( D_w \) is a decay parameter of the connections and \( \eta \) notates the learning rate. This equation is based on the Hebbian learning rule. This rule means that if activities of neurons correlate, the strength of connection between these neurons increases. The rate of this increase depends on the learning parameter.

**Change of activations:**

\begin{align*}
    a_i(t+1) &= (1-D_a) a_i(t) + E (\sum m [a_m(t) W_{im}(t)] + I_i(t)) (1-a_i(t)), \\
    &\text{if } \sum m [a_m(t) W_{im}(t)] + I_i(t) > 0. \\
    a_i(t+1) &= (1-D_a) a_i(t) + E (\sum m [a_m(t) W_{im}(t)] + I_i(t)) (1+a_i(t)), \\
    &\text{if } \sum m [a_m(t) W_{im}(t)] + I_i(t) < 0.
\end{align*}

where \( D_a \) is the decay parameter of the activations, \( E \) is the excitation parameter and \( t \) denotes the iteration step within each recursion. The second equation is repeated several times before the first equation is updated. Repeated application of the second equation
within each recursion $k$ shows that $a_i$ is a non-linear function of $W_{ij}$, while $W_{ij}$ is a nonlinear function of all $W$.

By computer simulations Van der Maas, Verschure, and Molenaar (1990) have shown that the learning parameter $\eta$ has a function which is equivalent to the growth rate parameter in the Verhulst model. For high values of the learning parameter oscillatory and chaotic behaviour of activities and weights occur.

A problem is that this behaviour has no direct functional interpretation. Skarda and Freeman (1987) suggest that chaos might be a flexible rest state which will disappear when some function has to be performed. The only claim is that even in very simple neural networks deterministic chaos occurs. Hence, the fact that this network is much too simple is not a real objection.

2.3 Empirical data

To prove that chaos is indeed relevant for psychology, we have to detect chaos in empirical time-series. The requisite techniques are advanced but not very applicable. One needs a lot (e.g. thousands) of reliable repeated measurements to test for the presence of chaos. Hence, these techniques have been predominantly used for psychophysiological variables. In the last ten years many publications present analyses of the chaotic dimension of, for instance, brain activity (EEG) and heart activity (ECG). Though the statistical tools for chaos detection in time-series are still in full development, there are sound reasons to believe that the irregularities in these signals are chaotic.

An example associated with developmental psychology is the study of cyclic motor activity (CM). Robertson, Cohen, and Mayer-Kress (1993) define CM as spontaneous and comprised of general movements of the limbs, trunk and head, and more isolated or stereotypes movements, with a cycle time on the order of a few minutes or less. This irregular cyclic motor activity emerges during gestation and does not alter by the birth itself. To investigate whether the irregularities in cyclic activity are random noise, or part of the underlying process and hence perhaps chaotic, time-series of motor activity of an awake 2-months old baby are analysed. They conclude that CM is indeed chaotic (with some caution), and that the irregularities are not random noise.

2.4 Measurement
The occurrence of deterministic chaos in various psychophysiologival variables shows that this strongly diverging tendency plays an important role in brain functioning. The presented models show that the appearance of chaos in psychological variables should not amaze us.

Besides consequences for predictability chaos could have implications for the way we think about measurement. The measurement of psychological variables by questionnaires, tests etc., is beset with complicated validity and reliability problems. Technically reliability is often understood as stability. Items of the same test as well as repeated measurements of the test should correlate. When the variable under study is governed by non-linear dynamics we can not be sure whether tests are unreliable or that there are large fluctuations in the variable (Van Geert, 1991). The irregularities in the cyclic motor activity are an example of this.

3. Self-organisation

3.1 Spiral waves

Though there are important implications of deterministic chaos for the study of psychological development they seem to be rather negative. Instability, lack of predictability, diverging developmental routes, etc. are generally not very productive insights. It might be that chaos has a function in terms of flexibility and adaptivity, as Skarda and Freeman suggest, or that chaos causes variation necessary for selection processes in, for instance, motor development.

However, non-linear dynamic system theory is also about converging tendencies. For the point attractors (see first columns of the time-series of the Verhulst equation) the initial condition does not matter. In all cases the same final state is reached.

The cyclic attractors are of more interest since they can be associated with self-organisation. To explain what self-organisation is we first discuss a biological example. This famous example concerns the evolution of self-replicating macro-molecules like RNA. The problem is that mutations destabilize the molecules when they become too long. The solution presented by Eigen and Schuster (1979) is the hypercycle. A hypercycle consist of a series of self-replicating macro-molecules that catalyse each other in a circle. Molecule A catalyses B, which catalyses C etc., where the last type of molecule catalyses A again. For this process equations quite similar to the Verhulst equations used by Van Geert can be used. It can be shown that the hypercycle is more stable and longer macro-molecules can be created.

The criticism of this model is that parasites, molecules that use the catalisation of one of the molecules of the hypercycle but do not catalyse another, destabilize the process. The parasite will quickly replace all other molecules.
Boerlijst and Hogeweg (1991a, 1991b) present an elegant solution for this problem. They use the paradigm of Cellular Automata (CA). In CA the process of self-replication is modelled on a grid with only local rules. That is, on each point of the grid the appearance and disappearance of molecules is based on the local presence of the molecule itself and the catalysing molecules. Whereas differential and difference equations model only the amount of molecules, the CA also include spatial dimensions.

This has an impressive effect. By coding nine molecules by different colours, and displaying these molecules on a computer screen, a grid of 300 * 300 points or cells, we can follow the evolution of the molecules visually. It appears that a spatial organisation of molecules emerges in the form of spiral waves (for figures, see Boerlijst and Hogeweg, 1991b). These rotating spirals isolate or destroy the parasites by pushing the parasites away from their catalysator resulting in a new level of stability.

The occurrence of spirals adds a new higher order level of description to the hypercycle model in the form of spatial structures. If we only look at the local rules (difference equations like the Verhulst equation) the parasite is the most fit molecule and we should predict that this molecule will win in the battle. It appears that the less fit, altruistic, molecules, beat the parasite. Hence, by adding a spatial dimension to the model of Eigen and Schuster a very functional form of self-organisation takes place.

3.2 Neural oscillations in the visual cortex

This nice form of self-organisation in a biological system gives an example of what might be possible in psychological systems. Yet, whereas self-organisation is often put forward as an important developmental mechanism, convincing evidence for this phenomenon is lacking. As before, we can look at models and real data.

Concerning models we can take another look at neural network models. Self-organisation as in CA is related with the complex cyclic behavior of non-linear dynamic systems that lays between the simple point attractors and the chaotic regimes (Langton, 1986). As has been dicussed above in the context of the autoassociator, neural networks show such periodic behavior.

An example of a neural network in which a functional form of cyclic behaviour occurs is the Schuster and Wagner (1990) model: the neural oscillator. Among others, Gray et al (1991) have shown empirically that external stimuli can induce oscillations of neurons in the visual cortex of cats and monkeys. Moreover, these oscillations appear to synchronize over long distances if the stimulus is large, for example, a long bar. If, however, two distinct small bars with the same intensity are presented synchronization occurs only within the group of neurons that are stimulated by one bar, but not between the two stimulated groups of neurons. Moreover, neurons that are stimulated weakly do not synchronize.
Schuster and Wagner have simulated this phenomenon by means of a simple neural network model of connected cortical columns, so-called neural oscillators. One neural oscillator consists of two groups of neurons: a group of excitatory neurons and a group of inhibitory neurons. Connections exist within the excitatory group, within the inhibitory group, and between the excitatory and the inhibitory group in both directions. Only the excitatory neurons receive external input. It appears that within a specific range of the input strength the neurons oscillate (cyclic attractor). If the input strength lies outside this range, the activity of the neurons reaches a point attractor. In addition, the activity makes discontinuous jumps between these attractor states due to gradual variation of the input strength.

In order to model a layer of cortical columns, single oscillators are coupled in a row. The connections between oscillators are distributed sparsely. The probability that a connection exists between two oscillators depends on their mutual distance. An external input is presented at the model by varying the strength of the external input of individual oscillators along the row. A stimulus consisting of two distinct bars presented at a row of 100 neurons is represented in the model by, for example, giving the first 20 neurons a low external input, neurons 21 to 40 a high external input, neurons 41 to 60 a low input, neurons 61 to 80 a high input, and neurons 81 to 100 a low input. The resulting behaviour of the model stimulated in this way is shown in Figure 2a.

The dashed line represents the intensity of the input for each oscillator in the row, the solid line represents the phase of the oscillation of each oscillator. Firstly, figure 2a, in which two distinct input bars are presented, shows that within the stimulated groups phases of the oscillators are locked, that is, oscillators synchronize. Whereas outside those groups the phases of the oscillations are unlocked. Secondly, the highly stimulated groups are desynchronized. In contrast, figure 2b, in which one long input bar is presented, shows that the oscillators at long distance synchronize if they are stimulated by the same bar, although they do not have direct connections. This self-organizing behaviour of connected neural oscillators possibly plays a role in foreground to background separation of visual inputs.

The constructive tendency of self-organisation in the model of Schuster and Wagner (1990) is important at least at the neuronal level. We might ask whether self-organisation also occurs at the level of symbolic reasoning and other psychological processes. The present neural network models often fail to model higher order psychological processes.

3.3 Transitions
Fortunately, dynamic system theory yields more ways to look at self-organisation. As Robertson et al. (1993) note, in case of many complex processes the state variables and precise dynamics are unknown. Concerning cyclic motor activity Robertson et al. (1993) show how chaos analysis can take place without this knowledge. A related possibility is to look at transitional processes. Tools are found in bifurcation and catastrophe theory.

Self-organising complex systems will undergo qualitative changes in the dynamics caused by continuous change in the parameters or control variables. These qualitative changes are classified by catastrophe theory in a small number of elementary types. Of special interest are the fingerprints of these qualitative changes (or phase transitions) that are derived from catastrophe theory (Gilmore, 1981). The most obvious characteristic is that such a change is discontinuous (the sudden jump). Other fingerprints concern the variance of the state variables, effects of perturbations, multi-modality and typical consequences of change in the independent parameters or control variables (Van der Maas & Molenaar, 1992). An example of the latter is hysteresis.

Hysteresis means that sudden jumps occur at different values of the independent variable depending on the direction of change in this independent variable. An example is the freezing and melting of water and ice. Temperature is the independent variable. If we increase the temperature ice will melt at 0 °C. If we decrease temperature in a perturbation free setting it will freeze suddenly at -4 °C. If perturbations or noise are present the locations of the sudden jump will converge at 0 °C. This typical property is unique for phase transitions. Demonstration of hysteresis requires advanced experimentation since noise should be excluded.

By these fingerprints we can recognize self-organising phase transitions even if we have very limited knowledge of the process involved. Such a process is the acquisition of cognitive rules. A famous example is the acquisition of the conservation rule as defined by Piaget. When young children (3 to 6 years) are faced with pouring of liquid from a high, small glass into a much wider glass, they think that the amount of liquid is decreased. Somewhere between 6 and 8 years they suddenly learn that the pouring does not matter at all. According to Piaget this insight reflects a major re-organisation of the knowledge system. To investigate whether this insight is really discontinuous and reflecting self-organisation, we can search for the fingerprints in empirical data.

In Van der Maas & Molenaar (1992) it has been shown that a number of these characteristics, sudden jump, bimodality and anomalous variance, indeed occur. We placed computers in classrooms on which children repeatedly made a conservation test. After two measurement sessions measurement could take place unsupervised, decreasing the costs of longitudinal experiments considerably. Testing continued for eight months resulting in 11 sessions. A quarter of the sample showed a clear jump. Moreover, scores on the test were bimodal on each session. Furthermore, evidence was found for unstable transitional strategies involving over-generalization.
In a separate experiment we investigated hysteresis. As independent variable we choose perceptual misleadingness (van der Maas & Molenaar, 1992). The test concerned conservation of weight. First two clay balls were shown. One of these balls was rolled in a sausage of 10 centimetres. The child had to judge whether the weights are still equal or that one has more weight. Then the sausage was rolled to 20 centimetres and child had to judge again. This procedure was repeated until the sausage reached 80 centimetres. Then the length of sausage decreased in steps until the original ball was recovered. It was expected that transitional children switch between correct and incorrect responses and that these switches show hysteresis. However only a few subjects showed the expected pattern. More experiments are required.

One important implication of this research is that the concept of attractor in non-linear dynamic system theory can be interpreted in terms of cognitive strategies. Strategies, in many domains of developmental psychology, are a basic unit of analysis. By dynamic system theory we can model and investigate the dynamic properties of these higher order cognitive organizations.

4. Conclusion

All these examples of non-linear change illustrate the large influence of chance and uncertainty in psychological development. If the butterfly in Brazil can change our weather then all kinds of insignificant accidental events in our life might change the life course dramatically. On the other hand, and this is often forgotten, non-linear dynamic system theory is also about converging routes. Non-linear system theory allows for two opposite possibilities 1) a principal lack of predictability even if all relevant variables are measured in a very reliable way 2) a large predictability in terms of single attractor points even when nothing is known about the values of initial conditions of the relevant variables. The first possibility is often associated with individual differences, whereas the second is associated with universal properties. Differences in verbal ability are difficult to predict but the presence of verbal ability is almost certain, whereas, in both cases, we do not know much of the underlying processes.

By non-linear system theory we understand these possibilities much better although the practical application of the new concepts and techniques has only just started. We can not predict whether this will be successful.
References


Table 1: Different kinds of behaviour simulated by the Verhulst equation

<table>
<thead>
<tr>
<th>r</th>
<th>L_0</th>
<th>L_1</th>
<th>L_2</th>
<th>L_3</th>
<th>L_4</th>
<th>L_5</th>
<th>L_6</th>
<th>L_7</th>
<th>L_8</th>
<th>L_9</th>
<th>L_{10}</th>
<th>L_{11}</th>
<th>L_{12}</th>
<th>L_{13}</th>
<th>L_{14}</th>
<th>L_{15}</th>
<th>L_{16}</th>
<th>L_{17}</th>
<th>L_{18}</th>
<th>L_{19}</th>
<th>L_{20}</th>
</tr>
</thead>
<tbody>
<tr>
<td>.5</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.10001</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td>.100000</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.000000</td>
<td>.200000</td>
<td>.250000</td>
<td>.300000</td>
<td>.350000</td>
<td>.400000</td>
<td>.450000</td>
<td>.500000</td>
<td>.550000</td>
<td>.600000</td>
<td>.650000</td>
<td>.700000</td>
<td>.750000</td>
<td>.800000</td>
<td>.850000</td>
<td>.900000</td>
<td>.950000</td>
<td>.100000</td>
<td>.105000</td>
<td>.110000</td>
<td></td>
</tr>
</tbody>
</table>

Verify that for $r = .5$, $L$ declines to zero, for $r = 2$, $L$ increases to a stable level, for $r = 2.5$, $L$ approaches its stable level oscillatory, for $r = 3.25$, $L$ switches between two stable points, and for $r = 4$, $L$ never converges. The last column has $r = 4$ too, but the initial level is slightly different, .10001. By comparing the last two columns we can see that this small difference accumulates to total divergence for the last iterations.
Figure captions

Figure 1: The autoassociator

Three neurons are shown coupled by excitatory (+) and inhibitory (-) connections. Input patterns enter from the left side. The activity of each of the neurons is a function of this input and the input from the other neurons. The connections are changed according to a learning rule. Verify that if only the upper neuron receives input (-1) the other neurons will still give the correct output.

Figure 2: The neural oscillator

The behaviour of a model consisting of 100 connected neural oscillators is shown. The x-axis denotes the number of the oscillator in the row. The dashed line represents the frequencies of the input, which is directly related to the input strength. In figure A the frequencies of the left bar are 1.1 ± 0.1 (mean value plus a noise term). The frequencies of the right spot are 1.0 ± 0.1. The background frequency is 0.1 ± 0.1. In figure B the input frequencies are 1.0 ± 0.2 inside the spot. In both figures, the solid line represents the phase of the oscillations of each neural oscillator.
+ , -
connections