A model of the development of the early infant object concept

George F Luger
Department of Computer Science, University of New Mexico, Albuquerque, NM 87131, USA
T R Bower, Jennifer G Wishart
Department of Psychology, University of Edinburgh, Edinburgh, Scotland
Received 7 October 1981, in revised form 19 July 1982

Abstract. A computational model is proposed for the early stages of development of the object concept in infancy. Stages in development are represented as a sequence of grammars or rewrite rules that pass a set of perceptual phenomena. The infant's changes between developmental stages can be described by differences between the grammar rules that model each stage. The program replicates five studies by Bower et al on the development of the object concept and reaffirms the primacy of rest and motion parameters as explanatory invariants in early object-concept development.

1 Introduction
The development of the object concept in infancy has attracted a great deal of psychological interest, particularly in recent years (for reviews see Eilkin and Sameroff 1970; Gratch 1975; Harris 1975, 1983; Schusterman 1982). One reason for this is that the rate of development of the object concept seems to be the only change in infancy that predicts later changes in cognitive functioning (Wachs 1975). Our paper is a beginning attempt at a computer simulation of that development. Some of the psychological data relevant to this attempt are set out in subsequent paragraphs. At this point we would like to clarify the presuppositions of our approach to this and other problems in cognitive development; these presuppositions have shaped not only this paper but also prior and emergent papers (e.g. see Bower 1979, 1983; Wishart and Bower 1983).

The basic problem which concerns us is development, i.e. change in functioning with age. For us this requires the description of the age-related changes in any cognitive change before any description of the nature of the transition between these stages. Descriptions of transition processes, it seems to us, are impossible without precise descriptions of the states at the beginning and the end of these transitions. Although the above may seem a truism, it leads into conceptual problems which psychologists at least rarely face: that is, the status of descriptions of behaviour. As will be seen below, our model is not based on records of specific S-R behaviour (e.g. eyes move left—stop—return 40°) but rather on higher-level descriptions of these. The term ‘level’ is used here in the sense used by Russell (1910) and introduced to psychology by Batten (1972). In our own research on object-concept development we have looked at behaviour and tried to formulate first- or second-level rules to account for that behaviour. We have then attempted to test the validity of these rules by presenting babies with new situations in which, if the baby operated by these rules, behaviour would go one way, if by different rules, in another way (e.g. see Bower and Paterson 1972; Bower 1974).

It should be clear that the rules begin in the mind of the adult experimenter. The adult experimenter is, however, trying to infer the rules that operate in the mind of the baby. If the two match up, we would have a complete theory of at least one segment of development. The computer model is important in that a computer, unlike one's fellow scientists, will not be confused by an ambiguous sentence. To try
Since featural information is incorporated in the rule for identifying an object, any event sequence that violates these new corollaries will now be treated by the infant as the replacement of the original object by another object rather than as a transformation process, as on rule 1. As a result, search behaviour in such situations (eg where an object is covered by a cup) will be directed first by the rule:

To find an object that has disappeared mysteriously, remove the object which has replaced it (which will lead to successful search in the stage III-IV transition task).

And then by the more specific rule:

To find an object that has disappeared mysteriously, remove the object which is now in the place where it was last seen (which will lead to successful search in the stage IV-V transition task).

It can be seen from the above that although a separate search rule is proposed to account for the behaviour seen at each of the first five stages of object-concept development, only two basic conceptual rules are believed to underlie these changes. Development between stages III and V results from elaboration of search behaviour rules rather than from any true advance in object understanding.

This paper describes a computational model of the above set of rules. The computational model will be run in several experimental situations and the results discussed in terms of the notions of motion and place proposed by Bower and Wishart as explanatory invariants of the early developing object concept.

In the next section the computational model will be described. Section 4 will contain the experimental tests of the model. The final section will discuss the results of these experiments and offer an explanation of the differences in development by comparing the rewrite rules that explain the behaviour at each stage. The paper will close with a description of the continuing research.

3 The computational model

3.1 Description

The rules proposed by Bower and Wishart (see the preceding section) can be formulated as a sequence of grammars or rewrite rules that translate perceptual phenomena into sets of behavioural responses. The perceptual phenomena will remain constant throughout the period of testing, i.e. across the running of the sequence of grammars that represents the early stages of development.

The a priori commitment of this model of development should be made clear. Two issues are critical: first, there is no interaction between the percept and the cognizing subject that in any way changes the nature of the percept. Rather the changes come in the cognizing subject's transformation of or interaction with the percept. This facet of the model is meant to capture the Bower-Wishart notion of the primary of motion and rest as explanatory invariants in the early stages of object-concept development.

Second, this belief in the primacy of perception allows description of its origin and presence according to a number of differing theories (Marr 1976, 1978; Ullman 1978; Kokers and von Grunau 1976). This particular program does not parse retinal arrays to detect edges or perform figure-ground separation. (However, it is able to detect boundary violations such as partial occlusion.) The perceptual variables given to the sets of grammars include position, size, colour, and shape. Further, the perception of motion and changes of motion by the calculation of differences in positions over time is an irrelevant implementation detail. That is, like feature extraction, how this is accomplished in the human is an empirical question to be addressed by researchers who are considering these aspects of human response (references as above).

This paper hypothesizes that the symbolic output of the featural and motion detection mechanisms is available to the cognizing subject. As stated above, this model makes no assumption of the physiological origins of such phenomena, but rather emphasizes the descriptive adequacy of the internal symbol structures and the interpretive adequacy of the cognizing subject's manipulation of such symbol structures. Further, the changes in the computational rules that express the interpretive adequacy of infants at various stages in development will offer explanation of that development.

Each experiment of this study will be composed of a sequence of 'snapshots' representing the physical situation according to a frame parameter. Snapshots represent objects by themselves, partially or totally obscured by occluders, and replaced by other objects. An example of such a set of snapshots may be seen in figure 1 (see below). This figure represents a subset of the snapshots taken from experiment 3 of section 4, where r indicates the time parameter.

A set of symbols represents each snapshot and a set of rules characterizes the grammar that interprets each sequence of snapshots. Each rule of a grammar represents a different interpretive capacity of the subject such as to locate an object symbol structure within a fixed radius r of a spatial position (x, y, z). This rule takes position (x, y, z) and radius r and returns an actual object structure at location (ox, oy, oz) such that

\[ r^2 = (x - ox)^2 + (y - oy)^2 + (z - oz)^2. \]

A further rule checks for parallax in an object symbol structure by checking whether the structure 'has' mass or volume. This is accomplished by testing it from two slightly different views. Each rule of a grammar is a procedure (in PROLOG, of next section) for interpreting the sequence of snapshots.

In particular the grammars of this model are designed to implement the rules outlined by Wishart (1979) and presented in section 2. An implicit assumption is that the infant is motivated to maintain contact throughout the event sequence with the object initially identified as interesting (see below).

3.2 Implementation

This computational description of the object-permanence phenomenon is written in PROLOG (Warren and Periera 1977). The action of PROLOG is that of a unification algorithm which operates on a set of record structures. These structures are of two general forms: a set of facts and a set of inference rules. The PROLOG facts are used to make up the object structure for the description of each snapshot. For example the facts

\( \text{color}(	ext{obj}, c), \text{size}(	ext{obj}, s), \text{location}(	ext{obj}, x, y, z) \)

indicates that object 'obj' has colour 'c', size 's', and location (x, y, z) at time 't'. The combination of these and other descriptors make up each snapshot of the experiment. The grammar rules interpret these object structures.

PROLOG rules are of the form 

\( \text{A} \rightarrow \text{B} \), \( \text{C} \), \( \text{D} \)

which may be described procedurally as

\( \text{to accomplish A attempt to accomplish B and C and D} \).

B, C, and D may be facts (checked to be true) or may themselves be rules that lead to the proof or performance of B, C, and D. For example, grammar 1 says to test for a permanent object at a location

(i) look for an object structure within a fixed radius r of the location (described in section 3.1).

(ii) Check whether previous snapshots would indicate that a permanent object should be at this location.

(iii) Test the object structure for interest (parallax, as described above).

(iv) Check whether the object structure is intact (that is whether it or any of its boundaries are occluded). And finally,
In experiment 5, grammar 1 posits a new object for each path of movement and each stationary object position. Grammar 2 sees only one object.

There are several directions for continued research. The first is to develop further the two grammars presented in this paper. With 'disappearing' objects, for example objects hidden by a screen, there is an age x time interaction that is not yet incorporated. It should also be possible to vary an object's path over time to test the robustness of the 'path' hypothesis.

An ever more important area for continued work is to develop a third grammar capable of interpreting object structures in more complex situations than those dealt with above: in experimental situations, for example, that involve partial occlusion of an object or close spatial interaction between a number of objects. Bower and Wishart have postulated a third conceptual rule (rule 3) to account for these final stages of the development of object understanding (Plaget's stage VI). This rule states that two or more objects cannot be in the same place or on the same path of movement at the same time unless they bear a spatial relationship to each other which involves the sharing of common boundaries. The interpretive adequacy of this rule could be evaluated against experimental situations in which just such boundary sharing occurs, as, for instance, when an object is placed on top of a platform, inside a cup, or behind a screen (Wishart and Bower 1983). Experiments of this form make up a valuable part of the research to explore the object concept and will be a focus of the next stage of this research.

The final aim of the research will, of course, be to move from describing each of the stages to modelling the rules which produce the actual changes upward from one conceptual stage to the next. A series of cost–gain acceleration studies with infants is at present in progress in an attempt to produce data which will give us some insight into these mechanisms for change (Bower 1981). Change seems to occur when the infant's erroneous descriptions of an object give rise to an uncooperative description of what is actually happening.

One last point is worth mentioning because of its implication for future attempts at modelling infant cognitive development. Prazdny has also recently produced a computer model of the rule 1 behaviours described above, a fact which might seem to render the present attempt redundant (Prazdny 1980). Prazdny examined twelve Bower tracking experiments. He then excluded one of them (experiment 5 here) from his model for reasons that are not entirely clear to us. In the other eleven situations he suggested that the experimental results could not fully support the place–movement analysis of Bower. He quoted one experiment in particular which he felt clearly demonstrated that stage II infants do attend to features and do not identify objects solely in terms of places and movements, an experiment in which infants showed upset at an unseen replacement of an object (Prazdny 1980, experiment 7). This would indeed undermine the analysis presented above. Such an experiment, however, although attributed to Bower (1977) has not to our knowledge been done by anyone as yet, least of all ourselves. Prazdny also rejected the Bower analysis of some of the other tracking results on the basis that an infant would not look for a 'missing' object (eg producing place or movement errors) if he was as dependent on immediate visual input as this analysis would have it. This represents a very basic misunderstanding of our analysis. It is central to the entire analysis that the infant's tracking behaviour is directed by conceptual rules and not by straightforward perceptual inputs. The missing experiment 12 of Prazdny (experiment 5 here) demonstrated this point conclusively—in that experiment the infant tracked forward or back to an empty space, a behaviour that obviously cannot be directed by any perceptual input.
APPENDIX

A more detailed explanation and trace of the computer analysis of experiment 3
To begin the experiment the object structures are created. In the case of experiment
3 this includes a sphere which moves across the field of vision and changes, in full
view of the infant, to a cube. The cube continues to move on the same path and
with the same speed as the sphere.
To start then, an object structure of type sphere is created at time \( t = 1 \). The
object is to move in three-dimensional space with coordinates \((x, y, z)\) in front of
the infant. The \( x \) parameter will change from 0 to 120, the \( y \), height, will be
constant at 4, and \( z \), the depth, a constant at 10. The infant is sitting at location
\((60, 0, 0)\). The object structure is then created over the time period, here from 1 to
30 with \( t = 15 \) occurring when the object is in the middle of the path at \((60, 4, 10)\).
For each time \( t \), the object structure, called OBJ1, is created (asserted into the data
base) with location, size, shape, and colour variables attached. For example, at
\( t = 1 \) the object structure is (comments in brackets):

<table>
<thead>
<tr>
<th>OBJ1 - time 1</th>
<th>loc(4,4,10)</th>
<th>shape(s)</th>
<th>size(3)</th>
<th>colour(g)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>location ((x, y, z))</td>
<td>shape is spherical</td>
<td>radius of 3</td>
<td>green</td>
</tr>
</tbody>
</table>

The sphere structure continues until \( t = 16 \) when OBJ1 turns into a cube. The
structure for \( t = 15 \) and \( t = 16 \) is:

<table>
<thead>
<tr>
<th>OBJ1 - time 15</th>
<th>OBJ1 - time 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>loc(60,4,10)</td>
<td>loc(64,4,10)</td>
</tr>
<tr>
<td>shape(s)</td>
<td>shape(s)</td>
</tr>
<tr>
<td>size(3)</td>
<td>size(3)</td>
</tr>
<tr>
<td>colour(g)</td>
<td>colour(g)</td>
</tr>
</tbody>
</table>

The cube continues to move until \( t = 30 \), when it is at location \((120,4,10)\), at which
time it disappears.
This concludes the 'object creation' phase of the program. Now any particular set
of rewrite rules (to represent analysis at that 'stage' of development) will consider
the object structures created. The code for grammar 1, presented in the text, is
called to consider a point \((3,7,10)\) near the object structure. The actual object
structure at \((4,4,10)\) is located, found interesting (parallax), and under the motion-rest
invariant is expected at the appropriate new location. OBJ1 is not occluded by any
other object throughout the sequence of time changes. In the trace it is referred to both
by its number (the number 1,2,3, ... of 'interesting' objects found in the sequence) and
by its shape, first a sphere and then changing to a cube. A sample of the trace follows.

Object 1 has interest at time 1
and is expected at place \((4,4,10)\).
The sphere called Object 1 is not occluded at time 1.

Object 1 has interest at time 2
and is expected at place \((8,4,10)\).
The sphere called object 1 is not occluded at time 2.