GENERATING SCRIPTS FROM MEMORY[1]

by

Walter Kintsch & Suzanne M. Mannes

University of Colorado

This is publication number 87-3 of the Institute of Cognitive Science at the University of Colorado, Boulder, CO 80309

[1] This research was supported by grant number MH - 15872 from the National Institute of Mental Health.

1. Scripts as Mental Structures

Scripts are representations of simple stereotyped event sequences. As such, they are a subtype of frames or schemata. They were originally introduced in computational models of language comprehension, where they fulfilled a number of important functions. First of all, they were needed for the kind of expectation-driven, predictive processing that was envisaged by programs such as SAM (Schank & Abelson, 1977) and FRUMP (DeJong, 1979). Given one event in a routine sequence, the script makes it possible to predict the next one. Furthermore, by knowing what is expected, the program is able to distinguish what is unusual and idiosyncratic. Scripts also are crucial to account for inferential processing in language comprehension. Their default values can be used to fill in where there are gaps in the text, scriptal information is also an important source of disambiguation, and last but not least, scripts save these programs and, presumably, human processors from the inference explosion which otherwise would quickly engulf them (e.g., Rieger, 1975).

In order to be useful, knowledge has to be organized in some way, and scripts, as well as frames, schemata, and semantic nets, provided that organization (e.g. Anderson, 1980; Graesser, 1981; Schank, 1980; Schank & Abelson, 1977). Scripts are claimed to be mental structures, and psychologists set out to demonstrate the psychological reality of these structures and to investigate their properties. Bower, Black, & Turner (1979) observed a high level of agreement when subjects were asked to list the characteristic events that they thought belonged to a script, and their order. Table 1, for instance, shows the results obtained in their study when subjects generated a "Grocery Shopping" script: subjects agreed quite well on the major events, and even about some of the minor ones. Bower et al. also show that scripts work much as they were supposed to in story understanding: subjects often falsely "remembered" default values, but they also remembered particularly well unusual events; scripts determined what was perceived as the gist of a passage. These results have been replicated by later investigators, as well as extended. The centrality of items in a script was shown to be an important factor (Galambos & Rips, 1982; Yekovich & Walker, 1986). Scripts were shown to be directional (Haberlandt & Bingham, 1984; Barsalou & Sewell, 1985); order information appears to be built into the script. Such findings led researchers to postulate scripts as permanent, stable mental structures.

However, the view that scripts are precompiled mental structures was soon challenged from several sides. Computationally, fixed mental structures like scripts turned out to be too inflexible to really serve the purposes for which they were originally designed (van Dijk & Kintsch, 1983; Schank, 1982). Fixed scripts just could not be fine-tuned sufficiently to fit widely different, ever changing contexts. Empirically, the data soon pointed away from the idea of stable mental structures. If scripts guide retrieval, how close events occur in the script structure should determine the time it takes to retrieve one event, given the other. However, such distance effects have not been observed; the data are messy (Haberlandt & Bingham, 1984; Galambos & Rips, 1982; Bower et al., 1979), but they clearly are not in agreement with script predictions. At most, it appears that immediately succeeding events are retrieved faster than events farther away (Bower et al.), but that argues more for a local relation like "next", rather than for a larger script structure. Furthermore, the good agreement among subjects about what belongs in a script and what does not may not reflect their reliance on a fixed, global script structure. It, too, depends a great deal on local cues (Mandler & Murphy, 1983).

For these reasons, there has been a general move away from considering scripts as fixed mental structures. Schank (1980, 1982) has turned to MOPs, that is smaller, more general memory organization packets from which scripts are to be constructed. But it is not clear that the same problems that proved too difficult to handle at the level of scripts will not eventually be encountered at the level of MOPs also. Others have proposed hybrid models which have not yet been worked out in detail. The suggestion is to combine an associative memory with a minimal script structure (e.g., the "bundle" notion of Haberlandt & Bingham, 1984; the associative-net-plus-procedures envisaged by Yekovich & Walker, 1986). In the present paper we follow up a more radical idea: we generate mental structures like scripts from a purely associative memory. We claim that knowledge is not pre-organized in terms of scripts and schemata, but that such structures are generated from an unorganized associative net in response to a specific task

Table 1
An Empirical Script Norm for Grocery Shopping. After Bower, Black, & Turner (1979).

ENTER STORE
GET CART
Take out list
Look at list
Go to first aisle
Go up and down aisles
PICK OUT ITEMS
Compare prices
Put items in cart
Get meat
Look for items forgotten
Go to checkout counters
Find fastest line
WAIT IN LINE
Put food on belt
Read magazines
WATCH CASHIER RING UP
PAY CASHIER
Watch bag boy
Cart bags out
Load bags into car
LEAVE STORE

Items in capital letters were mentioned by most subjects, items in italics by fewer subjects, and items in small case letters by the fewest subjects.

demand in a specific context. Only in this way can the flexibility and context sensitivity that characterize human script use be achieved. A detailed argument to this effect has been presented by Kintsch (in press). Here, we shall show how a behavior that has been taken as prima-facie evidence for the existence of scripts as mental structures can be generated from a knowledge base in which there are no pre-existing global structures like scripts. Specifically, we shall simulate how people go about listing the events which constitute common scriptal activities, such as going to a grocery store, as in Table 1.

2. Scripts and Categories

The approach taken here extends earlier work on generating conceptual categories to scripts. Considerable work has been done on how people generate members of conceptual categories like *animals, cars,* etc., following Bousfield & Sedgewick (1944). Walker & Kintsch (1985) have modelled this process by making two crucial assumptions.

First of all, they assumed that knowledge retrieval obeys the same laws as retrieval from episodic memory. The latter, of course, has been studied extensively in the laboratory, and several workable models of this process are now available. In particular, Walker & Kintsch assumed that the automatic component of the retrieval process can be described by the Raaijmakers & Shiffrin (1981) theory. Given a particular retrieval cue, this model predicts what will be retrieved, and when. On the other hand, the control processes which are necessary to put together an appropriate retrieval cue in the present situation are outside the Raaijmakers & Shiffrin model, and will be discussed below.

Secondly, Walker & Kintsch assumed that the retrieval process operates on an associative net in which categories are not explicitly represented. That is, nowhere in such a net is there an exhaustive list of all animals, all types of cars, etc. Some typical members of the category might be directly associated with the category name, but most would have to be retrieved indirectly, via one or more intervening member-nodes. Of course, there must be information stored with each category member that identifies it as such. In the simplest case this might be an associated IS-A proposition, though there are undoubtedly other tests of category membership. For instance, in a system with imagery, similarity to a prototype might be computed.

Consider how, according to this model, members of the category *cars* are retrieved. The original retrieval cue is the category name cars itself. This cue activates associated items according to the probabilistic process described by Raaijmakers & Shiffrin. Any item X thus activated is tested for category membership by searching the knowledge base for the proposition ISA[X,CAR]. If this test is positive, the item is retrieved and produced as a category member; if it is negative, another retrieval attempt is made. After L unsuccessful retrieval attempts, a new retrieval cue is constituted by adding to the element "cars" the last item retrieved. Then, the whole process is repeated with this new, composite retrieval cue. Examples of such composite retrieval cues used by the subjects of Walker & Kintsch are "cars in my dormitory parking lot", "cars my parents have owned", "cars I wrecked". The process stops after a certain number of composite retrieval cues have been tried without producing a (new) category member.

The category naming task is, of course, quite trivial, compared with many other retrieval tasks that people are faced with. In more difficult situations, the formation of a retrieval cue may involve a great deal more problem solving activity, such as the "retrieval by reformulation" observed by Williams & Hollan (1981).

The model sketched above generates lists of category members similar to the ones generated by human subjects. For instance, a plot of the cumulative number of retrievals against time is negatively accelerated and consists of sharp bursts of responses separated by gradually lengthening pauses approximating a scallop. Figure 1 shows the recall function typical of this task.

Scripts, however, appear to be generated in a quite different way. More specifically, the way retrieval cues are selected in the script generation task is quite different and more controlled than in category naming tasks[2]. Scriptal events, like category members, differ in typicality, but in addition, they are ordered whereas category members are not. For instance, scriptal events are easiest to retrieve in their natural order, while it is hard to generate category members according to some arbitrary ordering criterion like size (Barsalou & Sewell, 1985). As a consequence, the retrieval function for scriptal events is not a heavily scalloped negatively accelerated curve as for categories, but is essentially linear and smooth (Barsalou & Sewell, 1985; Gronlund & Shiffrin, 1986; Walker & Kintsch, 1985). In script generation, whenever a retrieval cue loses its effectiveness, the subject knows where to look next. Hence there are no long pauses when the subject is searching for a better retrieval cue. Obviously, the subject is using the order information associated with scriptal events for this purpose.

[2] Walker & Kintsch (1985) tried unsuccessfully to model script generation in the same way as category naming.

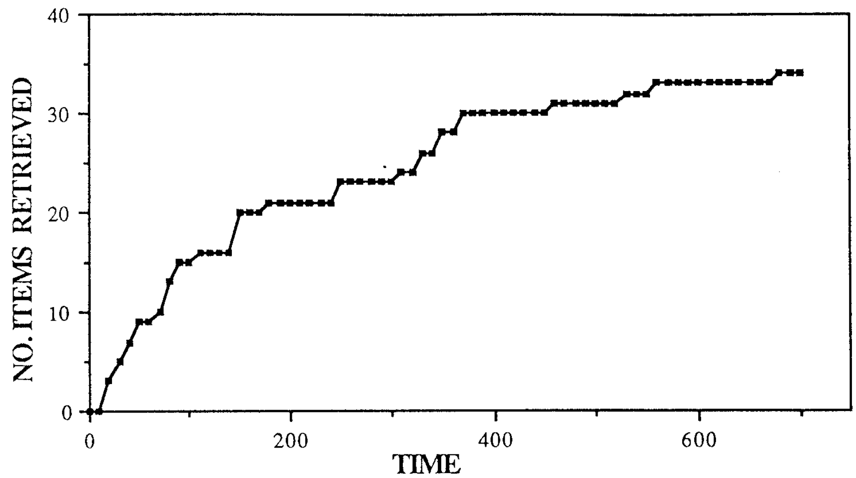

Figure 1. A representative retrieval function for one subject in a category retrieval task. Taken from Walker (1982). (Plotted in seconds.)

A model of script generation will be outlined below which is analogous to the category generation model of Walker & Kintsch, except that it utilizes a different control process to reconstitute exhausted retrieval cues: instead of randomly picking an associated node and adding it to the retrieval cue, local information about what comes next is used. In this way the unproductive search phases which are characteristic of category retrieval are avoided, and an essentially linear retrieval function is obtained. Furthermore, the retrieval process does not slowly peter out, but stops when the end of the chain is reached. Before describing this model in more detail, however, it is necessary to look more closely at how people generate scripts, so that we have a more solid data base with which to compare our model.

3. Generating Scripts: Experimental Results

When subjects are asked to generate members of some conceptual category, the data obtained are not very rich: we know which items they generate, and when. Such information is hardly sufficient to constrain theoretical speculations about the generation process. Walker & Kintsch (1985) have shown that richer, more powerful data can be obtained by having subjects provide concurrent verbal protocols about everything that goes through their minds as they are performing this task (Ericsson & Simon, 1984). In the experiment to be reported here, this protocol method is used with a script generation task.

3.1 Method

3.1.1 **Subjects.**

Six undergraduates from the University of Colorado participated in the experiment.

3.1.2 **Procedure.**

After a warm-up task, each subject participated in three script generation tasks. The three situations were *going to a restaurant for a meal, going to a grocery store to buy groceries, and going to a doctor's office for a checkup.* These were the same scripts that were used by Walker & Kintsch (1985). Each subject performed these three tasks in a different order. Subjects were asked to tell what typically happens in these situations, as if they were telling a stranger from another culture who knows nothing about them.

Subjects were instructed to "think-aloud" as they named the script events. It was emphasized that they were to express whatever thoughts crossed their minds and not to worry about producing proper English sentences. The subjects' protocols were tape recorded. The protocols were later transcribed with 5 sec time marks. The experimenter remained in the same room with the subject and reminded the subject to think out loud on the rare occasions when longer pauses occurred. Subjects were allowed 7 min for each naming task, though hardly any responses occurred after 5 min.

3.2 Results

Each subject's protocol was divided into idea units, roughly corresponding to propositions in the sense of van Dijk & Kintsch (1983). As will be explained below, these units were classified as either scriptal events or elaborations. Events were always single propositions, while elaborations were sometimes more complex. For example, the event *make a list* was elaborated with *of what you want to buy*. On the other hand, some events received more complex elaborations; for instance *go to store* was elaborated by *however you get there, driving or walking,* which was scored as a single unit.

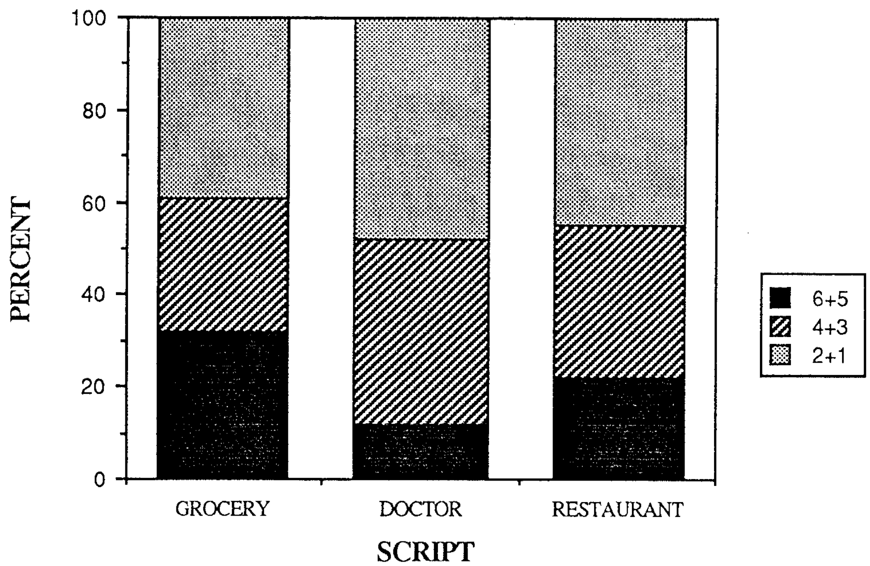

Figure 2. Percent of events agreed upon by 6+5, 4+3, and 2+1 subjects.

As in other experiments of this kind, there was good agreement among subjects about the events they thought belonged to each of the three scripts. Figure 2 shows the percentage of all responses which were produced by 5 or 6 subjects, 3 or 4 subjects, or by only 1 or 2 subjects. More than half of all responses were given by a majority of the subjects. The *doctor, restaurant, and grocery store* scripts elicited about the same amount of inter-subject agreement. Not only did our subjects agree among themselves, they also agreed with Bower et al. (1979). For instance, all of the high-frequency items in the grocery store script of Bower et al. (Table 1) were produced with a high frequency by our subjects.[3] Almost all items were generated in their natural order; only 1% of all responses were out of order.

Figure 3 shows the cumulative number of scriptal units as a function of retrieval time for one of our subjects. The important points to notice are first that the rate at which a script is generated is approximately constant. Certainly, there is no indication here (or with any of the other subjects) of a negatively accelerated retrieval function as is characteristic of category member naming. Secondly, the curve is relatively smooth, in contrast to the severe scalloping observed for category naming tasks. Both of these properties are true not only for individual curves but also for the average results shown in Figure 4. These data are in agreement with those of Barsalou & Sewell (1985), Gronlund & Shiffrin (1986), and Walker & Kintsch (1985).

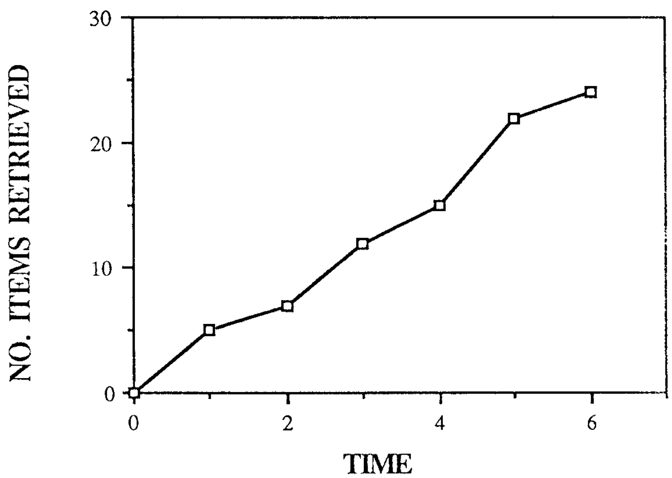

Figure 3. Number of grocery store items retrieved by a single subject as a function of time (plotted in 30 second intervals).

In listening to these protocols, one is struck by their fluency. Subjects always seem to know what to say next, with little hesitation. There was apparently no need for an extended search for a new retrieval cue. We therefore closely inspected each protocol for possible evidence as to the nature of the retrieval cues that permitted such smooth transition from event to event. In the course of these analyses we arrived at the distinction between scriptal events and elaborations. As already mentioned, events are single-proposition units, expressing an action either by the main actor or by some other participant (checkout-clerk, nurse, waiter) which can stand on their own. That is, events were meaningful by themselves, and did not require reference to some other unit for their interpretation. Elaborations on the other hand, have to be interpreted in the context of some other unit, usually an event. Thus, the event *make a list* was elaborated by *of what you want to buy*. Actually, the term *elaboration* is somewhat imprecise. What we classified as *elaborations* were true elaborations as in the example above, but also justifications (get a cart *if you plan to buy a lot*), and descriptions (*there are always some magazine racks there* (when speaking about the check-out counter)).

[3] This high level of inter-subject agreement contrasts sharply with the results obtained in a pilot study in which the same three scripts were employed with slightly different instructions. When subjects were asked to report what "you do when shopping in a grocery store" (instead of "what typically happens") highly idiosyncratic protocols were obtained, with much less overlap. See also Mandler & Murphy (1983).

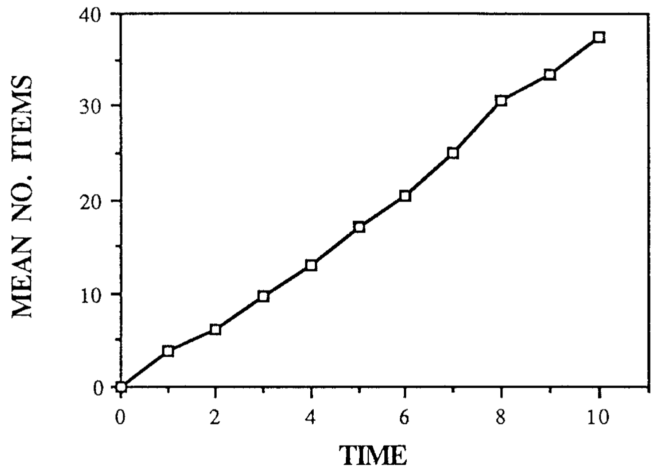

Figure 4. Mean number of items retrieved by 6 subjects as a function of time (plotted in 30 second intervals).

Thus, elaborations appear to be dependent on events. Events, however, either follow directly previous events, or are preceded by explicit temporal connectives, such as *then*, or *after that*. A striking asymmetry was noted in this respect: temporal connectives were used by most or all of the subjects in certain places, while they were rare otherwise. For instance, each subject used a temporal marker (connective) between the last item that dealt with preparation for shopping and entering the store and the first item that dealt with the actual shopping. The transition between shopping and checking-out was similarly marked by all subjects, and 5 of the 6 subjects separated checking-out and leaving the store with a temporal marker. Within each of these episodes, in contrast, temporal connectives were rare and used idiosyncratically. Thus, explicit temporal connectives appear to segment the script into separate episodes. In Figure 5 this pattern is shown for the three scripts. The labels we used to name these episodes are, of course, not very important and are entirely derived from our own intuitions. What is important, is that there exist objective indicators - the use of temporal connectives - that segment scripts into subclusters. We shall suggest below that the existence of these subclusters is crucial for an understanding of how scripts can be generated from an associative knowledge base.

Scriptal episodes may have quite different properties. In the case of the Grocery Store script, *Prepare, Check-out*, and *Leave* are quite similar, while *Shop* is a much less well constrained portion of the script. Relevant data are shown in Table 2. The three similar episodes are comparable in length (their mean lengths were used in drawing Figure 5) - subjects name about 4-6 scriptal events in each episode on the average. They agree quite well (the average intersubject agreement for events belonging to these episodes is about 3 out of 6). Between 9 and 11 different events are mentioned per episode. In contrast, the *Shop* episode is longer, many more different events are mentioned, and there is considerably less agreement among the subjects. If elaborations rather than events are considered, a similar pattern emerges: the *Shop* part of the script has many more elaborations than the other episodes, and is least constrained. In general, however, it is about the events belonging to a script that subjects agree so well - elaborations tend to be considerably more idiosyncratic.

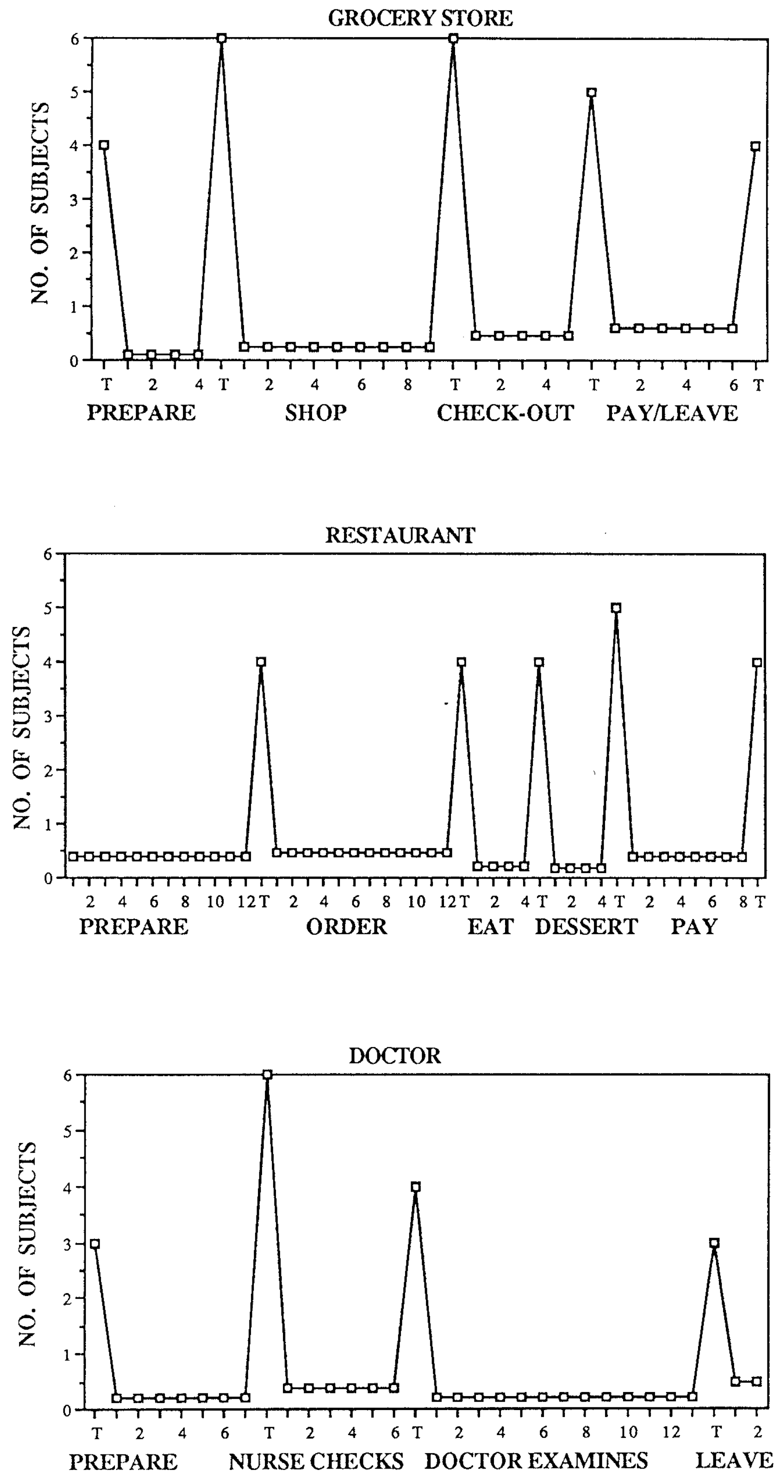

Figure 5. Subjects' placement of temporal markers for the three scripts.

Table 2
Scriptal episodes for Grocery Shopping: Number of events and elaborations, and inter-subject agreement.

	Episode: PREPARE	SHOP	CHECK-OUT	LEAVE
EVENTS:				
No. unique events	9	42	11	10
Total no. mentioned	26	56	32	36
Avg. mentions/event (max.6)	2.9	1.3	2.9	3.6
Avg. event/subject	4.3	9.2	5.3	6.0
ELABORATIONS:				
No. unique elaborations	12	45	10	4
Total no. mentioned	21	51	13	8
Avg. mentions/event (max.6)	1.7	1.1	1.3	2.0
Avg. event/subject	3.5	8.5	2.2	1.3

3.3 A Hypothesis about Retrieval

In Table 3 the events and elaborations for the grocery store script which were produced by our subjects are shown, broken down into episodes as suggested by Figure 5. Items generated by 5 or 6 subjects are shown with asterisks; inferred items (the episode labels) are shown in brackets. The reponses made by one subject are connected by a continuous line, to show this subject's retrieval path.

We are now in a position to state a possible hypothesis about the retrieval cues which control the process of script generation. The answer is simple for elaborations: since all elaborations can be co-ordinated with a specific event, we assume that the events serve as their retrieval cue. For the events themselves, we propose a dual process: some events are retrieved via specific temporal information, while others are retrieved associatively, much as category members are. Our data suggest that goal-directed retrieval on the basis of specific temporal cues, such as X FOLLOWS Y operates at the level of episodes: when enough information within an episode has been generated, the subject does not search for a new retrieval cue by checking random associations, but uses specific temporal information to establish a new episode cue. Thus, once the subject is done with *checking-out*, the knowledge base is searched for a proposition AFTER[CHECK-OUT, $], yielding AFTER[CHECK-OUT, LEAVING], and LEAVING becomes the next retrieval cue. The episode cue itself works much like the category cue in the category naming task: repeated retrieval attempts are made using this cue in conjunction with recently retrieved information. Thus, within-episode retrieval is locally governed, and stops when a certain number of unsuccessful retrieval events have occurred. At that point, a new episode cue is generated in the manner discussed above. Temporal information available in the knowledge base specifies what that cue has to be, and hence a long search is superfluous, giving the script generation process its characteristic smoothness and fluidity. In the next section, this

GROCERY STORE

[PREPARATION] T	TIME VISIT	SO THE STORE ISN'T CROWDED
	MAKE LIST	OF WHAT YOU WANT TO BUY
	EAT	IF FOOD SHOPPING OR MIGHT GET TOO MUCH
	GO TO STORE	HOWEVER YOU GET THERE, DRIVING OR WALKING
	PARK	IN FRONT OF STORE
	GO IN *	THROUGH DOORS
	GET CART *	WHICH ARE BIG METAL CARTS IF YOU ARE GETTING LOTS
	GET A BASKET	
[SHOPPING] T		
	START AT ONE END	POSSIBLY WITH FRUIT
	GO DOWN AISLES	IF YOU KNOW WHERE THINGS ARE
		AISLES ARE DESIGNATED BY NUMBERS
	START AT TOP OF LIST	
	LOOK IN DIRECTORY	CATEGORIES ARE LISTED ABOVE
	ASK SOMEONE	
	GO THROUGH PRODUCE	
	PICK LITTLE BAGS	POSSIBLY
	PUT ITEMS IN BAGS	THEN THEY WEIGH THEM FOR YOU WHEN YOU CHECK OUT
	GO TO EACH SUBSECTION	TO GET ANY SORT OF FRESH MEAT
	GET LETTUCE	IF YOU NEED IT FOR SALAD ONE NIGHT
	...	
	GO TO DAIRY	IF YOU NEED EGGS OR CHEESE
	NOTICE PRICES	

Table 3. Items which subjects produced in response to the Grocery Shopping cue. The lines show one subject's path through the network. See text for further explanation.

	GO TO FRONT	THERE ARE MAGAZINE RACKS
[CHECKING OUT] T *		
	GO TO CHECK OUT	THERE'S A CASHIER
	TAKE OUT GROCERIES	
	PUT BASKET UP	
	MACHINE READS CODES	
	SEE PRODUCTS OVER SCANNER	EVERY ITEM HAS A CODE
		PRICES ARE REGISTERED IN COMPUTER
	RING UP ITEMS	
	SEE ITEMS DOWN BELT	THERE'S A SACKER THERE
	CLERK PUTS GROCERIES IN BAGS *	DEPENDING ON THE NUMBER OF ITEMS
	CLERK PUTS BAGS IN CART	
[PAY/LEAVE] T *		
	CASHIER TELLS YOU THE AMOUNT *	
	YOU TAKE WALLET OUT	
	WRITE CHECK	
	PAY *	THE RIGHT AMOUNT YOU OWE
	GET CHANGE	
	WALK OUT	
	GET HELP WITH BAGS	IF YOU'RE OLD OR JUST LAZY
	CARRY BAGS OUT	
	LOAD CAR	
	GO HOME *	

Table 3. (continued).

hypothesis about how scripts are generated will be specified further to the point where simulations can be performed. The goal of these simulations will be to determine whether simulated script retrieval is at least qualitatively similar to the human data shown in Figures 2-5 and Tables 2-3.

4. Generating Scripts: A Simulation

Memory is an associative network with concepts and propositions as nodes. Nodes are related either positively, negatively, or not at all, with connections ranging from -1 to +1. Mathematically, this network will be represented by a matrix, **K**. The rows *(i)* and columns *(j)* of this matrix represent the nodes of the network, and the entries $(s_{i,j})$ represent the connection strengths between the nodes. The complete memory matrix would of course be huge, but in a simulation it is sufficient to work with a small sub-matrix, say of size $n \times n$.

A retrieval cue consists of one or more nodes of this network. Each retrieval attempt with a particular retrieval cue results in the retrieval of a single node from the network (though not necessarily in an overt response). The retrieval process can be conceptualized in two equivalent ways: once as activation spreading from the retrieval cues to other nodes in the net, followed by a probabilistic choice among the activated nodes; alternatively, one can think of retrieval in this network in terms of the Raaijmakers & Shiffrin (1981) model. Consider the spreading activation interpretation first. On a retrieval attempt, the nodes which constitute the retrieval cue are activated. Activation from these nodes then spreads to all nodes directly connected with them in proportion to the strengths of their connections.[4] However, how much these nodes will be activated depends on how much activation they receive from each of the nodes in the retrieval cue. That is, the total activation is the product (not the sum) of its components.

Suppose a retrieval cue consists of k components, r_i, $i=1,k$. If the activation of r_i is 1, the amount of activation that spreads to the other nodes in the system will be $s_{i,j}$, for $j=1,n$. Since each of the k components of the retrieval cue thus activates its neighbors, the total activation of the *j-th* node will be given by the product of these activation values,

$$(1) \qquad a_j = \prod_{i=1}^{k} s_{i,j} \; .$$

That is, each element of the activation vector **A** is the product of the inputs it receives from all components of the retrieval cue; in order to become activated, a node in the network must receive inputs from all components of the retrieval cue.

Given **A**, the probability that each element j will be retrieved is proportional to its activation value, relative to the total amount of activation (Luce, 1959):

$$(2) \qquad Pr(j) = \frac{a_j}{\sum_{j=1}^{n} a_j} \; .$$

Retrieval is thus an intersection-based spreading activation process, followed by a probabilistic response selection. As will be discussed in the final section of this paper, this allows us to relate the present model with a closely related model for knowledge use in discourse processing. We shall argue that the same knowledge base and the same activation processes are involved in both cases, but that they are used in different ways. Thus, otherwise separate

[4] This one-step version of the model could obviously be generalized.

models become building blocks for a more general theory of human knowledge structure and knowledge use. Except for these considerations, the model could have been stated more concisely (and without the need of introducing a spreading activation process) in terms of the well-known model of memory retrieval proposed by Raaijmakers & Shiffrin (1981). Memory in their model, is represented as an associative net, and the retrieval process which they describe results in the following equation (using the present notation, and neglecting other features of their model)

$$(3) \qquad Pr\,(j\,/\,i_{1...k}) = \frac{\prod_{i=1}^{k} s_{i,j}}{\sum_{j=1}^{n} \prod_{i=1}^{k} s_{i,j}}$$

which simply combines Eq.(1) and (2) in a single step.

Suppose that node j has now been retrieved. If it has not already been retrieved previously, it will be output as an overt response. An unsuccessful retrieval attempt occurs when the retrieved node duplicates an earlier retrieval.[5]

The newly retrieved node is added to the retrieval cue, which is now of size $k+1$. If this number exceeds some maximum value m, an old element must be dropped. We assume that this will be the most recently added element. Thus, the nature of the retrieval cue changes dynamically in response to local effects, but at the same time retains a stable core.

If a retrieval attempt did not result in an overt response (e.g. because an already retrieved item was produced), another retrieval attempt will be made. If L successive retrieval attempts are unsuccessful, a new retrieval cue is formed. The process that controls how retrieval cues are reconstituted determines what sort of memory structure is generated - it is the only feature in our model that determines whether a script or some other structure, e.g. a list of free associations, will be generated.

If subjects are instructed to generate a script, with some header H (such as Grocery Shopping, Doctor, etc), we assume that temporal information is used to generate the initial retrieval cue. That is, the proposition BEGIN[H, $] is formed, which yields via pattern matching BEGIN[H, H_1]. Thus, the initial retrieval cue is constituted, consisting of H and H_1. Propositions retrieved by that cue are added to it, in the manner described above. Once this cue becomes ineffective, i.e. L successive retrieval attempts have failed, it is abandoned, and a new cue is constituted. This is again done by a search guided by temporal information, except that we are now looking for what follows after H_1. Therefore, AFTER[H_1, $] becomes the basis of the pattern match, yielding H_2. H and H_2 now form the core of a new retrieval cue, and the process is repeated as long as another H_i can be found.

Free association differs from script generation only in that no temporal cues are used to guide retrieval. H is used to retrieve some associate which then takes on the role H_1. After it has become ineffective, some other associate of H takes its place. Thus, we obtain clusters of related associates, with occasional unpredictable jumps to new ones.

4.1 <u>Simulating Grocery Shopping</u>

In order to test the model outlined above, we attempted to simulate the generation of a grocery-shopping script. The goal of this simulation was to account qualitatively for the data

---

[5] In other cases this editing process must be more complex. E.g., in category naming a check needs to be made to determine whether the retrieved node is, in fact, a category member.

described in Section 3 of this chapter. Specifically, we wanted to see whether the simulation yielded an output comparable to the human data in Table 3, and whether the rate with which it was produced was constant, as in Figures 3 and 4. For this purpose, an associative knowledge system containing information about grocery shopping had to be constructed first.

4.1.1 **The Knowledge Base.**

Since merely a qualitative account is attempted here, we need not concern ourselves with all the knowledge people have about grocery shopping. Instead, fifteen typical items from Table 3 were selected more or less randomly. These were, grouped by episodes: *(go in/ get cart/ if you're getting lots) (getting vegetables/ noticing prices/ going down aisles/ categories are listed above/going to front) (take out groceries/ put basket up/ring up items/see products over scanner) (pay /the right amount you owe/walk out/ load car).* In addition, the script heading, the four episode headings as well as the temporal markers between them were included in the knowledge base. Furthermore, for each of the 15 scriptal events plus the 5 headings free associations were obtained from a group of 16 subjects. The two most frequently produced associations for each item were also added to the knowledge base. This resulted in the addition of 23 unique nodes, for in many cases scriptal events elicited the same high frequency associations (e.g. *grocery shopping, noticing prices, scanning prices,* and *paying* are all related to *money)*, or were already included in the original list. All the items in the knowledge base thus far are related to *grocery shopping*. To make the simulation more challenging, items unrelated to the grocery script are needed, to see whether or not the retrieval process suffers interference from such items. Therefore, a closely related script was selected, for which three of the four episodes overlapped. Subjects were asked to provide us with a *shoe-store* script. This yielded the three episodes *Entering, Shopping,* and *Leaving,* and the following 8 script events: *(entering store / looking over selection) (finding salesman / telling him shoe size / noticing prices / deciding how shoes fit / buying shoes) (leaving store).* Thus, in addition to the three episode headings, four of the scriptal events are associated with both scripts. If the two scripts will interfere with each other, there is certainly opportunity to do so. For each of the *shoe-store* events, two high-frequency associates were also added to the knowledge base. Once again, there was a certain amount of overlap: *money* appeared again; *noticing prices* was the top association for *looking over selection,* but later on it appeared as one of the scriptal events itself. (If this were a larger knowledge base, with more scripts in it, there would be even more common elements between them and hence there would be a fair degree of interscriptal connection.) There is, of course, in the knowledge base no distinction between scriptal events, episode headers, or associates: all are just nodes in a network, - we categorize them in this way only on the basis of the protocols subjects generate.

The total matrix thus constructed had 63 rows and columns - a tiny fragment of a human knowledge system. The 3,969 connection strengths $s_{i,j}$ in this matrix were estimated from actual script and free association data. Only rough estimates were made: Whenever a stimulus elicited the same response in 75% of the subjects, a strength value of 1 was used; responses given by fewer subjects were assigned a connection strength of .5; and responses which did not occur in our sample were given a strength of 0. All associations were assumed to be symmetric.

In addition, connections were established on the basis of the script data in the following way. The script headers *grocery shopping* and *shoe store* were connected with a strength of 1 to their respective episode headers and with a value of .5 to the scriptal events within each episode. The episode headers were connected with a strength of 1 to their respective script events. Finally, script events within each episode were connected to each other by a value of .5. However, the strengths of these connections were not necessarily symmetric: *picking up vegetables* and *going through the aisles* are connected both ways, and both are connected to *going to the check-out counter*, but the latter has no strength connecting it to either of the two former nodes. Thus, $s_{(vegetables,aisles)} = s_{(aisles,vegetables)}$, but $s_{(aisles,checkout)} \neq s_{(checkout, aisles)}$. Furthermore, the backwards connections between episode headers and their script headers, as well as between script events and their episode headers, were made less strong

(.5) than the forward connections between these elements. Thus, scriptal events may be directional. (It is of course possible that the same may be true for associations in general, but this possibility was not explored here.) The temporal connectives - BEGIN[GROCERY-SHOPPING, ENTER], BEGIN[SHOE-SHOPPING, ENTER], AFTER[ENTER, SHOPPING], AFTER[SHOPPING, CHECK-OUT], AFTER[SHOPPING, LEAVE], AFTER[CHECK-OUT, LEAVE] - which are needed to control the retrieval process in the script generation task are connected by a value of +1 to their second argument, and a value of -1 to their first argument. The only other thing that needs to be said is that all $s_{i,i}=1$, and that even though in some cases the sum of the association and script strengths would be larger, 1 is the maximum value. Obviously, some decisions which were made in constructing this matrix are difficult to justify in detail. However, we were guided by objective data in selecting script events and associations, as well as assigning connection strengths as far as possible. The decision to use only the values 1 and .5 for connection strengths is both crude and arbitrary, but by restricting ourselves to such rough approximations we have reduced the need for subjective judgments. It should also be noted that precise strength values matter relatively little here: It is the over-all pattern of connections that determines the results.

4.1.2 **Retrieval .**

Instead of estimating the parameters of the model from the data and trying to fit the data quantitatively, all the parameters were specified a priori.[6] Specifically, we assume that the retrieval cue has three components, two stable (which would be the script header and an episode header) and the third one variable (the last item retrieved). Hence, $m = 3$, which is a compromise assumption: the present model could not generate a script with a smaller m, while higher values would make a retrieval cue quickly ineffective in the relatively small network used here (not enough items are associated to all components). We also assume $L = 3$, that is, a retrieval cue is abandoned after three successive retrieval failures. Higher values of L would result in retrieval of a greater proportion of nodes in the network. Consider a particular simulation run with this model, then. What we observe is an interplay between the controlled search for a new retrieval cue, and the automatic retrieval process, based on Eqs. 1 and 2. Given Grocery Shopping, the proposition BEGIN[GROCERY-SHOPPING, $] is used to retrieve ENTER via a pattern matching process. These two nodes then form the first retrieval cue, activating 7 nodes in the system to some extent or other. By Eq. 2, *Entering Store* is selected. It is produced as a response, and becomes the third component of the retrieval cue. This modified cue now activates four other nodes, and, after 2 failures (the cue retrieves a component of itself), *Getting a Cart* is retrieved. This now replaces *Entering Store* as a component of the retrieval cue, which retrieves, after one failed attempt, a node which entered the network as an elaboration of Getting a Cart: *is fun*. The new retrieval cue activates 5 nodes, 3 of them new ones, but fails anyhow, because the same, old node is retrieved coincidentally three times in a row. Now the control process takes over, again: A memory search is made on the basis of AFTER[ENTER, $], which yields SHOPPING as the new episode cue, and the process once again shifts into its automatic retrieval mode. This interplay continues until the last episode cue is exhausted. In the simulation run under discussion here, a total of 12 nodes is thus retrieved, among them 5 nodes which we have called - on the basis of the data our subjects had given us - scriptal events; the rest were episode headers and elaborations. The data from this simulation run are shown in Table 4. This network looks much like an abbreviated version of Table 3: obviously, what the model does is quite close to what people do when they generate scripts. Most importantly, the retrieval process stayed on

[6] While it would be possible to estimate the parameters of the retrieval model statistically, a quantitative fit would presume an adequate simulation of the knowledge base, which is beyond our possibilities at present.

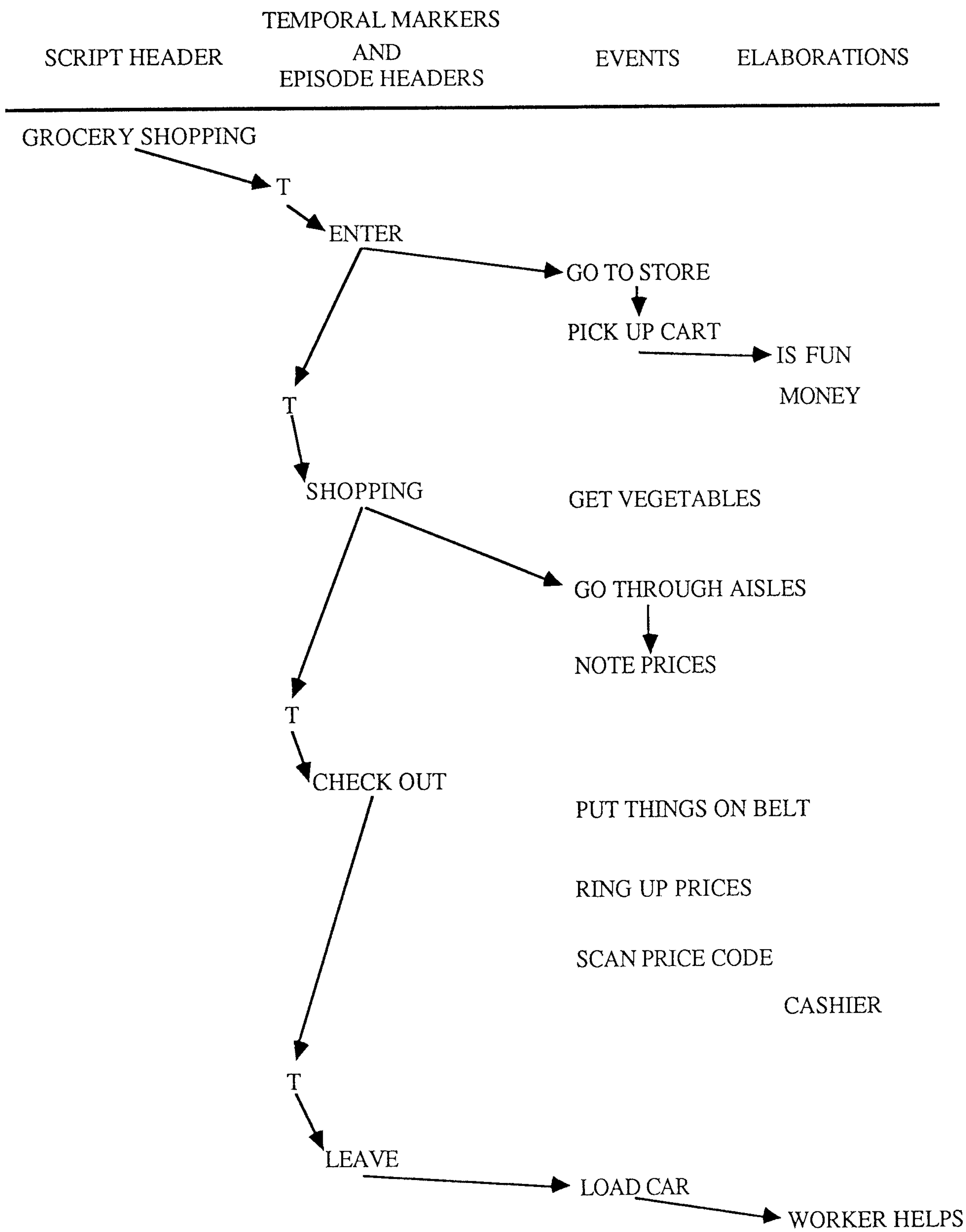

Table 4. Items which the model retrieved over three runs in response to the Grocery Shopping cue. The lines show the result of one run.

track: no intrusions from the shoe-store domain occurred, in spite of the fact that several of the items produced as part of the grocery script were also associated with the former domain. Secondly, the items were produced in the correct order, wherever order mattered, as in the actual data.

If each retrieval attempt is assumed to have a constant duration, overt responses can be plotted against retrieval attempts to estimate the time course of retrieval. This has been done in Figure 6 for three independent simulation runs with the grocery script. These simulation results have the same features as the data obtained from the individual subject in the script generation task (Figure 2): the rate at which scriptal events are produced is constant, there is no slowing down towards the end, and the curves are relatively smooth, without the large scallops which characterize category retrieval.

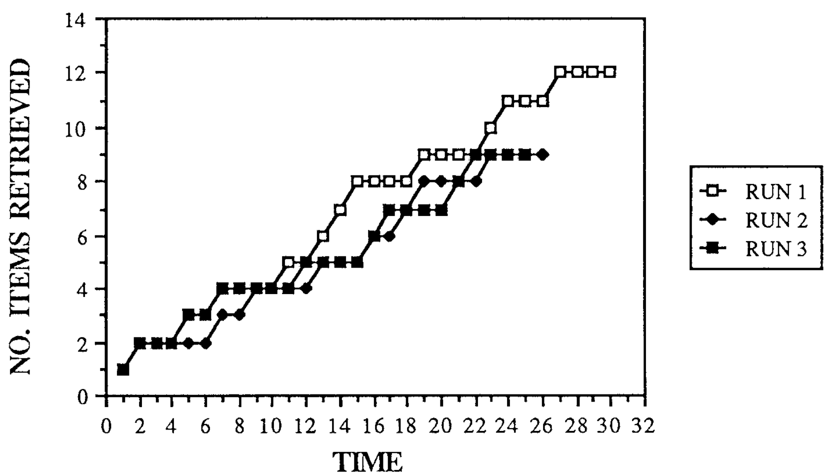

Figure 6. Number of scriptal items retrieved by the model in three runs as a function of time.

4.1.3 **Free Associations.**

Figure 7 shows the rate at which free associations to both grocery store and shoe store are produced by the model. To arrive at this figure, a different control process was assumed than for script generation, which does not involves the use of temporal cues. Given the script header, an associated item was produced. These two items then formed the first retrieval cue. As the third item was retrieved it was added to this retrieval cue, but the next item retrieved replaced it, and so on. After three unsuccessful retrieval attempts, this cue was abandoned, and the process was started all over again with the script header as the sole component of the retrieval cue. Once the script header itself lead to three successive retrieval failures, the whole process stopped.

The retrieval functions shown in Figure 7 are negatively accelerated and somewhat scalloped. They look more like the category naming function in Figure 1 than script generation functions. Nevertheless, they were generated from the same knowledge base, by the same retrieval process (Eqs. 1 and 2) as Figure 6 - only the control process was different. Scripts, categories, and associative structures do not reflect the organization of memory; rather, they are generated from an unorganized knowledge base - an associative net with only local connections between concepts and propositions. Structures need not be in the mind; they may result from particular kind of control processes - that is what this model suggests.

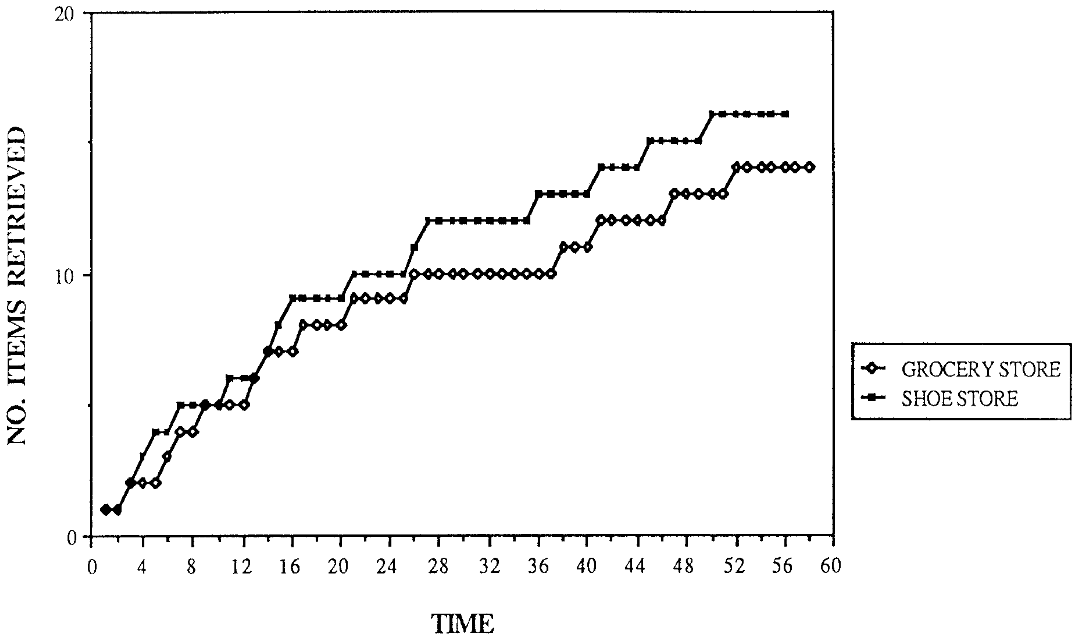

Figure 7. Number of items retrieved by the model in a free association task as a function of time.

5. Musings

Theories must be evaluated relative to their alternatives. In the present case, the principal alternative is the idea that scripts are mental structures in the sense of Schank & Abelson (1977). It is obvious that this alternative theory can account for the script generation data presented in Section 3 as well as the present model - and more simply. Why, then, bother with the present model? Why not be parsimonious, and choose the simplest account? There are two strong reasons for preferring the present model. First, as was outlined in Section 1, there are several empirical findings at variance with the classical script model; the present model does not have these problems, as will be shown below. Second, the idea of fixed mental structures is unsatisfactory for much more general reasons. As has been argued by Kintsch (in press), fixed mental structures may not be flexible enough to model human knowledge use. Much more context sensitivity is required than can be achieved with fixed structures, suggesting that mental structures need to be generated in the task contexts in which they are used. If they are prefabricated, they will never fit any particular context well enough to fulfill their intended function. Hence, it was suggested by Kintsch (in press) to explore associative networks as models of human knowledge organization. A construction-integration model was proposed to account for knowledge use in discourse comprehension. At the heart of this is a spreading activation mechanism, whereby each concept and proposition constructed during comprehension serves as an (independent) retrieval cue for related information - a bit as described above for the script generation task. The semi-random information thus retrieved is then integrated into a coherent textbase via a relaxation type mechanism. If knowledge use in discourse comprehension can be modelled as a spreading activation process in an associative net, it becomes very important to investigate whether other types of knowledge use can be modelled within the same framework. It is not obvious that scripts can be generated from a knowledge base which lacks precompiled structures. But scripts, schemata, and categories are certainly "psychologically real" in the sense discussed in Section 1 of this chapter. What we have done here is to show that this bit of psychological reality can be accounted for without having to postulate fixed mental structures, and, more specifically, in a way quite consistent with the proposed model of knowledge use in discourse comprehension. The model of script generation we have proposed here is in excellent agreement with the empirical data on script generation, beyond those aspects specifically investigated here. The well-known importance of the centrality of scriptal events (e.g. Yekovich & Walker, 1986) is given by associative strength. The directionality of scripts (e.g. Haberlandt & Bingham, 1984) is built into the present model by the choice of temporal connectives (AFTER, rather than BEFORE), and the use of asymmetric connections among scriptal events. Unlike the classical script notion, the present model does not predict general distance effects, though "next" effects at the episode level might be expected - which agrees reasonably well with the empirical evidence. Strategic control processes are, of course, the name of the game in the present model, and hence the evidence that under some conditions subjects exhibit high levels of agreement but not under others (e.g. Mandler & Murphy, 1983) can be taken as fairly direct support. As far as the role of scripts in memory and comprehension is concerned, more detailed processing models will have to be constructed, but this appears quite promising within the present framework. It is easy to see how networks of the kind proposed here would produce gist inferences if a text deals with scriptal activities. Similarly, predictions about false recognition of highly probable scriptal acts (e.g. Bower et al., 1979) could be derived within the present framework. Thus, whatever evidence psychologists have obtained about the psychological reality of scripts appears to lie well within the scope of the present model.

References

Anderson, J. R. (1980). Concepts, propositions, and schemata: What are the cognitive units? *Nebraska Symposium on Motivation, 28,* 121-162.

Barsalou, L. W. & Sewell, D. R. (1985). Contrasting the representation of scripts and categories. *Journal of Memory and Language, 24,* 646-665.

Bousfield, W. A. & Sedgewick, C. H. W. (1944). The analysis of sequences of restricted associative responses. *Journal of General Psychology, 30,* 149-165.

Bower, G. H., Black, J B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology, 11,* 177-220.

DeJong, G. F. (1979). Prediction and substantiation: A new approach to natural language processing. *Cognitive Science, 3,* 251-273.

Ericsson, K. A. & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Galambos, J. A. & Rips, L. J. (1982). Memory for routines. *Journal of Verbal Learning and Verbal Behavior, 21,* 260-281.

Graesser, A. C. (1981). *Prose comprehension beyond the word.* New York: Springer-Verlag.

Gronlund, S. D. & Shiffrin, R. M. (1986). Retrieval strategies in recall of natural categories and categorized lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 550-561.

Haberlandt, K. & Bingham, G. (1984). The effect of input direction on the processing of script statements. *Journal of Verbal Learning and Verbal Behavior, 23,* 162-177.

Kintsch, W. (in press). The representation of knowledge and the use of knowledge in discourse comprehension. In C. Graumann & R. Dietrich (Ed.) *Language in the social context.*

Luce, R. D. (1959). *Individual choice behavior.* New York: Wiley.

Mandler, J. M. & Murphy, C. M. (1983). Subjective judgments of script structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 534-543.

Raaijmakers, J. G. & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93-134.

Rieger, C. J. (1975). Conceptual memory and inference. In R. C. Schank (Ed.) *Conceptual information processing.* Amsterdam: North-Holland.

Schank, R. C. (1980). Language and memory. *Cognitive Science, 4,* 243-284.

Schank, R. C. (1982). *Dynamic memory.* New York: Cambridge University Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Erlbaum.

van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Walker, W. H. (1982). Retrieval of knowledge from memory: A generalization of the Raaijmaker-Shiffrin retrieval model. Unpublished Doctoral Dissertation, University of Colorado.

Walker, W. H. & Kintsch, W. (1985). Automatic and strategic aspects of knowledge retrieval. *Cognitive Science, 9,* 261-283.

Williams, M. D. & Hollan, J. D. (1981). The process of retrieval from very long term memory. *Cognitive Science, 5,* 87-119.

Yekovich, F. R. & Walker, C. H. (1986). Retrieval of scripted concepts. *Journal of Memory and Language, 25,* 627-644.