Psychological Research (2004) 69: 1–10
DOI 10.1007/s00426-003-0160-7

ORIGINAL ARTICLE

Johannes Engelkamp · Kerstin H. Seiler
Hubert D. Zimmer

Memory for actions: Item and relational information in categorized lists

Received: 19 June 2003 / Accepted: 13 October 2003 / Published online: 23 December 2003

Abstract Enacting action phrases in subject-performed tasks (SPTs) leads to better free recall than hearing or reading the same materials in verbal tasks (VTs). This enactment effect is usually explained by better item-specific information in SPTs than in VTs. The role of relational information is controversial. In the present paper, we will take the multiple recall approach to study the role of item and relational information in memory for actions by computing the number of item gains and the number of item losses over trials. This approach has previously been applied to lists of unrelated action phrases. We applied it to categorically related lists, also allowing a measure of relational information by clustering scores. It was found that SPTs produced more item gains than VTs. This finding confirmed the assumption that SPTs provide better item-specific information than VTs. The number of item losses did not differ between VTs and SPTs. This finding suggests that relational information is equally provided by VTs and SPTs. However, the organizational scores showed a more differentiated picture. Clustering was greater in SPTs than in VTs with randomly presented lists, but not with blocked lists. We suggested that these results, as well as other findings from the literature, could be explained by distinguishing automatic and strategic processes and the types of item associations that are addressed by these processes.

Introduction

The distinction between item-specific and relational information has proved to be fruitful in explaining explicit remembering (e.g., Einstein & Hunt, 1980; Hunt & McDaniel, 1993; MacLeod & Bassili, 1989; Smith & Hunt, 2000). Item-specific information refers to the information that is specific for each individual item and allows the item to be reintegrated and discriminated from other items. In contrast, relational information refers to associations among items. It serves particularly to support search processes in memory. Both types of information contribute to explicit memory performance (e.g., Hunt & Einstein, 1981; McDaniel, Einstein, Dunay, & Cobb, 1986). However, it is difficult to measure both components—item and relational information—directly at the same recall protocol. Relational information is mainly estimated from clustering scores that measure the extent to which the recalled items are organized according to the categories of the study list (e.g., Bousfield, 1953; Gollin & Sharps, 1988; Sternberg & Tulving, 1977). In contrast, item information is measured indirectly as old-new discrimination in a recognition memory test (e.g., Hunt & Einstein, 1981; Mohr, Engelkamp, & Zimmer, 1989). The only possibility of measuring item and relational information simultaneously is the calculation of gains and losses in a multiple recall procedure (Klein, Loftus, Kihlstrom, & Aseron, 1989). In the present paper, we will use this procedure to study the role of item and relational information in memory for actions of categorically structured lists.

J. Engelkamp (✉) · K. H. Seiler · H. D. Zimmer
FR Psychologie, Saarland University,
66123 Saarbrücken, Germany
E-mail: engelkamp@rz.uni-sb.de

Item and relational information and the multiple recall technique

With this technique, Klein et al. (1989) demonstrated that increased item encoding was reflected in an increased number of item gains, whereas increased relational encoding reduced the number of item losses. Item gains are defined as the number of items that are recalled in the actual trial but have not been remembered in the preceding test trial. Item losses refer to the number of items that are missed in the actual trial although they have been recalled in the preceding test trial. A number of studies (e.g., Burns, 1993; Burns & Gold, 1999; Burns & Schoff, 1998; McDaniel, Moore & Whiteman, 1998) confirmed and extended the finding of Klein et al. (1989) that gains and losses vary with item-specific information and relational information respectively.

The finding that good relational encoding protects against losing items in subsequent recalls was explained by the assumption that relational information serves to guide the retrieval of items. The presence of good and stable relational information provides reliable access to the learned items. In contrast, with weak relational information, consistent retrieval plans would not easily develop, but rather the generated retrieval cues would vary from trial to trial. As a consequence, the loss of items over trials is likely because formerly efficient retrieval cues are not used again in the actual recall (e.g., Burns, 1993; McDaniel et al., 1998).

The second finding stating that enhanced item information produces more item gains is explained by higher memory strengths of the encoded items. Focusing attention on each specific item produces a rich and extensive set of features that constitute that item. If some of these features are accessed during retrieval, the item is recovered with a high probability. However, because the items are not well integrated with each other, whether some item features are accessed or not strongly depends on the accidental retrieval context. Such items are available in memory, but not necessarily accessible. Therefore, if in a succeeding recall trial suitable retrieval cues are provided, a formerly not recalled item may now be recalled. An item with a less extensive set of features is generally less available, and is therefore less likely to be recalled in a succeeding trial. Hence, richly encoded items with good memory strength will produce more item gains than poorly encoded items (cf. McDaniel et al., 1998, Nairne, Pusen, & Widener, 1985).

Because there is ample evidence that item encoding is reflected in the number of item gains and relational encoding in the number of item losses, these measures can be used as indicators of the amount of item and relational information provided by specific encoding tasks (e.g., Mulligan, 2001; Olofsson, 1997). Such an approach is particularly interesting because it allows the contribution of item and relational information for free recall to be assessed directly and separately within the same memory task.

Item and relational information and the enactment effect

Memory for action phrases is usually compared under two task conditions. In verbal tasks (VTs), participants listen to a list of action phrases such as "open the door" or read such items and try to memorize them. In subject-performed tasks (SPTs), participants are requested to perform the denoted action on listening to it, usually without using real objects. A comparison of both tasks yields a robust memory advantage for SPTs, the so-called SPT effect (see Engelkamp, 1998, for an overview).

The SPT effect is generally attributed to better item-specific information in SPTs than in VTs (e.g., Engelkamp, 1988; Knopf, 1991; Kormi-Nouri, 1995; Zimmer & Mohr, 1986). A key argument for this claim is that enactment forces participants more to focus on the individual items than verbal learning does (e.g., Engelkamp, 1995; Zimmer & Engelkamp, 1999), and that encoding of action-specific features enhances the representation of each item (Zimmer, 2001). Hence, SPTs should provide good item-specific information and more item gains than VTs. The assumptions with regard to relational information are less consistent. Moreover, they differ depending on whether related or unrelated lists of action phrases are considered.

For unrelated actions, Engelkamp (1986, 1995; Engelkamp, Mohr, & Zimmer, 1991) assumed that enactment hinders the building of inter-item associations. In SPTs, attention is focused on the individual items and distracted from relational encoding due to the necessity of performing the action. In contrast, VTs allow for a more flexible allocation of attention to item and relational encoding processes. Participants in VTs get a less clear objective than participants in SPTs. Linked with the objective of retaining the items active item processing is permitted. Participants in VTs can deploy active processing strategies and search actively for inter-item associations (e.g., Engelkamp, 1998, pp. 96–98). The main evidence for this view was seen in findings of pair-associated learning of unrelated actions in VTs and SPTs. The effects observed in free recall of single actions were different from those in cued recall for the second action of a pair. In cued recall, the first element of a pair is given as a cue for the second (e.g., Engelkamp, 1986; Engelkamp, Mohr, & Zimmer, 1991). In single item-free recall, the usual SPT advantage over VTs was observed, whereas in cued recall the effect was reversed. Cued recall performance was poorer for SPTs than for VTs. This impairment was explained by the difficulty of associating unrelated actions under enactment. In other words, SPT encoding seems to be unsuited to providing relational encoding of unrelated items. The assumption that participants in VTs deploy more active relational processing strategies than in SPTs is also in line with the finding that free recall of VTs is more disrupted than that of SPTs by secondary tasks that also require controlled processes (Bäckman, Nilsson, & Chalom, 1986, Experiment 1; Engelkamp & Zimmer, 1996, Experiment 2). The idea that encoding is more strategic in VTs than in SPTs has already been suggested by Helstrup, (1987) and by Cohen (1981, 1983, 1985), although they focused more on the strategy-free encoding of SPTs than on the strategic encoding of VTs.

Therefore, Engelkamp and Seiler (2003) concluded that the more strategic relational processes of VTs should lead to a smaller number of item losses in a multiple recall paradigm than SPT encoding. This is exactly what they observed. With unrelated lists of action phrases, they observed better item-specific encoding, i.e., more item gains after SPTs than after VTs, and poorer relational encoding, i.e., more item losses after SPTs than after VTs. The same effects were observed by Olofsson (1997).

It would be interesting to know whether this pattern also holds for categorically related lists. On the one hand, it may be that SPTs hinder the use of categorical information too, because processing is still focused on the item due to enactment. On the other hand, unlike the unrelated lists, the categorical relational information exists pre-experimentally. Hence, in contrast to unrelated lists it is not necessary to generate new relations during encoding. If category information automatically comes into effect during retrieval, the categories may be used for retrieval even in the SPT condition. To our knowledge, no studies exist in which the multiple recall procedure was used with related lists, and the studies that analyzed clustering led to findings that are inconclusive.

Clustering in these studies was computed by the Adjusted Ratio of Clustering (ARC) scores (Roencker, Thompson, and Brown, 1971). This score assesses the degree to which the items belonging to the same category are recalled together (in a cluster). In most studies with categorically related lists, it turned out that the ARC scores under enactment and verbal learning did not differ statistically (Engelkamp & Zimmer, 1996, 2002; Norris & West, 1993; Zimmer & Engelkamp, 1989). However, there are also some exceptions that will be addressed below.

Engelkamp and Zimmer (e.g., 1996) explained the lack of finding different ARC scores in VTs and SPTs as follows. If action phrases are presented (notice that this is the case in VTs and SPTs) that belong to the same category (e.g., "put on the hat"; "put on the shoe") next to the specific action concept, they automatically activate their categorical super-ordinate concept (e.g., "dressing"). The category concept in turn activates other members of the category. Hence, a spreading activation process is assumed to take place similar to what is postulated in the context of semantic priming processes (e.g., Anderson, 1983; Meyer & Schvaneveldt, 1971; Neely, 1977; Perea & Rosa, 2002). If many exemplars of the same category are presented or recalled, the processing of category relational information benefits from pre-activation, and the category specific spreading activation processes are reinforced. These processes happen without the subjects' intention. The automatic spreading activation processes should support clustering in recall equally in VTs and in SPTs.

According to these assumptions, relational encoding processes in VTs and in SPTs differ for unrelated and for categorically structured lists. Relational processing of categorical relations is based on pre-experimentally established knowledge and occurs automatically (Mandler, 1979), whereas relational processing of unrelated lists is based on associations that were newly generated in the actual episode, which require strategic encoding processes (Engelkamp, 1998, pp. 96ff; Wippich, 1980, cf. also Zimmer & Mohr, 1986).

These assumptions would explain why with unrelated lists the number of item losses is smaller for VTs than for SPTs. They would also explain why ARC scores do not differ between VTs and SPTs with categorically related lists. Moreover, they would lead to the prediction that item losses between VTs and SPTs should not differ with related lists either because in both tasks automatic relational processes would be used.

However, the same predictions are made if it is assumed that categorical associations would also be activated strategically. In this case, automatic and strategic processes would provide the same categorical associations; again, the ARC scores and the number of item losses should not differ between VTs and SPTs.

Moreover, repeated recall tests should enhance both relational processes. The consecutive recalls should make the categories more and more obvious and enhance their strategic use, and they should reinforce activation of the super-ordinate category and also make the automatic process more efficient. Therefore, the ARC scores should increase and, correspondingly, the number of item losses should decrease over the recall tests.

Overall, the assumption that VTs allow for more strategic encoding processes than SPTs may have different effects on the number of losses in related and unrelated lists. There may be more stable retrieval processes in VTs than in SPTs with unrelated lists because SPTs lack an appropriate basis for automatic relational organization and stable retrieval processes.[1] This should be different with categorically related lists. In this case, there is also usually a material-induced basis for stable retrieval processes for SPTs, and the strategic processes of VTs may turn out to be redundant. Therefore, with related lists not only the clustering effects may be comparable in VTs and SPTs but also the number of losses. The present study was conceived to test whether item losses and ARC scores yield similar results with categorically structured lists. Moreover, due to better item encoding in SPTs than in VTs, we expect more item gains and better free recall in SPTs than in VTs.

Experiment 1

We presented categorically structured lists either in VTs or in SPTs. List presentation was followed by multiple recall tests. This procedure allowed us to compare VTs and SPTs with regard to ARC scores as well as to the number of item losses and item gains across the recall tests.

[1] It was demonstrated by Engelkamp and Seiler (2003) that order information (i.e., encoding of the accidental item neighborhood) is not the basis of the number of losses in unrelated lists.

Method

Participants

Fifty-two persons took part in this experiment. They were all students of Saarland University and were paid for participation. Half of them were randomly assigned to the VT condition and half to the SPT condition.

Material

As learning material, three different lists were constructed each consisting of 42 action phrases. The phrases were arranged in categories with 7 phrases each. The categories were created using taxonomic structures of the objects of the actions. For instance, the category "food" consisted of phrases such as "to grate the cheese," "to salt the potatoes," "to crack the egg," etc. During presentation, the items in the different categories appeared in random order with the constraint that at least two items from other categories had to appear between two items of the same category. We used three different lists in order to rule out our results being dependent on the specific categories used. The order of the items within a list was varied across participants, and each list was used in each encoding condition equally often.

Procedure

Participants were tested individually. Each subject studied one of the lists followed by five recall tests. Participants were instructed to learn the list of action phrases for a later memory test. It was not mentioned that the phrases could be organized into categories. Half of the participants were requested to learn the phrases by attentive reading (VT) the other half was instructed to learn the items by performing the actions without using real objects, i.e., pretending to perform the action (SPT). The action phrases were presented one after the other on a computer screen. Each phrase was preceded by a warning tone and was displayed for 5 s. After an interval of 500 ms, the next phrase appeared on the screen. After learning the phrases, the participants were requested to write down all of the phrases they could remember in any order they liked. This test was followed by another four successive free recalls. These recall tests were repeated without an intervening restudy phase. Each free recall lasted for 4 min.

Results and discussion

An action phrase was considered to be correct if the verb and the object of the phrase were correctly recalled. The recall performances are presented as a proportion of correct recall. Because the lists did not influence the results the data were collapsed across lists for further analyses.

The data of the recall test appear at the top of Table 1.

A 2×5 ANOVA with the factors encoding condition and recall tests yielded a main effect of encoding condition, $F(1,50) = 12.48$, MSE = .0883, $p < .001$, and a significant main effect of recall test, $F(4,200) = 16.22$, MSE = .0012, $p < .001$. This means that more items were recalled after SPTs (.42) than after VTs (.29), and recall increased in the course of the recall tests. The interaction between both factors was also significant, $F(4,200) = 3.70$, MSE = .0012, $p < .01$. Linear trends showed that recall increased over trials for SPTs, $F(1,50) = 37.24$, MSE = .0021, $p < .0001$, as well as for VTs, $F(1,50) = 5.46$, MSE = .0021, $p < .05$, but it increased more steeply for SPTs, $F(1,50) = 7.09$, MSE = .0021, $p < .01$.

In order to control the different recall rates in SPTs compared with VTs, we did not base our statistical analyses on the absolute numbers of item gains and item losses, but on their relative numbers (e.g., Mulligan, 2000). In the case of item losses, we divided the absolute number of losses by the number of items recalled in the previous trial. *Relative loss* therefore refers to the proportion of previously recalled phrases that were lost. This scoring takes into account that it is easier to lose an item from a large set than from only a small set of recalled items so that a lost item counts more at a low recall level than at a high recall level. The scoring also has the advantage that the scores can be compared more easily with the ARC scores, which are also relative. While it is plausible to weight item losses at the level of recalled items because participants with a high recall have a higher chance of missing items, the opposite may be expected for the number of item gains. Participants who have already recalled a high proportion of items do not have the same possibility of remembering additional items in the next recall trial as the participants who recalled a lower proportion. Therefore, we divided the number of item gains by the maximal number of items that could still be gained. *Relative gain* therefore gives the proportion of maximal possible gain that was realized by this participant.

The data for the mean relative number of item gains and item losses are depicted in Table 2.

The analysis of the number of item gains with a 2×4 ANOVA with the factors type of encoding and recall test showed an effect of type of encoding. There were more gains after SPTs (.10) than after VTs (.04), $F(1,50) = 20.43$, MSE = .0081, $p < .0001$. Neither the factor recall test nor the interaction between both factors was significant.

The corresponding analysis of item losses showed no main effect of type of encoding ($F < 1$). There were .09 item losses in VTs and .09 in SPTs. The factor recall test was significant, $F(3,150) = 7.55$, MSE = .0074, $p < .0001$. The number of losses decreased over test trials (linear trend: $F(1,50) = 16.55$, MSE = .0088, $p < .0001$). Both factors did not interact ($F < 1$).

The ARC scores are depicted in Table 3. The ARC score is "0" if no clustering takes place, and the degree of organization, i.e., the number of intra-category repetitions, is at the chance level. The

Table 1 Mean proportion recalled (standard errors in parentheses) as a function of recall test and encoding condition in Experiment 1, and in addition as a function of list presentation (random/blocked) in Experiment 2. *VT* verbal task, *SPT* subject performed task

	Recall test					
	FR1	FR2	FR3	FR4	FR5	Mean
Experiment 1: Random						
VT	.27 (.026)	.28 (.026)	.29 (.027)	.30 (.027)	.30 (.027)	.29
SPT	.39 (.024)	.40 (.026)	.42 (.027)	.43 (.027)	.46 (.029)	.42
Experiment 2: Random						
VT	.33 (.037)	.32 (.039)	.33 (.039)	.34 (.041)	.35 (.038)	.33
SPT	.50 (.023)	.50 (.019)	.52 (.020)	.55 (.020)	.57 (.021)	.53
Experiment 2: Blocked						
VT	.39 (.026)	.39 (.030)	.39 (.031)	.43 (.031)	.43 (.034)	.41
SPT	.47(.026)	.50 (.028)	.52 (.030)	.54 (.035)	.57 (.030)	.52

Table 2 Mean relative number of item gains and item losses (standard errors in parentheses) as a function of recall test and encoding condition in Experiment 1, and in addition as a function of list presentation (random, blocked) in Experiment 2

		Between tests				
		1–2	2–3	3–4	4–5	Mean
Experiment 1: Random						
Relative gains	VT	.05 (.010)	.05 (.009)	.03 (.008)	.03 (.009)	.04
	SPT	.11 (.016)	.09 (.015)	.09 (.009)	.10 (.083)	.10
Relative losses	VT	.15 (.021)	.08 (.019)	.07 (.017)	.06 (.014)	.09
	SPT	.12 (.025)	.09 (.017)	.10 (.018)	.06 (.015)	.09
Experiment 2: Relative gains						
Random	VT	.03 (.014)	.03 (.010)	.03 (.011)	.04 (.010)	.03
	SPT	.11 (.019)	.10 (.018)	.08 (.019)	.08 (.017)	.09
Blocked	VT	.08 (.019)	.04 (.011)	.05 (.010)	.03 (.014)	.05
	SPT	.13 (.017)	.10 (.018)	.11 (.023)	.10 (.019)	.11
Experiment 2: Relative losses						
Random	VT	.10 (.025)	.03 (.014)	.03 (.013)	.03 (.015)	.05
	SPT	.10 (.023)	.07 (.019)	.05 (.012)	.03 (.008)	.06
Blocked	VT	.12 (.027)	.08 (.029)	.02 (.012)	.04 (.013)	.07
	SPT	.10 (.027)	.06 (.010)	.10 (.023)	.03 (.010)	.07

Table 3 Adjusted ratio of clustering (ARC scores, standard errors in parentheses) as a function of recall test and encoding condition in Experiment 1, and in addition as a function of list presentation (random, blocked) in Experiment 2

	Recall test					
	ARC1	ARC2	ARC3	ARC4	ARC5	Mean
Experiment 1: Random						
VT	.28 (.106)	.24 (.119)	.59 (.126)	.61 (.097)	.71 (.092)	.49
SPT	.44 (.066)	.54 (.069)	.59 (.069)	.63 (.066)	.78 (.039)	.60
Experiment 2: Random						
VT	.23 (.017)	.43 (.138)	.56 (.090)	.71 (.084)	.64 (.101)	.51
SPT	.46 (.056)	.65 (.058)	.73 (.059)	.78 (.058)	.84 (.062)	.69
Experiment 2: Blocked						
VT	.71 (.072)	.80 (.073)	.86 (.051)	.81 (.060)	.88 (.066)	.81
SPT	.59 (.060)	.69 (.049)	.78 (.052)	.87 (.042)	.89 (.027)	.76

score is "1" if clustering is perfect and all items from the same category are recalled in succession. It is negative if clustering deviates systematically from categorical clustering. The score takes into account that the chance level for repetitions is a function of the recall level.

The ARC scores were analyzed in a 2×5 ANOVA with the factors type of encoding (VT vs. SPT) and recall trial. This analysis yielded no effect of type of encoding, $F(1,50) = 1.66$, MSE = .4952, $p = .20$, but the main effect of recall trial was significant, $F(4,200) = 10.37$, MSE = .1331, $p < .001$. The ARCs showed a highly significant linear trend, $F(1, 50) = 74.66$, MSE = .0701, $p < .0001$. They generally increased from early to late recall trials. Both factors did not interact, $F(4,200) = 1.44$, MSE = .1331, $p = .22$. However, if contrasting the ARCs in the early and late trials (first versus last two trials),[2] we yielded an interaction with type of encoding, $F(1,50) = 6.29$, MSE = .0651, $p < .05$. The first two ARC scores were higher in SPTs (.49) than in VTs (.26), $F(1, 50) = 4.50$, $p < .05$, whereas after three recalls both scores were comparable (SPT = .71, VT = .66, $F(50) < 1$).

The results confirmed the expectations. First of all and not surprisingly, an SPT effect in free recall was observed. Also, as expected, there were more item gains in SPTs than in VTs and, importantly, the number of item losses between VTs and SPTs did not differ. Globally, the ARC scores of VTs and SPTs were also comparable. Finally, as predicted, the number of item losses decreased and the ARC scores increased over recall trials. However, a closer look revealed a more differentiated picture for the clustering scores.

The ARC scores showed a substantial numerical SPT advantage over VTs for the first two ARC scores, which disappeared in later recall tests. The SPT advantage during the early recall trials—though not predicted—is of great interest because a corresponding finding was reported by Bäckman et al. (1986) in a one trial recall test, but it could not be replicated until now. Perhaps the finding that item losses and ARC scores showed different patterns reflects that different types of relational information were used in VTs and in SPTs. Participants in the VT condition may have formed subjective associations between list items that did not correspond to the intra-category associations. The subjective associations could have lowered the ARC scores. The categories may have become evident only over the recall trials, and the items were also clustered according to the categories in VTs.

The assumption that some VT participants may have organized the lists differently from the given list structure should be reflected in ARC scores that were more negative in the first two recall trials of VTs than in those of SPTs. In order to test this consideration, we counted the number of participants showing a zero or a negative ARC score. Out of the 26 persons in SPTs, 4 showed a zero or negative score in Trial 1 and/or 2, and out of the 26 persons in VTs, 11 produced a zero or negative score. This difference is significant

[2] In this analysis the third trial was not included because it is exactly the midpoint that divides the recall sequences into the two halves. However, the effect remains the same if we contrast the first two against the last three recall trials.

(χ^2 = 4.59, df 1, p < .05). A corresponding test yielded no significant effect for Trials 4 and/or 5.[3]

These post-hoc assumptions would be further strengthened if we were able to show that the advantage of ARC scores in SPTs over those in VTs is replicable, and that it disappears with a more obvious categorical list structure.

Experiment 2

Experiment 2 served two purposes: First of all, to replicate and generalize the findings of Experiment 1, and secondly, to test whether there will be no differences between ARC scores in VTs and SPTs if the list structure is sufficiently obvious.

In order to enhance memory and clustering, we shortened the list length to 36 items by using 6 instead of 7 items per category. Moreover, to manipulate automatic clustering we presented the lists in two versions. As before, the items were presented in random order to replicate the results of Experiment 1. In a second condition the items were presented blocked by categories to make the categorical structure obvious. Hence, there was a random version of the list and a blocked version of the list. With regard to ARC scores, we expected higher scores in SPTs than in VTs under random presentation and equal scores under blocked presentation. Moreover, the scores should be higher in the blocked version for VTs because the blocked version should enhance the strategic encoding of categorical information and reduce the number of deviant idiosyncratic associations. For SPTs, no such enhancement was expected.

In contrast, the number of losses should not differ between VTs and SPTs, and they should be the same for the random and the blocked versions because there are stable retrieval plans in all cases, no matter whether they follow the categories or some other structure.

Method

Participants

Seventy-two persons took part in this experiment. They were students of Saarland University and were paid for participation. Half of them were randomly assigned to the VT condition and half to the SPT condition. Half of the participants in each group received a random list and half received a blocked list.

Material

Three different lists each consisting of 36 action phrases were used. Each list consisted of 6 categories of 6 phrases each. The categories were again created according to the taxonomic structures of the objects of the actions. The lists were presented in two versions. They were presented in a random order as in Experiment 1 or in categorical blocks, i.e., all items of the same category followed each other during presentation of the list. Lists were counterbalanced over type of encoding.

Procedure

The procedure was identical to Experiment 1.

Results and discussion

Scoring of free recall performance was identical to Experiment 1. The three different study lists had no differential effects on the results. Therefore, the data were collapsed for further analyses. The data of the recall test are depicted in Table 1.

A 2×2×5 ANOVA with the factors type of list (random, blocked), type of encoding task (VT, SPT), and recall tests (1–5) yielded a significant effect of type of encoding, $F(1,68) = 26.71$, MSE = .0739, $p < .0001$, a significant effect of recall test, $F(4,272) = 31.28$, MSE = .0014, $p < .0001$, and a significant interaction of type of encoding task and recall trial, $F(4,272) = 5.71$, MSE < .0014, $p < .001$. None of the other effects was significant. More items were recalled after SPTs (.53) than after VTs (.37), and recall increased over the recall trials and it did so more for SPTs (.48 recall test 1, .57 recall test 5) than for VTs (.36 recall test 1, .39 recall test 5). For both encoding conditions the linear trends were significant; however, the linear trend components also interacted with type of encoding, $F(1,68) = 10.24$, $p < .005$. Recall of blocked lists did not differ from that of random lists (F = 1).

The data of gains and losses were computed in the same way as in Experiment 1. The data are shown in Table 2.

A 2×2×4 ANOVA of the relative number of item gains with the factors list type, type of encoding and recall trial yielded a significant effect of type of encoding, $F(1,68) = 35.97$, MSE = .0074, $p < .0001$. The relative number of item gains in SPTs (.10) was higher than in VTs (.04). Additionally, the number of gains generally decreased over recall trials, the linear trend was again significant, $F(1,68) = 6.94$, MSE = .0032, $p < .01$. No other effect was significant.

A corresponding ANOVA of the relative number of item losses showed only one effect. The main effect for recall trials was significant, $F(3,204) = 13.08$, MSE = .0054, $p < .0001$. The relative number of losses from recall 1 to 2 was .11, and it decreased to .03 from recall trial 4 to 5.

Table 3 gives an overview of the ARC scores.[4] The 2×2×5 ANOVA with the factors type of encoding, list type, and recall trial yielded a main effect of list type, $F(1,66) = 10.32$, MSE = .2792, $p < .01$. The ARC scores of blocked lists (.78) were higher than those of the random lists (.60). The ARC scores also differed over the recall tests, $F(4,264) = 24.83$, MSE = .0449, $p < .0001$. In general, they increased, but this was different for the two list types. There was an interaction of list type and recall test, $F(4,256) = 2.99$, MSE = .0449, $p = .05$. The increase of ARC scores over recall tests was stronger with random than with blocked lists as the interaction of the linear components revealed, $F(1,66) = 8.38$, MSE = .0462, $p < .005$. This is partially due to the fact that ARC scores of blocked lists started on a higher level (.65) than those of random lists (.34) and both scores were near to the maximum at the fifth recall. However, even at the beginning, when ARCs are not affected by ceiling effects, clustering in random lists showed a higher increase (.20) from first to second recall than in the blocked condition (.09). Overall, the ARC scores of VTs and SPTs did not differ ($F(1,66) = 1.14$, $p = .29$). But, as predicted, type of encoding and list type interacted, $F(1,66) = 4.24$, MSE = .2792, $p < .05$. Clustering in SPTs was the same in random and in blocked lists (.69 vs. .76, $F < 1$). In contrast, in VTs clustering was much better in blocked than

[3] Nairne, Riegler, and Serra (1991) and Golly-Häring, and Engelkamp (2003) showed that order information is not the basis of categorical clustering.

[4] In each VT condition, one participant was excluded from the analysis because one of their ARC scores was not defined due to division by zero. The corresponding ANOVA is therefore based on 70 participants.

in random lists (.81 vs. .51, $F(1,66) = 13.51$, $p < .001$). In other words, VTs had a strong benefit from a blocked presentation, which made the categorical relational structure obvious. Accordingly, in within list comparisons, the ARC scores of SPTs were higher than those of VTs with random presentations (.69 vs. .51), $F(1,66)$ 4.88, $p < .05$, but ARC in SPTs and VTs were equal with blocked lists (.76 vs. .81), $F < 1$. No other effects were significant.

The global pattern of results in Experiment 2 corresponded to that of Experiment 1. Again, there was an SPT effect in free recall, and more gains were observed in SPTs than in VTs. Again, the number of losses did not differ between VTs and SPTs, as was the case in Experiment 1. These findings support the assumption that the SPT effect is mainly due to item-specific information and that relational information plays a minor role. It seems that by and large relational encoding and retrieval with categorically related lists is comparable in VTs and SPTs. However, we conceived it as possible that different types of relational information might be provided in VTs and SPTs. This difference should be reflected in different ARC scores. Comparable ARC scores should be observed only as long as both encoding conditions make use of pre-experimentally existent knowledge, e.g., categorical list structures, and the information provided by strategic processes is redundant to automatically processed information.

Due to the latter restriction, we expected equal ARC scores in VTs and SPTs with a blocked list presentation. With a random list presentation, some participants in VTs should tend to organize the lists actively according to criteria that are not congruent with the categorical list structure. Therefore, the mean ARC scores in VTs (especially in the early recall trials) should be substantially lower than in SPTs. Moreover, with VTs, ARC scores should increase from random to blocked presentation. However, no such increase should be observed with SPTs because SPTs rely mainly on automatic clustering processes. These considerations are reflected in the interaction of type of list presentation and type of encoding.

Importantly, the variations in the relative number of item losses did not correspond to the variations in ARC scores. Item losses were independent of list presentation. This outcome may reflect that participants who used other organizational criteria than the categorical list structure could nevertheless have stable retrieval plans. This interpretation would mean that the number of item losses only reflected the stability of the retrieval plans, no matter what kind of associations the organization was based on.

A final effect worth mentioning is the fact that the free recall level is independent of the change of clustering from random to blocked lists in VTs. Such an independence of the recall level has already been observed before for other types of relational encoding (e.g., Engelkamp, Jahn, & Seiler, 2003; Serra & Nairne, 1993). We assume that the main reason lies in a flexible use of item and relational information in free recall dependent on the relative degree with which these types of information are available (cf. McDaniel, DeLosh, & Merritt, 2000; Seiler & Engelkamp, 2003; Zimmer, Helstrup, & Engelkamp, 2000).

General discussion

In the present study, we explored item-specific and relational encoding in categorically structured lists of action phrases by means of the multiple recall technique. The main findings can be summarized as follows. First of all, the well-established SPT effect was also observed in the present experiments. Free recall of SPTs was better than that of VTs. Secondly, there were more item gains over the test trials in SPTs than in VTs. Thirdly, the relative number of item losses did not differ between VTs and SPTs. Fourthly, ARC scores did not differ between VTs and SPTs if the list structure was obvious as in the case when the items of each category were presented in blocks, but they differed if the list structure was not immediately obvious as with a random presentation of the list items. In this case, the ARC scores of SPTs were greater than those of VTs. Correspondingly, the ARC scores of VTs increased from a random to a blocked presentation, but those of SPTs did not.

This pattern of findings and other results reported in the literature (e.g., Bäckman et al., 1986; Zimmer & Engelkamp, 1989) can be explained if the following assumptions are made.

With regard to item-specific information, we assumed, as did others (e.g., Knopf, 1991; Kormi-Nouri, 1995), that the instruction to enact the actions enhances item-specific encoding compared with VTs because enactment causes a cascade of action specific processes that enrich the item's memory trace (cf. Zimmer, 2001). The instruction forces the participants to understand the action, to concretize it in order to be able to plan and finally to perform it. Such an instruction does not only guarantee a deep conceptual encoding (e.g., Zimmer & Engelkamp, 1999). It also activates the motor system of the brain (e.g., Nilsson et al., 2000; Nyberg et al., 2001). This assumption can explain the SPT effect. We speculated that item-specific encoding should not necessarily differ between related and unrelated lists. It should be better in SPTs than in VTs, and therefore there should be more item gains in SPTs than in VTs. This pattern was clearly observed in Experiments 1 and 2.

With regard to relational encoding, the situation is more complex. In order to explain our findings, we suggested distinguishing between new associations based on idiosyncratic information in the study episode and categorical associations based on pre-experimental taxonomic knowledge. Additionally, we claimed that different types of processes exist to support encoding: Automatic processes based on spreading activation and strategic processes, which are voluntarily initiated. Like Engelkamp and Seiler (2003), we assumed that relational encoding in VTs is more strategic than in SPTs. Moreover, we assumed that this difference should hold for unrelated lists as well as for related lists. In order to explain the fact that we nevertheless observed an equal number of item losses in VTs and SPTs with related lists, we speculated that item losses only reflect the stability of retrieval processes, no matter whether this stability is based on automatic or strategic processes.

In order to explain our complex of ARC score findings, we assumed that automatic and strategic relational encoding processes should equally rely on the categorical structures in memory if related lists are used. This effect was observed with a blocked list presentation. In order to explain the finding that ARC scores were greater in SPTs and in VTs with a random presentation, we speculated that VTs and SPTs only address the categorical structures in memory in the same way if the list structure is sufficiently obvious. If the list items are presented randomly, the list structure may be not sufficiently obvious to avoid the strategic processes leading to associations that are inconsistent with the categorical list structure. Associations that are incongruent with the list structure could also explain the corresponding effect of Bäckman et al. (1986)—greater ARC scores in SPTs than in VTs. The fact that they presented real objects in SPTs, but not in VTs may have made the list structure in SPTs more obvious than in VTs.

In other studies, it has been shown that the strategic processes can also enhance the encoding of categorical information in VTs compared with SPTs. Under random presentation of the list items it helped participants in VTs more than in SPTs if they were pre-informed about the categories used in the list (Engelkamp, Seiler, & Zimmer, 2003; Zimmer & Engelkamp, 1989).[5]

Overall, it seems as if strategic processes may modulate and complete what is going on automatically if categorically structured lists are presented. Usually, the strategic processes that are assumed to be more efficiently used in VTs than in SPTs do not change the categorical clustering compared with the automatic processes. However, the equal clustering effects that were observed most frequently may be changed to lower ARC scores of VTs than of SPTs if the strategic processes do not use the categorical information from the list, and to higher ARC scores of VTs than of SPTs if the strategic processes are explicitly directed to the categorical list structure. In general, it seems that ARC scores vary less in SPTs than in VTs due to the fact that relational processes in SPTs are mainly automatic.

The variations in ARC scores do not necessarily correspond to the variations in the number of item losses. Item losses are assumed to indicate the stability of retrieval processes. We suggested that such a stability may be the result of automatic as well as strategic processes. Only ARC scores should reflect the content that the different processes address.

Admittedly, our theoretical assumptions so far remain somewhat speculative, and not all theoretical properties of our assumptions have been tested directly in this study. Nevertheless, we can summarize that they may explain not only the findings of the present study, but also results from the literature (e.g., Bäckman et al., 1986; Engelkamp & Seiler, 2003; Zimmer & Engelkamp, 1989), which until now could not be explained by one set of theoretical assumptions. In particular, our theoretical assumptions offer the possibility of explaining the variability of findings with regard to ARC scores, which previously seemed to be inconsistent.

A concluding note illustrates that the findings of item losses and ARC scores might be different in other research paradigms. Mulligan (2001, Experiment 4), for example, observed consistent differences in ARC scores and item losses as a function of type of encoding task in the generate-read paradigm. He found fewer item losses and better ARC scores for items that were generated than for those that were read. In that experiment, participants in the generate condition received blocks of word pairs in which the second words were nouns from the same category (e.g., stripes—zebra, bark—dog, mane—lion). The second word had to be generated on the basis of a fragment (e.g., z-br-) in the generate condition. This task should lead not only to a deeper conceptual encoding of the second words, but also reinforce their active intra-category relational encoding more than only reading the word pairs. Hence, more strategic encoding and better categorical-relational encoding are expected under these specific conditions for the generation task than for the reading task resulting in better ARC scores and fewer item losses in generation than in reading.

Therefore, it should be noted that the specific tasks must be analyzed carefully and that the findings of memory of actions cannot simply be generalized and applied to other materials and tasks. It seems that verbal tasks generally leave space for flexible encoding processes and that they correspond to strategic relational encoding processes. For action phrases, this suggestion can be specified by proposing that VTs allow for more flexible and strategic relational encoding processes than SPTs, which force the participants to focus on individual item processing and rely mainly on automatic relational encoding processes.

Acknowledgements This research was supported by a grant from the Deutsche Forschungsgemeinschaft (En 124/13). We would like to thank Nicola Ferdinand, Jochen Glössner, Hedda Janssen and Annabelle Saffran for their assistance with data collection.

[5] Whether the strategic processes in VTs are helpful seems to depend also on the specific categorical structure of the list. In most studies, the categorical list structure was based on the categorical status of the objects (e.g., actions with fruits such as "eat the apple," "peel the banana," "pick up the peach," etc.; e.g., Bäckman et al., 1986; Engelkamp Zimmer, 1996; Norris West, 1993). In some studies, list structure was based on categories of actions or on actions that belong to certain episodes (e.g., Zimmer Engelkamp, 1989, Experiments 1 and 2), or that were based on scripts, i.e., on typical action sequences (Engelkamp Zimmer, 2002). It seems, for example, as if VTs benefit somewhat more from script information than SPTs. A further factor that is differentially determining the organization processes comes to the fore if the lists are organized according to two structures. Koriat and Pearlman-Avnion (2003) used a list that could be organized according to conceptual categories (much like episodes) as well as according to motor categories, i.e., according to certain movement patterns (e.g., "to spread a cover over the car," or "to cover a patient with a blanket"). In this case, VTs were organized more than SPTs according to conceptual categories and SPTs more than VTs according to motor categories.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*, 261–295.

Bäckman, L., Nilsson, L. G., & Chalom, D. (1986). New evidence on the nature of the encoding of action events. *Memory & Cognition*, *14,* 339–346.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, *49*, 229–240.

Burns, D. J. (1993). Item gains and item losses during hypermnesic recall: Implications for the item-specific-relational information distinction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19,* 163–173.

Burns, D. J., & Gold, D. E. (1999). An analysis of item gains and losses in retroactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 978–985.

Burns, D. J., & Schoff, K. M. (1998). Slow and steady often ties the race: Effects of item-specific and relational processing on cumulative recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1041–1051.

Cohen, R. L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology*, *22*, 267–281.

Cohen, R. L. (1983). The effect of encoding variables on the free recall of words and action events. *Memory & Cognition*, *11*, 575–582.

Cohen, R. L. (1985). On the generality of the laws of memory. In L. G. Nilsson & T. Archer (Eds.), *Animal learning and human memory*. Hillsdale, N.J.: Erlbaum.

Einstein, G. O., & Hunt, R. R. (1980). Levels of processing and organization: Additive effects of individual item and relational processing. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 588–598.

Engelkamp, J. (1986). Nouns and verbs in paired-associate learning: Instructional effects. *Psychological Research*, *48*, 153–159.

Engelkamp, J. (1988). Modality-specific encoding and word class in verbal learning. In M. Gruneberg, P. Morris, & R. N. Snykes (Eds.), *Practical aspects of memory* (Vol. 1, pp. 415–420). Chichester: Wiley.

Engelkamp, J. (1995). Visual imagery and enactment in memory of actions. *British Journal of Psychology*, *86*, 227–240.

Engelkamp, J. (1998). *Memory for actions*. Hove: Psychology Press.

Engelkamp, J. & Seiler, K. (2003). Gains and losses in action memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *56A*, 829–848.

Engelkamp, J., & Zimmer, H. D. (1996). Organization and recall in verbal tasks and in subject-performed tasks. *European Journal of Cognitive Psychology*, *8*, 257–273.

Engelkamp, J., & Zimmer, H. D. (2002). Free recall and organization as a function of varying relational encoding in action memory. *Psychological Research*, *66*, 91–98.

Engelkamp, J., Mohr, G., & Zimmer, H. D. (1991). Pair-relational encoding of performed nouns and verbs. *Psychological Research*, *53*, 232–239.

Engelkamp, J., Jahn, P., & Seiler, K. H. (2003). The item-order hypothesis reconsidered: The role of order information in free recall. *Psychological Research*, *67*, 280–290.

Engelkamp, J., Seiler, K. H., & Zimmer, H. D. (2003). Differential relational encoding of categorical information in memory for action events. Manuscript submitted for publication.

Gollin, E. S., & Sharps, M. J. (1988). Facilitation of free recall by categorical blocking depends on stimulus type. *Memory & Cognition*, *16*, 539–544.

Golly-Häring, C., & Engelkamp, J. (2003). Categorical-relational and order-relational information in memory for subject-performed and experimenter-performed actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 965–975.

Helstrup, T. (1987). One, two or three memories? A problem-solving approach to memory for performed acts. *Acta Psychologica*, *66*, 37–68.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, *20*, 497–514.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, *32*, 421–445.

Klein, S. B., Loftus, J., Kihlstrom, J. F., & Aseron, R. (1989). Effects of item-specific and relational information on hypermnesic recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1192–1197.

Knopf, M. (1991). Having shaved a kiwi fruit: Memory of unfamiliar subject-performed actions. *Psychological Research*, *53*, 203–565.

Koriat, A., & Pearlman-Avnion, S. (2003). Memory organization of action events and its relationship to memory performance. *Journal of Experimental Psychology: General*, *132*, 435–454.

Kormi-Nouri, R. (1995). The nature of memory for action events: An episodic integration view. *European Journal of Cognitive Psychology*, *7*, 337–363.

MacLeod, C. M., & Bassili, J. N. (1989). Are implicit and explicit tests differentially sensitive to item-specific versus relational information? In S. Lewandowsky & J. C. Dunn (Eds.), *Implicit memory: Theoretical issues* (pp. 159–172). Hillsdale, NJ: Erlbaum.

Mandler, J. M. (1979). Categorical and schematic organization in memory. In C. R. Puff (Ed.) *Memory, organization and structure* (pp. 303–317). New York: Academic Press.

McDaniel, M. A., Moore, B. A., & Whiteman, H. L. (1998). Dynamic changes in hypermnesia across early and late tests: A relational/item-specific account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 173–185.

McDaniel, M. A., DeLosh, E. L., & Merritt, P. (2000). Order information and retrieval distinctiveness: Recall of common versus bizarre material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 422–435.

McDaniel, M. A., Einstein, G. O., Dunay, P. K., & Cobb, R. E. (1986). Encoding difficulty and memory: Toward a unifying theory. *Journal of Memory and Language*, *25*, 645–656.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234.

Mohr, G., Engelkamp, J., & Zimmer, H. D. (1989). Recall and recognition of self-performed acts. *Psychological Research*, *51*, 181–187.

Mulligan, N. (2000). Perceptual interference at encoding enhances item-specific encoding and disrupts relational encoding: Evidence from multiple recall tests. *Memory & Cognition*, *28*, 539–546.

Mulligan, N. (2001). Generation and hypermnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 436–450.

Nairne, J. S., Pusen, C., & Widener, R. J., Jr. (1985). Representation in the mental lexicon: Implications for theories of the generation effect. *Memory & Cognition*, *13*, 183–191.

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 702–709.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.

Nilsson, L. G., Nyberg, L., Klingberg, F., Åberg, C., Persson, J., & Roland, P. E. (2000). Activity in motor areas while remembering action events. *Neuro Report*, *11*, 2199–2201.

Norris, M., & West, R. (1993). Activity memory and aging: The role of motor retrieval and strategic processing. *Psychology and Aging*, 8, 81–86.

Nyberg, L., Petersson, K. M., Nilsson, L. G., Sandblom, J., Åberg, C., & Ingvar, M. (2001). Reactivation of motor brain areas during explicit memory for actions. *NeuroImage*, *20*, 1–8.

Olofsson, U. (1997). Win some, lose some: Hypermnesia for actions reflects increased item-specific processing. *Memory & Cognition*, *25*, 797–800.

Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research*, *66*, 180–194.

Roencker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin*, *76*, 45–48.

Seiler, K., & Engelkamp, J. (2003). The role of item-specific information for the serial position curve in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 954–964.

Serra, M., & Nairne, J. S. (1993). Design controversies and the generation effect: Support for an item-order hypothesis. *Memory & Cognition*, *21*, 34–40.

Smith, R. E., & Hunt, R. R. (2000). The effects of distinctiveness require reinstatement of organization: The importance of intentional memory instructions. *Journal of Memory and Language*, *43*, 431–446.

Sternberg, R. J., & Tulving, E. (1977). The measurement of subjective organization in free recall. *Psychological Bulletin*, *84*, 539–556.

Wippich, W. (1980). *Bildhaftigkeit und Organisation: Untersuchungen zu einer differenzierten Organisationshypothese*. Darmstadt, Germany: Steinkopff.

Zimmer, H. D. (2001). Why do actions speak louder than words: Action memory as a variant of encoding manipulations or the result of a specific memory system? In H. D. Zimmer, et al. (Eds.), *Memory for action: A distinct form of episodic memory?* (pp. 151–198). New York: Oxford University Press.

Zimmer, H. D., & Engelkamp, J. (1989). Does motor encoding enhance relational information? *Psychological Research*, *51*, 158–167.

Zimmer, H. D., & Engelkamp, J. (1999). Levels of processing effects in subject-performed tasks. *Memory & Cognition*, 27, 907–914.

Zimmer, H. D., & Mohr, M. (1986). *Organisation und Organisierbarkeit von Verben und Substantiven bei einer verbal-semantischen bzw.* 'modalitätsspezifischen' Lernweise (Arbeiten der Fachrichtung Psychologie Nr. 100). [Organisation and facilities of organisation of verbs and nouns after semantic or modality-specific encoding.] Saarbrücken, Germany: Universität des Saarlandes.

Zimmer, H. D., Helstrup, T., & Engelkamp, J. (2000). Pop-out into memory: A retrieval mechanism that is enhanced with the recall of subject-performed tasks. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, 658–670.