Catastrophe analysis of discontinuous development

van der Maas, H.L.J.; Molenaar, P.C.M.

Published in:
Categorical variables in developmental research methods of analysis

Catastrophe analysis of discontinuous development

Han L.J. van der Maas & Peter C.M. Molenaar

Department of Psychology

University of Amsterdam

Roetersstraat 15

1018 WB Amsterdam

The Netherlands

This research was supported by the Dutch organization for scientific research (NWO) and the department of Psychology of the University of Amsterdam (UVA).
Correspondence should be addressed to Han van der Maas, Developmental Psychology, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands.
E-mail : op_maas@macmail.psy.uva.nl

Table of Contents

I. Introduction

The theme of the conference and this book is categorical data-analysis. Advanced statistical tools are presented for the study of discrete things behaving discretely. Many relevant social science questions can be captured by these tools. There is tension between categorical and non-categorical data-analysis on basis of ordinal, interval or ratio scaled data.

Zeeman (1993) presents an original view on this conflict (see table 1).

Insert table 1 about here

He distinguishes 4 types of applied mathematics according as to whether things are discrete or continuous, and whether their behavior is discrete or continuous. Of special interest is Pandora's box. In contrast to the other three types, discrete behavior of continuous things give often rise to controversies. The mathematical modeling of music, the harmonic of vibrating strings, led to a major debate in the eighteenth century. The foundations of quantum theory are still controversial. The last example is catastrophe theory which is concerned with the modelling of discontinuities. In developmental psychology we are aware of the controversial nature of discontinuities in for instance cognitive development.

Note that categorical behavior is not explicitly included in Zeeman's boxes. Categorical data, nominal measurements, should be labeled as discrete whereas measurements on a ordinal, interval or ratio level are covered by continuous behavior. Probably, the critical distinction here is in quantitative and qualitative differences in responding. If this distinction is crucial then the definition of discreteness in catastrophe theory includes categorical data.

We will not further discuss this classification of Zeeman. We hope that it illustrates the relation between categorical data-analyses and the contents of this paper. The goal of this paper is to present our results obtained by the application of catastrophe theory to the study of conservation acquisition. Conservation acquisition is a kind of bench-mark problem of the study of discontinuous cognitive development. Discontinuous development is strongly associated with the stage theory of Piaget (1960, 1969, 1971). The complexity of debate on this theory is enormous (the Pandora box), and in our presentation we will necessarily neglect a large body of literature on alternative models, other relevant data and severe criticism on Piaget's work. It is common knowledge that Piaget's theory, especially concerning stages and discontinuities, has lost its dominant position in developmental psychology. Yet, we will defend that our results generally confirm Piaget's ideas on the transition of nonconservation to conservation.

We start with short introduction to catastrophe theory, issues in conservation research, and alternative models of conservation acquisition. We proceed by an explanation of our so-called cusp model of conservation acquisition, and the possibilities to test this model. The last part of this chapter discusses the experiments that we performed to collect evidence for the model. Herein, we first explain our test of conservation, a computer variant of the traditional test. Finally, data of a cross-sectional and a longitudinal experiment are discussed in relation to 3 phenomena of catastrophe models: sudden jumps, anomalous variance, and hysteresis.

II. Catastrophe theory

There are several good introductions to catastrophe theory. In a sequence of increasing difficulty we can recommend Zeeman (1976), Saunders (1980), Poston & Stewart (1978), Gilmore (1981) and Castrigiano & Hayes (1993).

Catastrophe theory is concerned with the classification of equilibrium behavior of systems in the neighborhood of singularities of different degrees. Singularities are points where besides the first derivative higher order derivatives of the potential function are zero. The mathematical basis of catastrophe theory consist of a prove that the dynamics of systems in such singular points can be locally modelled by 7 elementary catastrophes (for systems up to 4 independent variables). The elementary behavior of systems in singular points depends only on the number of independent variables. In case of two independent variables systems characterized by singular behavior can be transformed (by a well defined set of transformations) to the so-called cusp form. Our preliminary model of conservation acquisition is formulated as a cusp model.

In contrast to the well-known quadratic minima, the equilibrium behavior of singular systems is characterized by strange behavior, like sudden jumps and splitting of equilibria. The quadratic minimum is assumed in the application of linear and non-linear regression models. In each point of the regression line a normal distribution of observed scores is expected. In contrast, singular minima lead to bimodal or multimodal distributions. If the system moves between modes, this change is called a qualitative change (of a quantitative variable).

The comparison with regression analysis is helpful. As in regression models catastrophe theory distinguishes between dependent (behavioral) and independent variables (control variables). These variables are related by a deterministic formula which can be adjusted for statistical analysis (see below). The cusp catastrophe is denoted by:

$$V(X;a,b) = 1/4\, X^4 - 1/2\, a\, X^2 - b\, X \qquad (1)$$

which has as equilibria (first derivative to zero):

$$X^3 - a\, X - b = 0 \qquad (2)$$

If we compare the latter function to regression models in the general form of

$$X = f(a,b) \qquad (3)$$

we can see an important difference. The cusp function (2) is an implicit function, a cubic, that can not be formulated in de general regression form.

This difference has far-reaching consequences. In the cusp sudden jumps occur under small continuous variation of the independent variables a and b. In contrast, in the regression models, either linear or non-linear, small continuous variation of independent variables may lead to an acceleration in X but not to genuine discontinuities.

Of course f, in the regression model, can be explicitly discontinuously, a threshold function, but then an purely descriptive position is taken. In catastrophe functions the equilibrium surfaces are continuous (smooth), hence discontinuities are not build in.

In catastrophe theory qualitative changes in quantitative behavior are modelled in a way clearly distinguished from not only (non-linear) regression models but also from Markov chains models (Brainerd, 1979) and special formulations of Rasch models (Wilson, 1989).

III. Issues with Conservation

Conservation is the understanding of the invariance of certain properties in spite of transformation of form. An example of a conservation of liquid task is shown in figure 1.

Insert figure 1 about here

The liquid of one of two identical glasses (B) is poured in a third, non-identical glass (C). The child should judge whether the quantity (C) is now more, equal or less than (A). A conserver understands that the quantities remain equal, a nonconserver systematically chooses the glass with the highest level as having more. A very typical property of conservation tests is that the lowest test score is systematically below chance score.

Piaget's and co-workers took conservation repeatedly as example to explain the general equilibration theory. Piaget's idea of epigenetic development implies that children actively construct new knowledge. In conservation development a number of events can be differentiated.

a) focusing on the height of the liquid columns only (nonconservation)

b) focusing on the width of the liquid columns only

c) the conflict between the height and the width cue

d) the conflict between the height cue and the conserving transformation

e) constructing new operations f) and g)

f) focusing on the conserving transformation

g) understanding of compensation (multiplication of height and width)

In Piaget's model the sequence of events is probably, a, b, c, e, f & g. In the Piagetian view f and g are connected. Whether d should be included or that d is connected to c, is unclear. Many authors give alternative sequences and deny the active construction of the new operations. Bruner (1967), for instance, argues that f and g already exist in the nonconservation period but are hampered by the perceptual misleading factor. Peill (1975) gives an excellent overview of the various proposals about the sequence of events a to g.

Thousands of studies have been performed to uncover the secrets of conservation development. Many models have been introduced which have been tested in training settings and sequence and concurrence designs by test criteria varying from generalization to habituation. Mostly, consensus has not been reached. To our opinion, conservation (or nonconservation) can not be reduced to a perceptual, linguistic or social artefact of the conservation test procedure, although perceptual, linguistic and social factors do play a role, especially in the transitional phase. In a later section we will explain how we developed a computer test of conservation to overcome the major limitations of clinical conservation testing.

A very important approach to the study of conservation focuses on conservation strategies (also rules and operations). At present the best methodology of assessing conservation and several related abilities has been introduced by Siegler (1981; 1986). He uses several item types to distinguish between 4 rules. These rules roughly coincide with Piaget's phases in conservation development. Rule 1 users focus on the dominant dimension (height). Rule 2 users follow rule 1 except if values of the dominant dimensions are equal, then they base their judgment on the subordinate dimension (width). Rule 3 users do as rule 2 users except if both the values of the dominant and subordinate dimensions differ, then they guess. Rule 4 users multiply the dimensional values (in case of the balance beam) and/or understand that the transformation does not effect the quantity of liquid. Multiplying dimensions (or compensation) is correct on conservation tasks too, but more difficult and rarely used (Siegler, 1981).

Siegler's methodology has been criticized mainly for rule 3. Guessing (muddling through) is a rather general strategy of another order. Undoubtedly, it can occur, and should be controlled for. Several alternative strategies have been proposed, for instance addition of dimensions (Anderson & Cuneo, 1978), maximization (Kerkman & Wright, 1988), and the buggy rule (van

Maanen, Been & Sijtsma, 1989). The detection of these alternatives will take a lot of test items but is not impossible.

Another important concept in the study of conservation is that of the cognitive conflict. Piaget himself associated the concept of cognitive conflict to conservation. Most authors agree that conservation performance on Piagetian conservation tasks should be explained as (or can be described as) a conflict between perceptual and cognitive factors. Concepts as salience, field-dependency and perceptual misleadingness have been used to define the perceptual factor in conservation (Pascual-Leone, 1989; Bruner, 1967; Odom, 1972). Siegler applies it in the dominant / subordinate distinction. We will use this factor as an independent variable in our model.

The cognitive factor is less clear understood. It can be defined as cognitive capacity, short term memory or as a more specific cognitive process like learning about quantities. Also the very general concept of maturation can be used. In cusp models one has to choose two independent variables. In the case of conservation learning this very important choice is difficult, one of the reasons to communicate our cusp model as preliminary.

A. Alternative models

Flavell and Wohlwill (1969) took Piaget's view on the acquisition of conservation as a starting point. Piaget distinguished between four phases, nonconservation, pre-transitional, transitional and conservation.

Flavell and Wohlwill (1969) formulated these Piagetian phases in the form of a simple equation.

$$P(+) = Pa * Pb^{1-k} \qquad (4)$$

where P(+) is the probability that a child given person parameters Pa and k and item parameter Pb succeeds on a particular task. Pa is the probability that the relevant operation is functional in a given child. Pb is a difficulty coefficient between 0 and 1. The parameter k is the weight to be attached to the Pb factor in a given child. In the first phase Pa and k are supposed to be zero, and consequently P(+) is zero. In the second phase Pa changes from 0 to 1, whereas k remains zero. For a task of intermediate difficulty (Pb = .5) P(+) = .25. According to Piaget, and Flavell & Wohlwill add this to their model, the child should manifest oscillations and intermediary forms of reasoning. Notice that this does not follows from the equation. In the third phase, a period of stabilization and consolidation, Pa is 1 and k increases. In the fourth phase Pa and k are 1 and full conservation has been reached.

Of course this model has its disadvantages. It predicts a kind of growth curve that is typical for many more models and it predicts little besides what it assumes. In addition, the empirical verification suffers from the lack of operationalization of the parameters. As Flavell and Wohlwill admit, task analysis will not easily lead to an estimation of Pb, Pa is a non-measurable competence factor and k has not a clear psychological meaning and can not be assessed either.

Yet, the reader will notice that this 25 year old model has a close relation to modern latent trait models. The simplest one, the Rasch model, can be denoted as:

$$P(+) = 1 / (1+ e^{-(Pa-Pb)}) \qquad (5)$$

Pa is the latent trait and Pb is the item difficulty. Several modifications are known. Parameters for discrimination, guessing etc. can be added. Pa and Pb have in both models a similar meaning. In latent trait models one can not find a direct equivalent for k, though some latent trait models contain person specific weights for item parameters. If we compare (4) and (5) we can conclude that (4) is a kind of an early latent trait model.

The theory of latent trait models has developed to a dominant approach in current psychometrics. Much is known of the statistical properties of these models. The most appealing advantage is that item parameters can be estimated independent of the sample of subjects and person parameters independent of the sample of items. Besides assumptions as local independence, some constraints on the item specific curves are required. For instance, in the Rasch model item curves have equal discriminatory power. More complex latent trait models do not demand this but other constraints remain. The important constraint for the present discussion is that in all latent trait models, for fixed values of Pa and Pb, one and only one value of P(+) is predicted. The item characteristic curve suggested by catastrophe theory does not accommodate this last constraint.

Hence, we can conclude that the behavior in Flavell & Wohlwill model, like more advanced latent trait models (for instance Saltus, proposed by Wilson, 1989), varies not discontinuously. The behavior may show a continuous acceleration in the increase of P(+), but is not saltatory in a strict sense. A second important conclusion is that the model of Flavell and Wohlwill, contrary to their suggestion, does not predict the oscillations and intermediary forms of reasoning in the transitional phases. We will see that the cusp model explicitly predicts these phenomena.

This latter point is also of importance with regard to Markov models of Brainerd (1979) and the transition model of Pascual-Leone (1970). In Molenaar (1986) is explained why such models do not test discontinuities.

Of course many more models on conservation are proposed, often in a non-mathematical fashion. But as far as they concern the discontinuity explicitly it seems that they will be in accordance with the catastrophe interpretation or not, i.e. belonging to class of models discussed in this section.

IV. The cusp model

In this section we will explain the cusp model in general and its specification as a model of conservation. The type of cusp model we apply, typical by a rotation of axes, is also found in Zeeman (1976). It is especially useful if the dynamics are understood in terms of a conflict. In our application of the cusp model we use perceptual misleadingness and cognitive capacity as independent variables and conservation as dependent variable. The model is represented in figure 2.

Insert figure 2 about here

A three dimensional surface consisting of the solutions of equation (2) represents the behavior for different values of the independent variables. The independent variables define the control plane beneath the surface. The bifurcation set consists of independent values for which three modes of behavior exists. Two of the three modes are stable, the one in the middle is unstable. Outside the bifurcation set only one mode exists. Changes in independent variables lead to continuous changes in the behavior variable except for places on the edges of the bifurcation set. There, where the upper or lower mode of behavior disappears, a sudden jump takes place to the other stable mode.

In terms of conservation we interpret this as that an increase in cognitive capacity, or a decrease in perceptual misleadingness leads to an increase of conservation, in most cases continuously, and sometimes saltatory. The reverse process is also possible but usually not the dominant direction in development. However, artificial manipulation of cognitive capacity (by shared attention tasks) or misleadingness (by stimulus manipulation) should lead to saltatory regressions.

This cusp model suggest the discrimination of four groups of children. At the neutral point, the back of the surface, the probability of a correct response equals chance level. Both independent values have low values. At the left front the scores are below chance level. Perceptual misleadingness is high, capacity is low. In middle both independent values are high. Two behavior modes are possible, above and below chance level. Finally, at the right front the high scores occur. Here perceptual misleadingness is low and capacity is high. Associated with these states are four groups which we call residual group (R), nonconservers (NC), transitional group (TR) and conservers (C), respectively. The residual group consist of children who guess or follow irrelevant strategies because of lack of understanding of test instructions or lack of interest. This group is normally not incorporated in models of conservation development. In this model it is expected that this group is characterized by low values for both independent variables. The nonconserver and conserver children are included in each model of conservation development. Most models also include a transitional group, defined, however, according to all kinds of criteria. In the cusp model the transition period is defined as the period that the subjects stay in the bifurcation set. In this period the possibility of the sudden jump is present. The sudden jump itself is part of this transition phase. The transition includes more than the sudden jump only. There are at least seven other behavioral phenomena typical for the transition period (see below).

A. Routes through control plane

The classification of four groups suggest an expected sequence but this does not directly follows from the cusp model. We need an additional assumption on the direction of change of independent variables as function of age (or as function of experimental manipulation).

To get at a sequence of residual, nonconservation, transitional, and conservation it should be assumed that first, both factors are low, second, perceptual misleadingness increases, third, cognitive capacity increases too, and fourth perceptual misleadingness decreases back to a low level. We discuss this assumption because from literature we know that very young children indeed score at chance level, four to six year olds score below chance level, and later on, they score above chance level (McShane & Morrison, 1983). The phases of the model of Flavell and Wohlwill (1969) coincide with the nonconservation, transitional, and conservation sequence. The phase in which very young children score at chance level is not included in their model nor in the model of Piaget. This is partly due to issues of measurement. Guessing is not helpful on the classical Piagetian test of conservation in which valid verbal arguments are required. Other test criteria (judgment-only criterion or measures of looking time and surprise) allow for false positive classifications whereas Piaget's criterion suffers from false negatives. In the cusp model we assume the use of the judgment-only criterion of conservation.

What should be clear now is that some assumption on the change of the independent variables is necessary to specify a sequence of behavior. Of course this is also required in the model of Flavell & Wohlwill (1969).

B. Test of the cusp model: the catastrophe flags

The transitional period, associated with the bifurcation set, is characterized by 8 catastrophe flags, derived from catastrophe theory by Gilmore (1981). Three of these flags can also occur outside the bifurcation set and may predict the transition. Together, they can be used to test the model.

These catastrophe flags are: sudden jump, bimodality, inaccessibility, hysteresis, divergence, anomalous variance, divergence of linear response and critical slowing down. Some are well-known in developmental research others only intuitively or not at all.

The *sudden jump* is the most obvious criterion. However, in the case of developmental research, quite problematic. Empirical verification of this criterion requires a dense time-series design. In spite of the many studies on conservation development a statistical reliable demonstration of the sudden jump in conservation ability is lacking. Below we will present data that demonstrate the sudden jump.

Bimodality is also known in conservation research. In van der Maas & Molenaar (1992) we present a re-analysis of data of Bentler (1970) which demonstrates bimodality clearly. Though this seems to be the most simple criterion, applicable to cross-sectional group data of the behavior variable only, some issues arise. In the prediction of bimodal score distribution we combine two flags, bimodality and inaccessibility.

Inaccessibility is implied by the unstable mode in between the two stable modes. In mixture distribution model (Everitt & Hand, 1981), this inaccessible mode misses. In these models a mixture of two normal (or binomial, etc) distributions is fitted on empirical distributions:

$$F(X) = p\, N(\mu_1,\sigma_1) + (1-p)\, N(\mu_2,\sigma_2) \qquad (6)$$

where N is the normal distribution, p defines the proportions in the modes, μ's and σ's the characteristics of the modes. This mixture model takes 5 parameters, whereas a mixture of binomials takes 3 parameters. The number of components can be varied and tested in a hierarchical procedure.

An alternative formulation is found in the work of Cobb and Zacks (1985). They apply the cusp equation to define distributions of a stochastic cusp model:

$$F(X) = \lambda \, e^{-(Z^4 - aZ^2 - bZ)} \qquad Z = s\,X - l \qquad (7)$$

where a and b (independent values) define the form of the distribution, s and l linearly scale X, and λ is the integration constant. The 4 parameters a, b, s and l are estimated. It is possible to fit unimodal and bimodal distributions by constraints on the parameters a and b, and compare the fits. The method of Cobb has some limitations (see Molenaar, this volume), and is difficult to fit to data (for instance in computing λ). Yet, it takes account of the impact of the inaccessible mode, and the computational problems are solvable.

A comparison of possible forms of (6) and (7) leads to the conclusion that their relationship is rather complex. In (6) distributions can be fitted to data defined by $\mu_1=\mu_2$ and $\sigma1>>\sigma2$. Such distributions are not allowed in (7), the modes must be separated by the inaccessible mode.

Hysteresis is easily demonstrated in simpler physical catastrophic processes, but probably very difficult in psychological experimentation. The degree of the hysteresis effect, the distance between the jumps up and down, depends on the disturbance or noise in the system (leading to the so-called Maxwell condition). Below we will discuss our first attempt to detect hysteresis in conservation.

Divergence has a close relation to what is usually called a bifurcation. In terms of the chosen independent variables it means that if, in case of a residual child, perceptual misleadingness as well as cognitive capacity are increased, the paths split between the upper and lower sheet, i.e. high and low scores. Two children with almost similar start values (both independent variables very low) and following the same path through control plane can show strongly diverging behavior. Again this is not easily found in an experiment. The manipulation of only one independent variable is already difficult. Another choice of independent variables (a motivation factor or optimal conditions factor in another rotation of independent variables, for instance) may lead to an empirical test.

Anomalous variance is a very important flag. Gilmore (1981) proves that the variance of the behavioral variable increases strongly in the neighborhood of the bifurcation set and that drops occur in the correlation structure of various behavioral measures in this period. In developmental literature many kinds of anomalous behavior, from oscillations to rare intermediate strategies, are known. In our view this prediction, used here as a criterion, is what Flavell and Wohlwill call 'manifestations of oscillations and intermediary forms of reasoning'.

Divergence of linear response and *critical slowing down* , the last flags concern the reaction to perturbations. They imply the occurrence of large oscillations and a delayed recovery of equilibrium behavior, respectively. These flags are not studied in our experiments. Reaction time measures are possibly the solution for testing these two flags.

The flags differ importantly in strength. Bimodality and the sudden jump, certainly when they are tested by regression analysis and mixture distributions, are not unique for the cusp model. It is very difficult to differentiate acceleration and cusp models by these flags. Inaccessibility changes this statement somewhat. It is unique, but as indicated before, usually incorporated in the bimodality test.

Anomalous variance is unique since other mathematical models do not derive this prediction from the model itself. In that respect the many instances of oscillations and intermediary forms of reasoning count as evidence for our model. Hysteresis and divergence are certainly unique for catastrophe models. Divergence of linear response and critical slowing down seem to be unique too. Such a rough description of the strength and importance of each flag is only valid in contrasts with predictions of alternative models. In the case of conservation research, bimodality and the sudden jump are thought of as necessary but not sufficient criteria, the others as both necessary and sufficient. The general strategy should be to found as many as flags as possible. The demonstration of the concurrence of all flags forms the most convincing argument for the discontinuity hypothesis.

Yet, the flags constitute an indirect method for testing catastrophe models. However the advantage is that, except for divergence and hysteresis, the flags only pertain to the behavioral variable. We suggested a perceptual and a cognitive factor in our cusp model of conservation acquisition, but the test of the majority of flags does not depend on this choice. These flags do answer the question whether a discontinuity, as restrictively defined by catastrophe theory, occurs or not.

The definitions of hysteresis and divergence include variation along the independent variables. In case of the cusp conservation model we can think of variation of perceptual cues of conservation items or divided attention tasks to force the occurrence of hysteresis and divergence. On the one hand, it is a pity that these appealing flags can not be demonstrated without knowledge of independent variables. On the other hand, by these flags hypotheses about the independent variables can be tested.

C. The statistical fit of cusp models

A direct test of the cusp model of conservation acquisition requires a statistical fit of the cusp model to measurements of the dependent as well as the independent variables. The statistical fit of catastrophe models to empirical data is difficult and a developing area of research.

In the comparison of (2) and (3), the cusp models and (nonlinear) regression models, we did not discuss the statistical fit. It is common practice to fit regression models with one dependent and two independent variables to data and to test whether the fit is sufficient. Fitting data to the cusp model is much more problematic.

There have been two attempts to fit (2) in terms of (3) (Guastello, 1988; Oliva et all. 1987). Alexander (19??) heavily criticize the method of Guastello. The main problem of Guastello's method puts forth the odd characteristic that the fit increases when the measurement error increases. Gemcat by Oliva et al (1987) does not have this problem but needs an additional penalty function for the inaccessible mode.

A different approach is taken by Cobb (Cobb & Zacks, 1985). They developed stochastic catastrophe functions which can be fitted to data. In (7) the basic function is shown. If one fits a cusp model a and b are, for instance linear, functions of observed independent variables. Examples of application of Cobb's method can be found in Ta'eed, Ta'eed & Wright (1988) and Stewart & Peregoy (1983). We are currently modifying and testing the method of Cobb. Simulation studies should reveal the characteristics of parameter estimates and test statistics, and the requirements on the data. At present we did not fit data of measurements of conservation, perceptual misleadingness and cognitive capacity directly to the cusp model of conservation acquisition.

V. Empirical studies

A. Introduction

In the experiments we have concentrated on the catastrophe flags. Evidence for bimodality, inaccessibility and anomalous variance can be found in literature (van der Maas & Molenaar, 1992). This evidence comes from predominantly cross-sectional experiments. Analysis of time-series from longitudinal experiments have been rarely applied.

The focus on cross-sectional research hampers the assessment of the other flags, most importantly the sudden jump itself. We do not know of any study which demonstrates the sudden jump (or sharp continuous acceleration) convincely. This is of course a difficult task, since dense time series are required.

The presented evidence for anomalous variance is also achieved by cross-sectional studies. In such studies conservation scores on some test are collected for a sample of children in the appropriate age range. Children with consistent low and high scores are classified as nonconservers and conservers, respectively. The remaining children are classified as transitional. Then some measure of anomalous variance is applied (rare strategies, inconsistent responding to a second set of items, inconsistencies in verbal/non-verbal behavior). By an analysis of variance (or better by a non-parametric equivalent) it is decided whether the transitional group shows anomalous variance more than the other groups. In comparison to the strong demands, in the form of the catastrophe flags, that we put on the detection of a transition this procedure of detecting transitional subjects is rather imprecise. However, it is a very common procedure in conservation research, also applied to decide whether training exclusively benefits transitional subjects. In the light of the importance of anomalous variance for our verification, we want a more severe test.

For these reasons a longitudinal experiment is required. Special requirements are dense time-series and precise operationalizations for the flags. This appears to be a difficult task when we rely on the clinical test procedure of conservation. Many re-tests in a short period will have a heavy load on our resources, as well as on the the time of children and the continuation of normal school activities. Moreover, this clinical test procedure has been heavily criticized.We choose to construct a new test of conservation. We describe this test and statistical properties of this test elsewhere (van der Maas et al, submitted), only a short summary will be given below.

B. Instrument: a computer test of conservation

1. Items

In the clinical test procedure the child is interviewed by an experimenter who shows the pouring of liquids to the child and asks a verbal explanation for the judgment. Only a few tasks are used and the verbal justification is crucial in scoring the response and detecting the strategy that is applied by the child.

The rule assessment methodology of Siegler (see above) has a different approach. Siegler applies much more items which are designed to detect strategies on basis of the judgements only. Although verbal justifications have an additional value, the level of conservation performance is determined by simple responses to items only.

The computer test is based on 4 out of 6 item types of Siegler's methodology. We call them the guess equality, guess inequality, standard equality and standard inequality item types. The guess equality and guess inequality item types compare to Siegler's dominant and equal items. In the guess item types the dimensions of the glasses are equal. In the equality variant amounts, heights and widths of the liquid are equal before and after the transformation. In the inequality variant one of the glasses has less liquid, is equal in width but differs in height. In these item types the perceptual height cue points to the correct response and should therefore be correctly solved by all children. Consequently, these items can be used to detect children who apply guessing or irrelevant strategies. In view of the criticism of the judgment-only criterion, concerning the possibility of guessing, items like this should be included.

The standard item types, equality and inequality, compare to conflict equal and subordinate items of Siegler. The standard equality item is shown in figure 1. The dimension of the glasses and liquids differ, whereas the amount are equal. In the initial situation the dimensions are clearly equal. Understanding of the conserving transformation is sufficient to understand that the amounts are equal in the final situation, though multiplication of dimensions in the final situations suffice too.

The standard inequality item starts with an initial situation in which widths are equal but heights, and amounts, differ. In the final situation the liquid of one of the glasses is poured in a glass of such a width that the heights of the liquid columns are then exactly equal (see figure 3).

Insert figure 3 about here

2. Strategies

For all these item types a large variation in dimensions is allowed. In, for example the item in figure 3, the differences in width can be made more or less salience. We expect that these variations do not alter the classification of children. Siegler makes the same assumption. A close examination of results of Ferretti & Butterfield (1986) show that only very large differences in dimensional values have a positive effect on classification.

Siegler uses 6 item types (each four items) to distinguish between 4 rules. We use instead the two standard item types (figures 1 and 3) and classify according to the strategy classification schema (table 2):

Insert table 2 about here

The conservation items have three answer alternatives: Left more, equal, and right more. These three alternatives can be interpreted as responses 'highest more', 'equal' and 'widest more' in case of equality items, and in responses 'equal', 'smallest more', 'widest more' in case of inequality items. The combination of responses on equality and inequality items is interpreted in terms of a strategy. Children that consistently chose 'highest more' on equality items and 'equal' on inequality items follow the nonconserver height rule (NC.h) or, in Siegler's terms, rule 1. Conservers (C) or rule 4 users chose the correct response on both equality and inequality items. Some children will fail the equality items but succeed on the inequality items ('widest more'). If they prefer the 'highest more' response on the equality items they are called NC.p or rule 2 users in Siegler's terms. If they chose the 'widest more' response, they probably focus on width instead of height. The dominant and subordinate dimension are exchanged (in other words, from a to b in the list of events in conservation development). We call this the nonconserver width rule (NC.w).

The nonconserver equal rule (NC.=) refers to the possibility that children prefer the 'equal' response to all items, including the inequality items. We preliminary interpret this rule as an over-generalization strategy. Children, discovering that the amounts are equal on the equality items, may generalize this to inequality items. If this interpretation is correct, NC.= may be a typical transitional strategy. The cell denoted by i1 has not a clear interpretation. Why should a child focus on width in case of equality items, and on the height in case of inequality items. This combination makes no sense to us.

The same is true for the second row, determined by the 'smallest more' response on inequality items. This response should not occur at all, so this row is expected to be empty. The results of the experiments will show how much children apply the strategies related to the nine cells. If a response pattern has an equal distance to two cells, then it is classified as a tie. The number of ties should be small. Notice that ties can only occur when more than one equality and more than one inequality item is applied. The possibility of ties implies a possibility of a formal test of this strategy classification procedure.

The NC.=, NC.w and the i1 rules do not occur in Siegler classification of rules, although they could be directly assessed in Siegler's methodology. Rule 3 of Siegler, muddle through or guessing on conflict items, does not occur in this classification schema. Proposed alternatives for rule 3, like addition or maximization, are neither included. Additional item types are required to detect these strategies.

3. Test

Items are shown on a computer screen. Three marked keys on the keyboard are used for responding. They represent this side more (i.e. left), equal, that side more (right). First the initial situation appears, two glasses filled with liquid are shown. The subject has to judge this initial situation first. Then an arrow denotes the pouring of liquid into a third glass which appears on the screen at the same time as the arrow. Then the liquid disappears from the second glass and appears in the third. Small arrows below the glasses point to the glasses that should be compared (A and C). In the practice phase of the computer test the pouring is, except by the arrow, also indicated by the sound of pouring liquid. This schematic presentation of conservation items is easily understood by the subjects.

The application of the strategy classification schema is more reliable when more items are used. We choose to apply 1 guess equality, 1 guess inequality, 3 standard equality, 3 standard inequality items in the longitudinal experiment and 2 guess equality, 2 guess inequality, 4 standard equality, 4 standard inequality items in the cross-sectional studies.

Guessers are detected by responses to guess items and the initial situations of the standard items, which are all non-misleading situations. Two types of data are obtained, number correct and strategies.

Apart from the test items discussed above, 4 other items are applied. We call these items compensation construction items. On the computer screen two glasses are shown, one filled and one empty. The child is asked to fill the empty glass until the amounts of liquids are equal. The subject can both increase and decrease the amount until it is satisfied with the result. Notice that a conserving transformation does not take place, hence the correct amount can only be achieved by compensation. We will only refer briefly to this additional part of the computer test. It is, however, interesting from a methodological point of view since the scoring of these items is not discretely but continuously, i.e. on a ratio scale.

C. Experiment 1: reliability and validity

The first experiment concerns a cross-sectional study in which 94 subjects in the age range 6 to 11 years old participated. We administered both the clinical conservation test (Goldschmid & Bentler, 1968) and an extended computer test.

The reliability of the conservation items of the computer test turned out to be .88. Four of 94 subjects did not pass the guess criteria, they failed the clinical conservation test too. This clinical test consisted of a standard equality and a standard inequality item, scored according to three criteria: Piagetian, Goldschmid & Bentler, and Judgment only. The classifications obtained by these three criteria correlate above .96, hence to ease presentations we only apply the Piagetian criterion.

The correlation between the numbers correct of the clinical test and the computer test is .75. In terms of classifications 79 of 94 subjects are classified concordantly in nonconservers and conservers. A more concise summary is given in table 3.

Insert table 3 about here

Here we can see that the classifications in NC.h and C do not differ importantly. For the NC.p, NC.= and NC.b strategies no definite statements can be made because of the small number of subjects. It seems that the ties appear among conservers and nonconservers on the clinical test. The percentages correct on both tests show that the tests do not differ in difficulty (62 % versus 59 %, paired t-test: $t = .98$, $df = 93$, $p = .33$).

We can state that the computer test appears to be reliable and valid (if the clinical test is taken as criterion).

D. Experiment 2: Longitudinal investigation of the sudden jump and anomalous variance

In the second study 101 subjects of 4 classes of one school participated. At the beginning of the experiment the ages varied between 6.2 and 10.6 years old. The four classes are parallel classes containing children of age groups 6, 7, and 8.

We placed a computer in each class and trained the children in using it individually. During 7 months 11 sessions took place. Except for the first session, children made the test by themselves as part of normal individual education. This method of closely following subjects looks like what Siegler & Jenkins (1989) call the microgenetic method, except for the verbal statements on responses which they use as additional information. We can speak here of a computerized microgenetic method. Its main advantage is that it takes less effort, since, ideally, the investigator has only to backup computer diskettes.

1. Sudden Jump

In order to find evidence for the sudden jump we classified subjects in 4 groups, nonconservers, transitional, conservers and residual on basis of the strategies applied (see below for details). The transitional subjects are subjects who applied both conserver and nonconserver strategies during the experiment. Twenty four of the 101 subjects show a sharp increase in the use of conserver strategies. We corrected the time-series for latency of transition points. The resulting individual plots are shown in figure 4.

Insert figure 4 about here

This figure demonstrates a very sharp increase of conservation score. To judge how sharp we applied a multiple regression analysis wherein conservation score serves as dependent and session (linear) and a binary template of a jump as independent variables. This latter jump indicator consists of zeros for session 1 to 10 and ones for session 11 to 19.

Together the independent variables explain 88 % of the variance ($F(2,183) = 666.6$, $p = 0.0001$). The t-values associated with the beta coefficient of each independent variable is t =

1.565 ($p = .12$) and $t = 20.25$ ($p = .0001$), for the linear and jump indicator, respectively. Actually, the jump indicator explains more variance than a sixth order polynomial of the session variable.

Yet, this statistical result does not prove that this sharp increase is catastrophic. The data can also be explained by a continuous acceleration model, since the density of sessions over time may be insufficient. What this plot does prove is that this large increase in conservation level takes place within 3 weeks, between two sessions, for 24 of 31 potentially transitional subjects. For the remaining 70 subjects, classified as nonconservers, conservers and residual subjects on basis of strategy use, no significant increases in scores are found. The mean scores on session 1 and 11 for these subjects do not differ significantly, $F(1,114) = .008$, $p = .92$. The scores of these subjects stay at a constant level.

The transitional group shows important individual differences. A few subjects show regressions to the nonconserver responses, some apply rare strategies during some sessions, but the majority shows a sharp increase between to sessions. One subject jumped within one test session, showing consistent nonconserver scores on all preceding sessions, and conserver scores on all subsequent test sessions.

2. Anomalous variance

A special problem in the analysis of anomalous variance is that the conservation test consist of dichotomous items. The variance depends of the test score. We looked for 3 solutions for this problem: inconsistency, alternations, and transitional strategies.

The inconsistency measure is achieved by an addition to the computer test, a repetition of four standard items at the end of test. These additional items are not included in the test score. The inconsistency measure is the number of responses that differ from the responses on the first presentation. We do not present the results here. In summary, inconsistencies do occur in the responses of transitional subjects but occur more in the responses of residual subjects.

The other two measures concern the application of strategies. The analysis of alternations did not indicate a transitional characteristic. The analysis of strategies, however, did. To explain this we will now describe the results in terms of strategies. The other measures are discussed in van der Maas, Walma van der Molen, and Molenaar (submitted).

The responses to 6 standard items are analyzed by the strategy classification schema. In table 4 the results are shown.

Insert table 4 about here

This table shows some important things. The number of ties is low, hence in the large majority of patterns the classification schema applies well. The NC.h and the C strategy are dominant (80 %). The other cells are almost empty or concern uncommon strategies. NC.p, rule 2 in Siegler's classification, NC.=, i1 and NC.w make out 14 % of the response patterns. Focusing on the subordinate dimension, NC.weight, is very rare. A statistical test should reveal whether the small number, 6 if guessers are removed, can be ascribed to chance or can not be neglected. Latent class analysis may be of help here.

For the classification of time series, i.e. the response patterns over 11 sessions, we make use of the translation of raw scores in strategies. The classification is rather simple. If NC.h occurs at least once, and C does not occur, the series is classified as **NC**. If C occurs at least once, and NC.h does not occur, the series is classified as **C**. If both C and NC.h occurs in the time series it is classified as **TR**. The remaining subjects and those who apply guess and irrelevant (i2, i3, and i4) strategies on the majority of sessions are classified as **R.** According to these criteria there are 31 **TR**, 42 **NC**, 20 **C** and 8 **R** subjects. Note that this classification of time series does not depend on the use of the uncommon strategies, NC.p, NC.=, i1, and NC.w. Hence it is allowed to look at the frequency of use of these strategies in the four groups of subjects (see table 5).

Insert table 5 about here

NC.p, rule 2 in Siegler's classifications, does not seem to be a transitional characteristic, in most cases of application it is used by the subjects of the nonconserver group. The strategies i1, and NC.w are to rare to justify any conclusion.

In 15 of 27 cases of application of NC.= it is applied by the transitional subjects. The null hypothesis that NC.= is distributed evenly over the four groups is rejected (chi-square = 11.3, df = 3, p = .01). Furthermore, seven transitional subjects apply NC.= just before they start using the C strategy.

This analysis demonstrates much more convincingly than cross-sectional studies that transitional subjects manifest, what has been called, intermediary forms of reasoning. Whether we can ascribe the occurrence of the NC.= strategy to anomalous variance is not entirely clear.

Note that the search for anomalous strategies comes from the fact that traditional variance is not useful in a binomially distributed test score. In this respect the compensation construction items are of interest. These measure conservation (in fact compensation) ability on a continuous ratio scale (amount of liquid, or height liquid level). In this measure a more direct test of anomalous variance may be possible.

E. Experiment 3: Hysteresis

The last flag that we will discuss here is hysteresis. We build the cusp model of conservation acquisition on the assumption that perceptual and cognitive factors conflict each other. Hysteresis should occur under systematic increasing and decreasing variation in these factors. Preece (1980) proposed a cusp model that is a close cousin of ours. Perceptual misleadingness plays a role in his model too.

We will not explain Preece' model here, but it differs from our preliminary model, for instance, with respect to the behavior modes to which the system jumps. The difference can be described in terms of the conservation events (see 'issues with conservation' section). Preece's model concerns event 3 whereas our model concerns event 4, resulting in a choice between the 'height more' and 'weight more', and between the 'height more 'and 'equal' response, respectively.

So Preece expects a hysteresis effect between incorrect responses and we between correct and incorrect responses.

The idea suggested by Preece is to vary the misleading cue in conservation tasks. In a conservation of weight test a ball of clay is rolled to sausages of lengths of 10, 20, 40, 80, 40, 20, and 10 cm, respectively (see table 6). In this test a child shows hysteresis if it changes the judgment twice; ones during the increasing sequence and ones during the decreasing sequence. If both jumps takes place, hysteresis occurs either according to the delay (Hyst D) or Maxwell convention (Hyst M). In the latter case the jumps takes place on the same position, for instance between 20 and 40 cm. If only one jump occurs we will classify the pattern as 'jump'.

This weight test was performed as a clinical conservation test and was administered to 65 children. Forty three subjects correctly solved all items. Two subjects judged consistently that the standard ball was heavier, and 2 judged the sausage as heavier. In table 6 the response patterns of all subjects are shown, the consistent subjects in the lower part. The codes -1, 0 and 1 mean 'ball more', 'equal weight', and 'sausage more'.

The first subject shows a good example of a hysteresis response pattern. When the ball is rolled to a sausage of a length of 10 cm he judges the ball as heavier. Then the sausage of 10 cm is rolled to a length of 20 cm, the child persist in his opinion. But when the sausage of 20 cm is rolled to a sausage of 40 cm he changes his judgment, now the roll is heavier. He continues to think this when the sausage reaches its maximum length of 80 cm. Then the sequence is reversed. The roll of 80 cm is folded to 40 cm, the child judges that the roll is still heavier. This continues until the roll is folded to 10 cm. On this last item the child changes his opinion, the ball is heavier again. The first subject is tested two times. The second time again hysteresis according to the delay convention takes place. However the jumps have exchanged relative position. The jump takes place earlier in the increasing than in the decreasing sequence of items. We denoted this strange phenomena as Hyst D'.

This subject shows hysteresis between incorrect responses (inc/inc), subject 42 between correct and incorrect responses (c/inc). This subject judges the ball and sausage as being equal until the sausage is rolled to a length of 80 cm. Then the ball is heavier. The child persist in this judgment when the sausage is folded to 40 cm. Finally, in the last two items the child responses 'equal' again. Eleven instances of hysteresis and jumps occur among the tested children. The other patterns are not interpretable in this terms.

Insert table 6 about here

The hysteresis results on the conservation of weight test are not conclusive. It would be naive to expect a convincing demonstration of hysteresis in the first attempt. The number of items in the sequence and the size of steps between items, are difficult to chose. Perhaps another kind of test is needed to vary misleadingness continuously. Furthermore, only the transitional subjects in the sample are expected to show these phenomena. In this regard the results in table 6 are promising.

We also constructed a computer test of liquid for hysteresis. Only 2 instances of the jump pattern and 4 of the hysteresis pattern occurred in the responses of 80 subjects (from the longitudinal sample). Three of these subjects came from the transitional group. One of them responded as nonconserver until we applied the hysteresis test (between session 7 and 8) and responded as conserver on all subsequent sessions.

These tests of hysteresis have some relationship to a typical aspect of Piaget's clinical test procedure of conservation. This aspect is called resistance to countersuggestion which is not applied in the computer test of conservation. It means that, after the child has made its judgment and justification of this judgment, the experimenter suggest the opposite judgement and justification to the child. If the child does not resist to this countersuggestion he or she is classified as transitional. That countersuggestion works has been shown in many studies. Whether subjects who accept countersuggestion are transitional (or socially adaptive) is not always clear, as said before, cross-sectional studies have serious limitation in making this decision.

The method of countersuggestion can be interpreted in the cusp model as pushing the behavior in the, new or old, mode of behavior and see whether this mode, still or already, exists and has some stability. If countersuggestion is resisted the second mode apparently does not exist. In the cusp model the second mode only exists in the bifurcation set. Note that we defined the transition as being the period in which the system remains in this set. If the subject switches opinion there and back as a result of repeated countersuggestions, this can be understand as hysteresis. Whether the manipulation of countersuggestion yields a independent variable or not, is less clear.

VI. Discussion

In this chapter we gave an overview of our work on the application of catastrophe theory to the problem of conservation development. Not all ideas and all empirical results are discussed but the most important outcomes are presented.

What can be learned from this investigation. One may think that catastrophe theory is a rather complex and extensive tool for something simple as conservation learning. Some readers may hold the saltatory development of conservation as evidently in itself.

With respect to the first criticism we would state that there are theoretical and empirical reasons for using catastrophe theory. Theoretically, the reason for applying catastrophe theory is that the rather vague concept of epigenese in Piaget's theory can be understood as equivalent to self-organisation in non-linear system theory. Empirically, the lack of criteria for testing saltatory development, implied in the Piagetian theory, has been an important reason for the lost of interest in this important paradigm (van der Maas & Molenaar, 1992).

The second criticism can be based on the argument that our conservation items are dichotomous items, so behavior is necessarily discrete. The criticism is not in accordance with test theory. A set of dichotomous items yield in principle a continuous scale. On other tests where a set of dichotomous items of the same difficulty are used, the sum score is unimodality distributed. In contrast, the sum scores on the conservation test are bimodally distributed. Yet, the use of dichotomous items complicates the interpretation of results. It raises a question about the relationship between catastrophe models and standard scale techniques, especially latent trait theory. This is a subject that needs considerable study. We suggested the use of compensation construction items to radically exclude these complications.

This criticism of intrinsic discreteness is not commonly used by developmental psychologists. The traditional rejection of Piaget's theory assumes the opposite hypothesis of continuous development. With regard to this position the presented data are implicative.

Firstly, the sudden jump in the responses of a quarter of our longitudinal sample has not been established before. It is disputable whether this observation excludes alternative models that predict sharp continuous accelerations. We can argue that at least one subject showed the jump within a test session during a pause of approximately 30 seconds between to parts of the computer test. Hence, we demonstrated an immediate sudden jump for at least one subject.

Secondly, evidence is presented for the occurrence of transitional strategies. Note that we, since this concerns a result in a longitudinal study, are much more certain that the subjects classified as transitional are indeed transitional. Transitional subjects appear to respond 'equal' to both equality and inequality items. This may be interpreted as that they over-generalize their discovery of the solution of standard equality items to the inequality items. Verbal justifications are required to decide whether this interpretation is correct. In our view the demonstration of a typical transitional strategy is quite important to Piaget's idea of knowledge construction, this result is confirmatory. It also makes an argument for using the rule assessment methodology in studying Piagetian concepts. We presented this result under the heading of anomalous variance. We relate the catastrophe flag anomalous variance to, what Flavell and Wohlwill (1969) call, oscillations and intermediary forms of reasoning. We suggest that the findings of Perry, Breckinrigde Church, and Goldin Meadow (1988) concerning inconsistencies between nonverbal gestures and verbal justifications should be understand in terms of this flag, too.

Thirdly, the few instances of hysteresis response patterns suggest a choice for the cusp model of conservation. We admit that more experimentation is required here. The test should be improved and perhaps we need other designs and manipulations. We already suggest to use countersuggestion as manipulation.

Fourthly, evidence for other flags can be found in literature. Especially for bimodality and inaccessibility already strong evidence exists (Bentler, 1970; van der Maas et al., 1994). Divergence, divergence of linear response and critical slowing down are not demonstrated.

Although the presented results make together a rather strong argument for the hypothesis of discontinuous development, we will attempt to find evidence for these flags too.

The presented evidence for the catastrophe flags offers a strong argument for the discontinuity hypothesis concerning conservation development. That is, the data bear evidence for catastrophe models of conservation development. However, we can not claim that this evidence directly applies to the specific cusp model that we presented. The choice of independent variables is not fully justified. The hysteresis experiment gives some evidence for perceptual misleadingness as independent variable, but this variable is, for instance, also used by Preece (1980). On the other hand, there are some theoretical problems for which the model of Preece (1980) can be rejected (see van der Maas & Molenaar, 1992). We did not put forward evidence for cognitive capacity as independent variable. In future studies we will focus on this issue. However, it is in accordance with the views of many supporters and opponents of Piaget that conservation acquisition should be understood in terms of a conflict between perceptual and cognitive factors.

We hope that we made clear that catastrophe approach to discontinuous behavior has fruitful implications. Catastrophe theory concerns qualitative (categorical) behavior of continuous variables. It suggests a complex relation between continuous and categorical variables, which falls outside the scope of standard categorical models and data-analytical methods. Yet, the catastrophe models are not unrelated to standard notions, but the question how catastrophe models should be incorporated in standard techniques is far from answered.

References

Alexander, R.A., Herbert, G.R., DeShon, R.P., & Hanges, P.J. (1992). An examination of Least squares regression modeling of catastrophe theory. Psychological Bulletin, 111(2), 366-374.

Anderson, N.H. & Cuneo, D.O. (1978). The Height + Width Rule in Childrens Judgments of Quantity. Journal of Experimental Psychology: General, 107,4, 335-378.

Bentler, P.M. (1970). Evidence Regarding Stages in the Development of Conservation. Perceptual and Motor skills, 31, 855-859.

Brainerd, C.J. (1979). Markovian Interpretations of Conservation Learning. Psychological Review, 86, 181-213.

Bruner, J.S. (1967). On the Conservation of Liquids. In Bruner, J.S., Olver, R.R., & Greenfield, P.M. (Eds.), Studies in Cognitive Growth (pp. 183-207). New York: John Wiley & Sons, Inc.

Castrigiano, D.P.L. & Hayes, S.A. (1993). Catastrophe Theory. Reading: Addison-Wesley P.C..

Cobb, L., & Zacks, S. (1985). Applications of Catastrophe Theory for Statistical Modeling in the Biosciences. Journal of the American Statistical Association, 80(392), 793-802.

Everitt, B.S., & Hand, D.J. (1981). Finite Mixture Distributions. London: Chapman and Hall.

Ferretti, R.P., & Butterfield, E.C. (1986). Are children's rule assessment classifications invariant across instances of problem types ? Child Development, 57, 1419-1428.

Flavell, J.H., & Wohlwill, J.F. (1969). Formal and Functional Aspects of Cognitive Development. In Elkind, D., & Flavell, J.H. (Eds.), Studies in Cognitive Development: Essays in Honour of Jean Piaget (pp. 67-120). New York: Oxford University Press.

Gilmore, R. (1981). Catastrophe Theory for Scientists and Engineers. New York: John Wiley & Sons.

Goldschmid, M.L. & Bentler, P.M. (1968). Manual conservation assessment kit. San Diego, Educational and Industrial Testing Service.

Guastello, S.J. (1988). Catastrophe Modeling of the Accident Process: Organizational Subunit Size. Psychological Bulletin, 103(2), 246-255.

Kerkman, D.D. & Wright, J.C. (1988). An exegenis of compensation development: Sequential decision theory and Information integration theory. Developmental Review, 8, 323-360.

McShane, J. & Morrison, D.L. (1983). How young children pour equal quantities: a case of pseudoconservation. Journal of Experimental Child Psychology, 35, 21-29.

Molenaar, P.C.M. (1986). Issues with a Rule-Sampling theory of Conservation Learning from a Structuralist Point of View. Human Development, 29, 137-144.

Odom, R.D. (1972). Effects on Perceptual Salience on the Recall of Relevance and Incidental Dimensional Values: A Developmental Study. Journal of Experimental Psychology, 92, 185-291.

Oliva, T.A., Desarbo, W.S., Day, D.L., & Jedidi, K. (1987). GEMCAT: A General Multivariate Methodology for Estimating Catastrophe Models. Behavioral Science, 32, 121-137.

Pascual-Leone, J. (1970). A Mathematical Model for the Transition Rule in Piaget's Development Stages. Acta Psychologica, 32, 301-345.

Pascal-Leone, J. (1989). An Organismic Process Model of Witkin's Field-Dependency-Independency. In Globerson, T., & Zelniker, T. (Eds.), Cognitive Style and Cognitive Development (Vol. 3, pp. 36-70). Norwood: Ablex Publishing Corporation.

Peill, E.J. (1975). Invention and discovery of reality: the acquisition of conservation of amount. London: John Wiley & Sons.

Perry, M., Breckinridge Church, R, & Goldin Meadow, S. (1988). Transitional Knowledge in the Acquisition of Concepts. Cognitive Development, 3, 359-400.

Piaget, J. (1960). The General Problems of the Psychobiological Development of the Child. In Tanner, J.M., & Inhelder, B. (Eds.), Discussions on Child Development (Vol. 4). London: Tavistock.

Piaget, J. (1971). The Theory of Stages in Cognitive Development. In Green, D.R., Ford, M.P., & Flamer, G.B. (Eds.), Measurement and Piaget . New York: McGraw-Hill.

Piaget, J., & Inhelder, B. (1969). The psychology of the Child. New York: Basic Books.

Poston, T., & Stewart, I. (1978). Catastrophe Theory and its Applications. London: Pitman.

Preece, P.F.W. (1980). A Geometrical Model of Piagetian Conservation. Psychological Reports, 46, 143-148.

Saunders, P.T. (1980). An introduction to catastrophe theory. Cambridge: Cambridge University press.

Siegler, R.S. & Jenkins, E. (1989). How children discover new strategies. New Jersey: Erlbaum.

Siegler, R.S. (1981). Developmental sequences within and between concepts. Monographs-of-the-Society-for-Research-in-Child-Development, Vol 46(2), 84.

Siegler, R.S. (1986). Unities across domains in childrens strategy choices. The Minnesota symposia on child psychology, 19, 1-46.

Stewart, I.N., & Peregoy, P.L. (1983). Catastrophe Theory Modeling in Psychology. Psychological Bulletin, 94(2), 336-362.

Ta'eed, L.K., Ta'eed, O., & Wright, J.E. (1988). Determinants involved in the Perception of the Necker Cube: An Application of Catastrophe Theory. Behavioral Science, 33, 97-115.

van Maanen, L., Been, P. & Sijtsma, K. (1989). The linear logistic test model and heterogenity of cognitive strategies. Mathematical Psychology in progress. Ed. Roskam, E.E. Berlin: Springer-Verlag.

van der Maas, H.L.J. van der, & Molenaar, P.C.M.. (1992). Stagewise cognitive development: an application of catastrophe theory. Psychological Review, 99(3), 395-417.

van der Maas, H.L.J., Walma van der Molen, J. & Molenaar P.C.M. Discontinuity in conservation acquisition: A longitudinal experiment. (submitted).

van der Maas, H.L.J., Molenaar P.C.M., van Gils, N., Pope, F.H.K., Geus, W de, & Bast, J.W. Nonverbal assessment of conservation (submitted).

Wilson, M. (1989). Saltus: A Psychometric Model of Discontinuity in Cognitive Development. Psychological Bulletin, 105(2), 276-289.

Zahler, R.S., & Sussmann, H.J. (1977). Claims and Accomplishments of Applied Catastrophe Theory. Nature, 269(10), 759-763.

Zeeman, E.C. (1976). Catastrophe Theory. Scientific American, 234(4), 65-83.

Zeeman, E.C. (1993). Controversy in science, the ideas of Daniel Bernoulli and René Thom. Johann Bernoulli Lecture, University of Groningen.

Table 1: Four types of applied mathematics (Zeeman, 1993).

Behavior	Things	
	discrete	continuous
discrete	Dice Symmetry DISCRETE BOX Finite probability Finite groups	Music, harmony Light Discontinuities PANDORA'S BOX Fourier series Quantum theory Catastrophe theory
continuous	Planets Populations TIME BOX Ordinary differential equations	Waves Elasticity CONTINUOUS BOX Partial differential equations

Table 2: Strategy classification scema.

		standard equality items		
		highest more	equal	widest more
standard inequality items	equal	NC.h / rule 1	NC.=	i1
	smallest more	i2	i3	i4
	widest more	NC.p / rule 2	C / rule 4	NC.w

Table 3: Classifications on the computer test versus the clinical test consisting of an equality and an inequality item. Four score patterns of correct (1) and incorrect (0) can occur on the clinical test. These are compared to the strategies found on computer test (strategies i1 to i4 did not occur at all). The cells contain numbers of subjects.

Scores clinical test		Strategy on computer test							
=	≠	NC.h	NC.p	NC.=	C	NC.b	ties	guess	Total
0	0	19	3	0	1	0	4	3	30
1	0	1	0	0	2	0	0	1	4
0	1	2	2	0	3	0	0	0	7
1	1	2	1	2	40	1	7	0	53

Table 4: For 11 sessions * 101 subjects = 1111 response patterns can be classified. The were 265 missing patterns. Forty three response patterns, ties, could not be classified because of equal distance to the ideal patterns associated with the nine cells. The 801 remaining patterns can be uniquely classified and are distributed as shown in the table. Between brackets the number of guessers is displayed. A response pattern is classified as a guess strategy if 2 guess items are incorrect, or, if 1 guess item and more than 25 % of the responses to the non-misleading initial situations of the standard items are incorrect.

ties 43 (8)		equality items		
missing 265		highest more	equal	widest more
inequality items	equal	NC.h 417 (15)	NC.= 36 (9)	i1 20 (5)
	smallest more	i2 4 (0)	i3 5 (3)	i4 2 (2)
	widest more	NC.p 50 (11)	C 258 (4)	NC.w 11 (5)

Table 5: Application of uncommon strategies by nonconservers, conservers, transitional and residual subjects. Other strategies are C, NC.h, guess, ties and irrelevant strategies (i2 to i4). Raw numbers and percentages of use in each groups are displayed.

	Group			
strategy	**NC**	**C**	**TR**	**R**
NC.p	21 (5.5 %)	4 (2.5 %)	7 (2.8 %)	7 (11.8 %)
NC.w	1 (0.3 %)	0 (0 %)	3 (1.2 %)	2 (3.4 %)
NC.=	5 (1.3%)	4 (2.5 %)	15 (6.0 %)	3 (5.1 %)
i1	5 (1.3 %)	0 (0 %)	7 (2.8 %)	3 (5.1 %)
other	349 (91.6%)	149 (95 %)	217 (87 %)	44 (74.6 %)
total	381 (100 %)	157 (100 %)	249 (100 %)	59 (100 %)

Table 6: Response patterns of 65 children in a cross-sectional experiment concerning hysteresis. Explanation of abbreviations is given in the text.

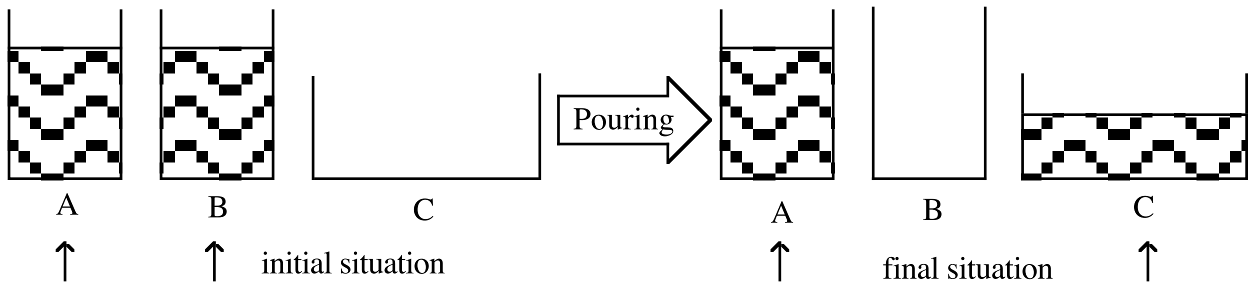
nr	10 cm	20 cm	40 cm	80 cm	40 cm	20 cm	10 cm	pattern	alternation
39	-1	-1	1	1	1	1	-1	hyst D	inc/inc
	-1	1	1	1	-1	-1	-1	hyst D'	inc/inc
63	-1	-1	-1	-1	1	1	1	jump	inc/inc
11	-1	1	1	1	1	1	-1	hyst M	inc/inc
51	0	0	-1	-1	0	0	0	hyst D'	c/inc
42	0	0	0	-1	-1	0	0	hyst D	c/inc
47	0	-1	-1	-1	-1	0	0	hyst D'	c/inc
18	-1	-1	-1	-1	-1	-1	1	jump	c/inc
	0	-1	-1	-1	-1	0	0	hyst D'	c/inc
	0	-1	-1	-1	-1	0	0	hyst D'	c/inc
3	0	0	1	1	1	1	1	jump	c/inc
21	1	-1	-1	1	-1	1	-1		
40	1	-1	-1	1	1	-1	1		
49	0	0	-1	0	0	0	0		
1	1	0	0	0	0	0	0		
62	-1	-1	1	1	0	1	-1		
64	1	1	1	0	-1	-1	-1		
65	1	-1	0	-1	1	0	0		
25	1	-1	1	1	1	1	0		
32	0	-1	-1	-1	0	0	1		
	0	0	0	0	1	0	-1		
27	0	1	-1	0	1	0	1		
	-1	-1	-1	-1	-1	-1	-1		2 subjects
	1	1	1	1	1	1	1		2 subjects
	0	0	0	0	0	0	0		43 subjects

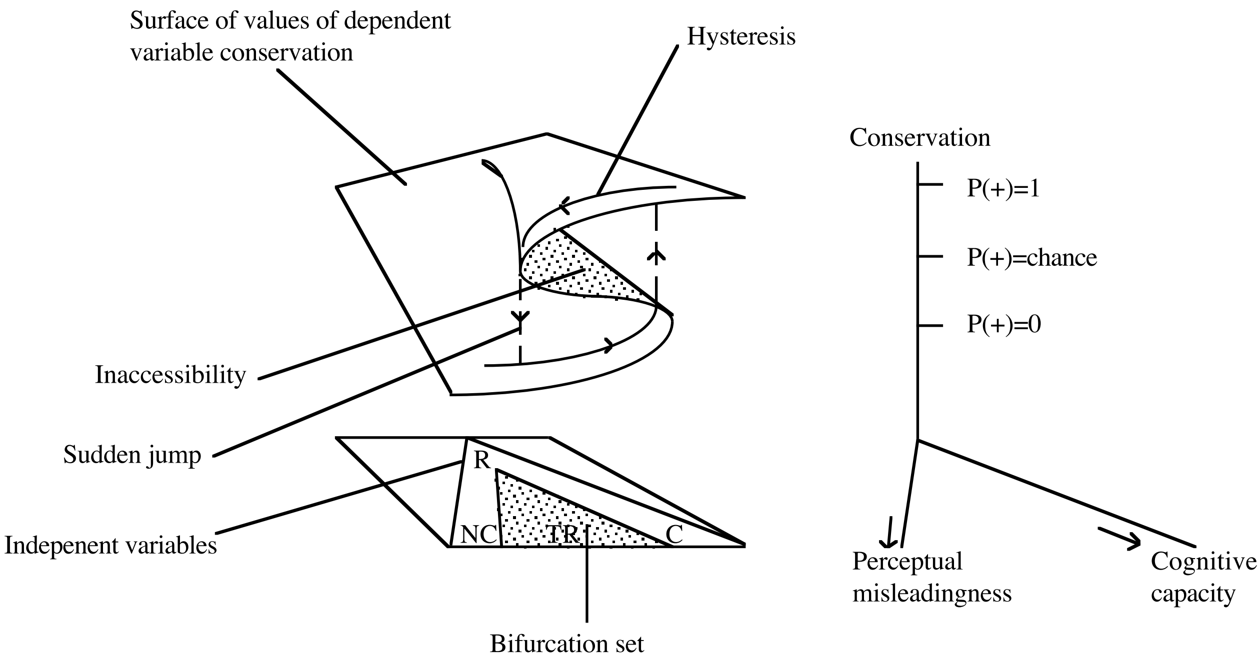
Figure 1: A standard equality item of conservation of liquid. In the initial situation A and B have to be compared, in the final situation, after the transformation, A and C (as indicated by the arrows below the glasses).

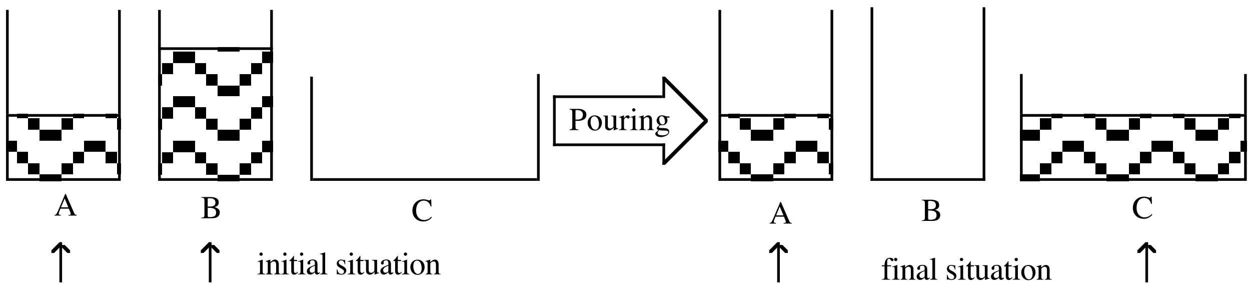
Figure 2: The cusp model of conservation acquisition holds two independent and one behavioral variable, perceptual misleadingness, cognitive capacity and conservation, respectively. The cusp surface is shown above the plane defined by the independent variables. The cusp surface has a folded form at the front that vanishes at the back of the surface. The folding implies that for certain values of the independent variables 3 values of the behavioral variable, i.e. 3 modes, are predicted. The mode in the middle is unstable and compelling and therefore called *inaccessible*. The remaining two modes lead to *bimodality*. The area in the control plane, defined by the independent variables, for which more behavioral modes exist is called the *bifurcation set*. If the system is in its upper mode and moves from right to the left a *sudden jump* will take place to the lower mode at the moment that the upper mode disappears. If the system then moves to the right another jump to the upper mode will occur when leaving the bifurcation set at the right. This change in position of jumps is called *hysteresis*. The placement of groups, R, NC, TR and C is explained in the text. Van der Maas & Molenaar, Stagewise cognitive development: an application of catastrophe theory, Psychological Review, 99(3), 395-417. Copyright 1992 by the American Psychological Association. Adapted by the permission of the publisher.

Figure 3: A standard inequality item of conservation of liquid.

Figure 4: Individual transition plots aligned to time point 11. On the vertical axis raw scores on the conservation test are shown. Two guess items and 6 standard items were used.


Pouring
A
B
C
initial situation
A
B
C
final situation


Surface of values of dependent variable conservation
Hysteresis
Inaccessibility
Sudden jump
Indepenent variables
R
NC
TR
C
Bifurcation set
Conservation
P(+)=1
P(+)=chance
P(+)=0
Perceptual misleadingness
Cognitive capacity


Pouring
A
B
C
initial situation
A
B
C
final situation