Cognitive Processing (2019) 20:45–53
https://doi.org/10.1007/s10339-018-0889-y

RESEARCH ARTICLE

The output monitoring of performed actions: What can we learn from "recall-recognition" performance?

Guangzheng Li[1] · Lijuan Wang[2] · Ying Han[3]

Received: 20 February 2018 / Accepted: 21 October 2018 / Published online: 26 October 2018

Abstract
Memories of simple action phrases are retained better following subject-performed tasks (SPT) than verbal tasks (VT), and this superior memory performance of SPT over VT is referred to as the SPT effect. Although research has been conducted on the SPT effect for more than 30 years, how well participants recognize whether they recalled the items successfully and the relationship between item-specific processing encoding and automatic retrieval have not been addressed. The present study used a 2 (instruction: with or without pleasantness rating) × 2 (type of encoding: SPT or VT) between-subject design and applied a "recall-recognition" paradigm to explore the above issues. For the "recall-recognition" performance, the SPT (and the VT with the pleasantness-rating task) produce very poor recognition of the correct recall of the last items (recency effect) and still poorer recognition of the middle items relative to VT alone, indicating that the retrieval process of these items in SPT needs less intention, effort, or monitoring, happens instantly, and involves a more non-obvious memory trace than that in VT alone. This was taken as support for the idea that an emphasis on item-specific information leads to automatic retrieval and thus poor awareness of the prior correct recall. We suggest that the SPT effect can be explained from the perspective of both encoding and retrieval.

Keywords SPT effect · "Recall-recognition" paradigm · Output monitoring · Item-specific processing

Introduction

Memories for simple action phrases (such as "open the book" or "break the tick") are retained better if participants are instructed to learn the phrases while performing the actions rather than learning the phrases by only listening or reading. The excellent memory performance for subject-performed tasks (SPT) compared to verbal tasks (VT) has been called the enactment effect or SPT effect (Cohen 1989; Engelkamp 1998). This SPT effect is very reliable and has been shown in associative recognition (Zhao et al. 2016), cued recall (Steffens et al. 2009), category recall (Engelkamp et al. 2005), and free recall (Schult et al. 2014) tests. The significance and generalizability of the SPT effect have greatly aroused the interest of researchers. However, most previous studies have investigated the mechanism underlying the SPT effect from a perspective of encoding or retrieval (for reviews, see Nilsson 2000; Roediger and Zaromb 2010), and few studies have combined encoding and retrieval in exploring the retrieval mechanism of SPT.

The strategy needed in encoding and retrieval

Several explanations have been offered to account for the SPT effect from the viewpoint of information processing. Strategic and automatic processes (e.g., Zimmer et al. 2001; Zimmer et al. 2000) have been, respectively, assumed in the encoding or retrieval of VT and SPT. Specifically, the non-strategic encoding view (Cohen 1981) and episodic integration theory (Kormi-Nouri 1995) have been assumed to explain the SPT effect. Cohen (1981) found that no levels-of-processing effects were obtained with SPT, and subjects reported that they did not use memory strategies under SPT condition. For this reason, researchers suggested that the encoding of SPT is non-strategic (Cohen 1981, 1983) or effortless (Kausler 1989). In contrast, Helstrup (1987) opposed Cohen's non-strategic encoding theory by

✉ Lijuan Wang
wanglj699@nenu.edu.cn

[1] Jiangsu Normal University, Xuzhou, China

[2] Northeast Normal University, Changchun, China

[3] Fordham University, New York, USA

proposing that subjects can also actively use intentional strategies to encode actions if they are specifically instructed to do so. Later, Cohen (1989) clarified that there are two underlying meanings of the non-strategic encoding theory. First, subjects do not need intentional strategies or active rehearsal in order to remember action phrases; second, even if they do, this probably would not enhance their memory performance. However, if the encoding of SPT is non-strategic, developmental effects should not be observed with SPT regardless of whether the subjects are elderly individuals or young children. In fact, the empirical findings from research in children (e.g., Mecklenbräuker et al. 2011) and elderly subjects (Schatz et al. 2010; Schatz et al. 2011) have reported age effects, which argue against the non-strategic encoding theory. The above studies show that the encoding of actions is not entirely non-strategic and that at least part of this encoding requires cognitive resources.

In other work, Kormi-Nouri et al. (1994) found that the physical and semantic components of SPT require equal attentional resources. Later, Kormi-Nouri (1995) suggested that the encoding of SPT is ultimately strategic and that it is not qualitatively different from VT encoding, arguing that the differences merely result from the high level of self-involvement while performing actions, which helps subjects to be more aware of their own actions. Enactment is considered the "glue" that cements the components of actions into a single or a tightly connected memory unit that is more easily accessible in retrieval. Findings from research in early development and aging support this view that the integration is a strategic ability which can be acquired or degenerate with age (Feyereisen 2009; Hainselin et al. 2014; Mangels and Heinberg 2006; Mecklenbräuker et al. 2011; Sahakyan 2010).

Nilsson and Bäckman (1989) analyzed the strategies required in SPT and suggested that SPT includes not only explicit components, but also implicit ones. However, unlike other researchers, they emphasized that the retrieval process in SPT might include automatic components, although this hypothesis was not supported by evidence. Since then, studies have gradually begun to focus on the retrieval process of SPT and reveal that SPT and VT have different serial position curves (e.g., Cohen 1983; Schatz et al. 2011; Seiler and Engelkamp 2003; Zimmer et al. 2000); VT exhibit a characteristic U-shaped curve, whereas SPT display an extended recency effect but lack a primacy effect. Primacy and recency effects mean that the first (primacy) and last (recency) items are remembered better than the middle ones. The extended recency effect is caused by automatic "pop-out" during retrieval (Zimmer et al. 2000). Automatic pop-out means that items near the terminal serial position can automatically pop into a person's memory without active searches. This specific mechanism especially extends the recency effect, which is seen as crucial for the SPT effect.

Seiler and Engelkamp (2003) pointed out that if items are highly distinct because of rich item-specific encoding, then a distinctiveness-based retrieval is dominant, which leads to an extended recency effect. However, if the item is learned through rehearsal, a search-based retrieval is dominant, which leads to a primacy effect. The lack of a primacy effect in SPT is interpreted as an indication of the non-strategic encoding. What's more, previous research revealed that the retrieval speed is much faster in SPT than in VT during the initial stage of recall, which supports the automatic pop-out mechanism (Spranger et al. 2008). In sum, the above results give reason for the assumption that SPT are more accessible and therefore retrieved faster from memory, which makes SPT less susceptible to awareness of the retrieval process. To our knowledge, no study has focused on the output monitoring mechanism of SPT. In addition, very few research efforts have investigated the mechanisms underlying the SPT effect from the perspective of a combination of encoding and retrieval.

The present study

We applied the "recall-recognition" paradigm to explore how well participants recognized whether they recalled the items successfully in SPT as well as the relation between item-specific processing encoding and automatic retrieval from the view of output monitoring. The output monitoring view proposed by Gardiner and Klee (1976) successfully explained the strategies needed in the retrieval process, assuming that subjects would apply search retrieval to some items and automatic retrieval to others. The output monitoring of search retrieval—the retrieval process of which involves intention, effort, and monitoring, happens over a long time, and contains an obvious memory trace—was shown to work well, with subjects remembering the items that were correctly recalled. In contrast, the output monitoring of automatic retrieval—the retrieval process of which does not need intention, effort, or monitoring, happens instantly, and involves a non-obvious memory trace—was less effective, with subjects failing to remember the items whether or not they were correct recalled. To test their theory, Gardiner and Klee designed the "recall-recognition" paradigm, which consists of three stages. The first stage is the learning phase, the second is the free recall phase, and the last is the "recall-recognition" phase, when all items that just learnt by the subjects are presented again and the subjects are asked to recognize whether they recalled the item in the free recall phase.

Studies have revealed that subjects have weak "recall-recognition" performance in the recency items; that is, subjects can hardly discriminate whether the items have been recalled successfully, indicating that the retrieval of the recency items is automatic and that the output monitoring of these items is less effective (Gardiner et al. 1977). In contrast, subjects have strong "recall-recognition" performance in primary items, implying that subjects can clearly remember whether these items were recalled successfully, suggesting that the retrieval of the primary items is strategic and that the output monitoring of these items is effective. As such, the "recall-recognition" paradigm is an excellent way to examine retrieval strategies and has been widely regarded as a suitable method of investigating output monitoring (e.g., Klee and Gardiner 1980; Koriat et al. 1988).

In the study, we aim to solve two key questions. First, we will apply the "recall-recognition" paradigm to explore how well participants recognize whether they successfully recalled the items of SPT from the viewpoint of output monitoring. If the SPT effect is due to the automatic pop-out mechanism in the retrieval process, meaning that items in SPT can automatically pop into an individual's consciousness without the need for an active search, then the subject may have difficulty recognizing the recalled items (especially for the recency items) in the "recall-recognition" test. We assume that the output monitoring will be much poorer for SPT than for VT—that is, that the "recall-recognition" performance of the middle and recency items in SPT will be worse than in VT. The serial position scores of the free recall are also collected to validate the existing research conclusions. We posits that SPT and VT will have different serial position curves, with VT having a U-shaped serial position curve with the primacy and recency effects, whereas an extended recency without a primacy effect shall be developed in SPT (Cohen 1983; Schatz et al. 2011; Seiler and Engelkamp 2003; Zimmer et al. 2000).

Second, we aim to explore the relationship between item-specific processing in encoding and automatic retrieval. Previous studies have tended to suggest that SPT promote item-specific processing, making the item more specific and distinguishable (e.g., Steffens et al. 2009; Kubik et al. 2014a, b, 2016; Mulligan and Peterson 2015; Schult et al. 2014). As mentioned above, excellent item-specific encoding generally leads to automatic retrieval and enhances the recency effect (Seiler and Engelkamp 2003). Therefore, we may speculate that the excellent item-specific encoding may cause the poor output monitoring. In order to test this hypothesis, we use a corresponding orienting task (subjects were asked to estimate the pleasantness of each single action in encoding) to enhance the item-specific processing during the encoding phase (Seiler and Engelkamp 2003). If the poor output monitoring is caused by item-specific processing, the memory performance will be enhanced by item-specific processing under the VT condition, but not the SPT condition. That is, the free recall performance will be equally strong for VT and SPT, and the output monitoring performance will be equally bad for SPT and VT in the additional item-specific processing.

Method

Subjects and design

Eighty-four students (37 males and 47 females, $M_{age} = 22.28$, $SD = 3.13$) from Jilin University participated in the experiment. All of the subjects reached naked eye or corrected visual acuity of 1.0 or more, and none of them had taken part in similar experiments before. A 2 (instruction: without or with pleasantness rating) × 2 (type of encoding: SPT or VT) between-subjects design was used.

Material

The experiment material consisted of 36 action phrases (e.g., "open the textbook," "pick up the pencil"), with each phrase made up of two to four Chinese characters. The actions represented by these phrases are physical movements that are common in daily life and that could be split into six different categories, such as cooking (e.g., cut the cabbage), mailing (e.g., paste the stamp), studying (e.g., open the textbook), cleaning the house (e.g., clean the window), driving (e.g., turn the wheel), and washing (e.g., wash the sweater). Twenty students who did not take part in the experiment were randomly invited to evaluate the phrase familiarity with a 7-point scale (1 point indicating the lowest familiarity and 7 points indicating the highest familiarity) before the experiment. The result showed that the mean of the phrase familiarity was 4.45 with a standard deviation of 1.0. In order to prevent subjects from applying category strategies in the learning process, all of the phrases were arranged in Latin square order according to their categories, and one of the arrangements was randomly chosen.

Procedure

The experimental procedure was similar to the one developed by Gardiner and Klee (1976) and was divided into four phases: an encoding phase, a free recall phase, a distracter task, and a "recall-recognition" phase. In contrast to the original experiment, there were four different encoding conditions in our work.

The encoding phase: First, a black "+" fixation point was presented on the computer screen (for 2000 ms), followed by an action phrase (for 6000 ms). Half of the subjects learned under the SPT condition; namely, they were asked to read the phrases silently and then perform the actions represented by the phrases. The other half of the subjects were asked to learn the phrases under the VT condition, meaning that they had to read the phrases silently but were prohibited from making any unnecessary hand movements. (We deleted the data if the participants made some unnecessary hand movements. In this experiment, four subjects' data were deleted because of unnecessary hand movements. Four other subjects were then recruited to take their place.) In each of the above conditions, half of the participants were required to judge the pleasantness of each single action on a scale from 1 (unpleasant) to 5 (pleasant) during the encoding phase both in VT and in SPT. The participants were asked to make quick decisions. Under the SPT condition, the rating followed the enactment.

The free recall phase: Directly after the presentation of the last item, a free recall test was conducted; subjects were required to report all of the phrases verbally within 5 min, and the contents of the report were recorded by the researcher using a voice recorder. During the free recall test, the participants were free to choose the order in which they retrieved the learned phrases.

The distracter task phase: After the free recall phase, the subjects were asked to perform 5 min of arithmetic tasks and write down their answers on an answer sheet.

The "recall-recognition" phase: In this phase, the original 36 action phrases were presented on the screen again, with each phrase appeared for 6000 ms. The subjects were required to recognize whether the presented phrase had been recalled in the free recall phase or not. If the subjects remembered that they had recalled it, they pressed the "X" key; if not, they pressed the "M" key.

The method of data analysis

The free recall scores in terms of the serial positions were analyzed separately. Three items were summed up into one triplet by referring to existing studies (Schatz et al. 2011), and the 36 items were divided into 12 triplets. As the memory performance of the middle items (the items of triplets 2–9) did not change much, the present study analyzed these items as a whole (Schatz et al. 2011).

In this study, the results of the "recall-recognition" tests were analyzed by signal detection methods. The hit rate (a "hit" meant a subject successfully recognized that the item had been initially recalled) and the false alarm rate (a "false alarm" meant that subjects recognized items that had not been initially recalled) were collected, and the mean "recall-recognition" performance (hit rate minus false alarm rate) for SPT and VT was calculated as the indices measuring output monitoring. The mean "recall-recognition" performance of triplet 1, triplets 2–9, and triplets 10, 11, and 12 was also calculated independently. Because the correct words recalled by each subject were different, the related data were processed manually.

Results

The results of the free recall performance

The free recall of an action phrase was considered correct if both the verb and the noun were correctly remembered. The serial position curves of different encoding conditions are shown in Fig. 1. In order to verify the previous research's conclusions, we analyzed the memory scores under the standard instruction conditions and the item-specific processing instruction conditions, respectively. Under the standard instruction condition, A 2 (encoding condition: SPT or VT) × 5 (serial position: triplet 1, middle triplets 2–9, triplets 10, 11, and 12) ANOVA was conducted to analyze the performance. The results showed a significant main effect of encoding condition, $F(1, 40) = 7.69$, $p < 0.01$, $\eta_p^2 = 0.16$, and serial position, $F(4, 160) = 8.34$, $p < 0.01$, $\eta_p^2 = 0.17$. Additionally, the interaction between encoding condition and serial position was significant, $F(4, 160) = 4.04$, $p < 0.05$, $\eta_p^2 = 0.09$.

A simple effect test revealed that under the VT condition, there is a significant difference in the memory performance in different serial positions, $F(4, 80) = 3.36$, $p < 0.05$, $\eta_p^2 = 0.14$. LSD tests demonstrated memory performance is significantly better in triplet 1 ($P < 0.05$) and triplet 12 ($p < 0.05$) than the middle triplets 2–9, and there are no other significant comparisons ($p > 0.05$), suggesting that a primacy effect and recency effect are observed under the VT condition. Under the SPT condition, the memory scores of different serial positions are also significantly different, $F(4, 80) = 8.21$, $p < 0.05$, $\eta_p^2 = 0.23$. The LSD test demonstrated that the performance is significantly better in triplets 10 ($p < 0.05$), 11 ($p < 0.01$), and 12 ($p < 0.01$) than the middle triplets 2–9. However, the performance of triplet 1 is not significantly different than the middle triplets 2–9 ($p > 0.05$). This indicates that SPT display an extended recency effect, but lack primacy effect.

At the same time, the present study also examined the memory performance of different encoding conditions under certain serial positions. The simple effect test showed there is a significant difference in different encoding conditions for triplets 10, 11, and 12. The corresponding values are, respectively, $F(1, 40) = 9.11$, $p < 0.05$, $\eta_p^2 = 0.19$, $F(1, 40) = 7.45$, $p < 0.01$, $\eta_p^2 = 0.16$, $F(1, 40) = 4.47$, $p < 0.05$, $\eta_p^2 = 0.10$; that is, for triplets 10, 11, and 12, the recall performance in SPT is better than in VT. A significant SPT effect was observed, indicating that the SPT effect is related to the extended recency effect.

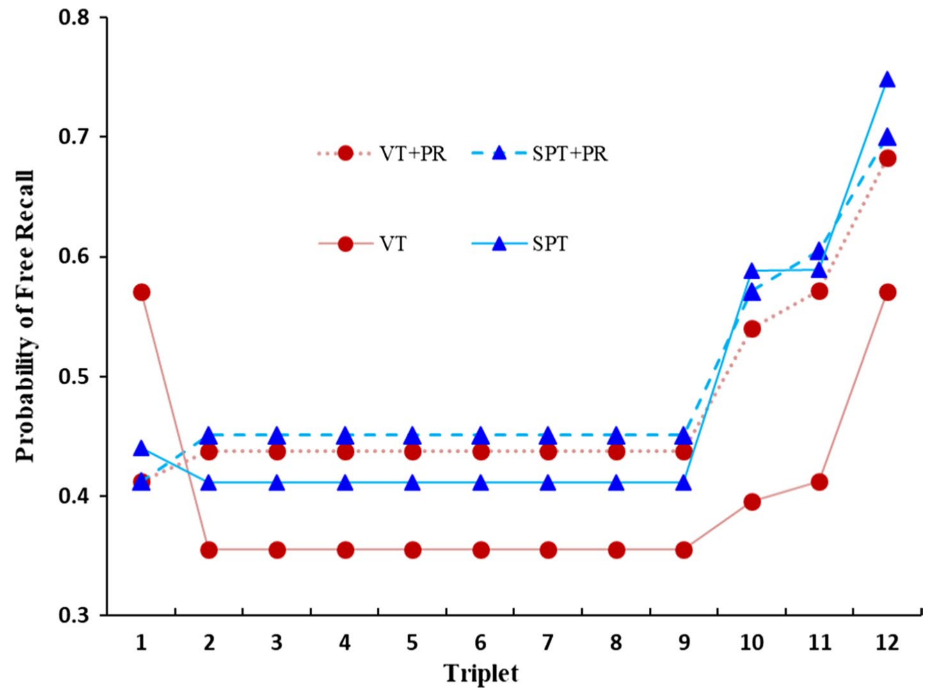

Fig. 1 Probability of free recall for different triplets under each condition

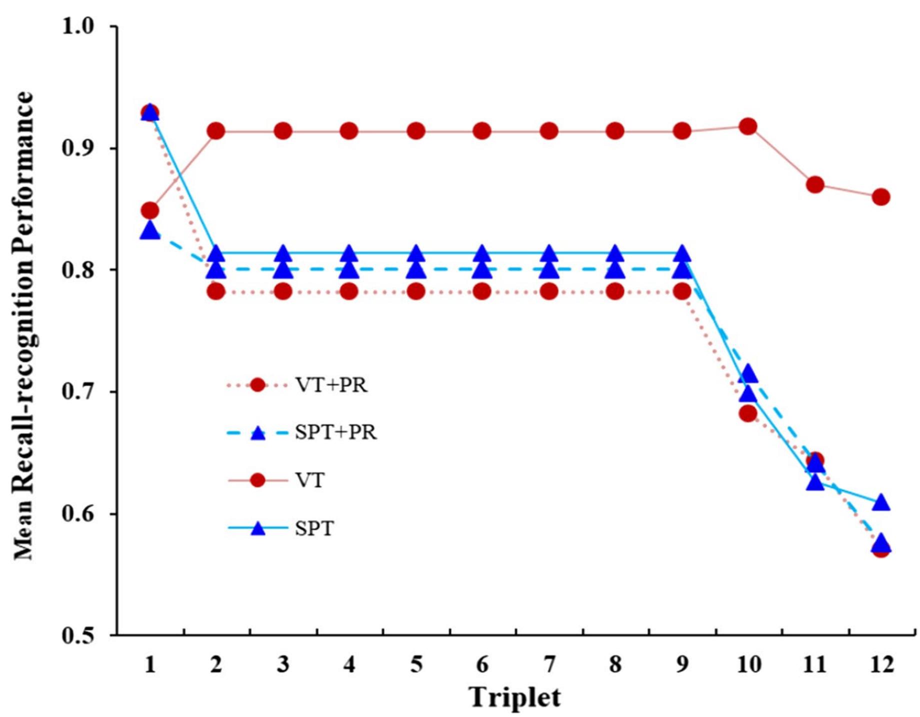

Fig. 2 Mean performance (hit rates minus false alarm rates) of "recall-recognition" test for different triplets under each condition

No significant difference was found with respect to the encoding condition for triplet 1 and triplets 2–9 ($p > 0.05$).

Under the item-specific condition, A 2 (encoding condition: SPT or VT) × 5 (serial position: triplet 1, middle triplets 2–9, triplets 10, 11, and 12) ANOVA on the free recall performance was conducted. The serial position curves of different encoding conditions are displayed in Fig. 2. The results showed a significant main effect of serial position, $F(1, 160) = 9.33$, $p < 0.01$, $\eta_p^2 = 0.20$. The LSD test showed the performance is significantly better in the triplets 10 ($p < 0.05$), 11 ($p < 0.01$), and 12 ($p < 0.01$) than the middle triplets 2–9. However, the performance of triplet 1 is not significantly different than the middle triplets 2–9 ($p > 0.05$). The main effect of encoding condition and interaction between encoding condition and serial position were not significant, all Fs < 0.16, all ps > 0.68, all η_p^2s < 0.01. The results supported the finding that the pleasantness-rating task cannot change the shape of the serial position of SPT, but can change the shape of the serial position of VT, which leads to parallel serial position curves in VT and SPT with no primacy effect but with a clear recency effect.

The results of "recall-recognition" performance

For the "recall-recognition" performance under different serial positions, a 2 (encoding condition: SPT or VT) × 2 (instruction: without or with pleasantness rating) × 5 (serial position: triplet 1, middle triplets 2–9, triplet 10, triplet 11 or triplet 12) ANOVA was calculated. The corresponding serial position curves are displayed in Fig. 2. The results showed a significant main effect of encoding, $F(1, 80) = 5.10$ $p < 0.05$, $\eta_p^2 = 0.06$, instruction, $F(1, 80) = 7.06$ $p < 0.05$, $\eta_p^2 = 0.08$, and serial position, $F(4, 320) = 11.59$, $p < 0.01$, $\eta_p^2 = 0.13$. The interaction between encoding condition and instruction was significant, $F(4, 152) = 4.08$, $p < 0.05$, $\eta_p^2 = 0.05$. Most importantly, these effects were all subsumed under the three-way interaction of encoding condition, instruction, and serial position, $F(4, 320) = 2.80$, $p < 0.05$, $\eta_p^2 = 0.03$. To follow up on the three-way interaction, separate analyses of each instruction condition were conducted subsequently.

For the standard condition, an interaction between encoding condition and serial position was observed, $F(1, 160) = 3.86$, $p < 0.01$, $\eta_p^2 = 0.09$, Further simple effect test showed the performance of different encoding conditions is not significantly different for triplet 1, $F(1, 40) = 1.31$, $p = 0.26$, $\eta_p^2 = 0.03$, whereas significantly different for triplets 2–9, triplets 10, 11, and 12. Corresponding values are $F(1, 40) = 5.41$, $p < 0.05$, $\eta_p^2 = 0.11$, $F(1, 40) = 7.22$, $p < 0.05$, $\eta_p^2 = 0.15$, $F(1, 40) = 7.32$, $p < 0.01$, $\eta_p^2 = 0.16$, $F(1, 40) = 9.60$, $p < 0.05$, $\eta_p^2 = 0.19$; namely, for triplets 2–9 and triplets 10, 11, and 12, the "recall-recognition" performance under the VT condition was better than that under the SPT condition. This indicates that, for those triplets, the output monitoring was much poorer in SPT than in VT. However, under the pleasantness-rating condition, the interaction was not significant, $F(1, 160) = 0.36$, $p = 0.83$, $\eta_p^2 = 0.01$, The main effect of encoding condition was not significant, $F(1, 160) = 0.02$, $p = 0.88$, $\eta_p^2 = 0.01$. The main effect of serial position was significant, $F(1, 160) = 8.23$, $p < 0.01$, $\eta_p^2 = 0.17$, with superior performance for triplet 1 than for triplets 2–9 ($p < 0.05$), triplets 10 ($p < 0.05$), 11 ($p < 0.01$), and 12 ($p < 0.01$), indicating that the pleasantness-rating task led to poor output monitoring of the middle and recency items in VT, making VT have a parallel serial position curve similar to that of SPT.

Discussion

The present study aimed to explore the retrieval mechanism from the perspective of output monitoring. Our main findings can be summarized as follows. Under the standard condition, the serial position curve of SPT lacked a primacy effect, yet showed an extended recency effect; the "recall-recognition" performance of SPT was significantly worse than that of VT, and the difference was specifically shown in triplets 2–9, 10, 11, and 12. Under the pleasantness-rating condition, the pleasantness rating improved the free recall performance of VT and changed the shape of the serial position of VT, which led to parallel serial position curves in VT and SPT with no primacy effect, but with a clear recency effect. The "recall-recognition" performance of the middle and recency items for VT was reduced by the pleasantness-rating task, making VT have a parallel serial position curve similar to that of SPT.

The often-reported serial position effects in free recall performance after VT and SPT are replicated in the present study. Specifically, VT have a typical U-shaped curve, while SPT have an extended recency effect instead of a primacy effect in the serial position curve, which supports the previous results (Cohen 1983; Schatz et al. 2011; Seiler and Engelkamp 2003; Zimmer et al. 2000). The lack of primacy effect indicates that SPT cannot take advantage of a rehearsal strategy. The presence of the extended recency effect might be because SPT can enhance the item-specific processing in the encoding process, and the recency items can decrease the distinctiveness of the former items and be more readily recalled (Schatz et al. 2011; Zimmer et al. 2000). However, we can only conclude that the extended recency effect accounts for the SPT effect if we use memory performance under different serial positions to analyze the origin of the SPT effect. These results do not indicate that the SPT effect might also benefit from the items presented in the middle triplet (Konpf 2005), although the memory performance of the middle items is also better in SPT than in VT from the data. The reason for this may be that this method is not sensitive enough to measure the retrieval mechanism of the SPT effect.

To overcome the disadvantage of the method described above, the present study performed a "recall-recognition" test to explore this issue. The results show that participants have difficulty in deciding whether the middle and recency items have been recalled under SPT conditions compared to VT conditions, indicating that a poor monitoring mechanism exists in the retrieval of those items; that is, these items may automatically enter into a subject's consciousness without an active search. Such a fleeting emerging process can be completed successfully without extra cognitive resources and consequently cannot leave an obvious memory trace. Hence, subjects tend to confuse the "items reported" and "items not reported" in the "recall-recognition" phase. For the VT group, search-based retrieval is mainly available for free recall (Seiler and Engelkamp 2003), which requires determination and effort that deepens the memory trace. Therefore, subjects can remember the retrieval items more clearly and distinguish "items reported" from "items not reported" more accurately in the "recall-recognition" test, reflected by the high performance.

However, for triplet 1, both groups had the same "recall-recognition" performance. It is because participants can voluntarily adopt a search strategy and tend to identify those items first presented as searching clues (Cohen 1983; Schatz et al. 2011; Seiler and Engelkamp 2003; Zimmer et al. 2000). That is, subjects from both the SPT and VT groups applied a search retrieval strategy for triplet 1. The results of "recall-recognition" are not completely consistent with the results of free recall in this study. These different results might be an artifact of the experimental methods used. It may be that the "recall-recognition" paradigm is more sensitive and direct in measuring the retrieval mechanism of the SPT effect, which extends the existing conclusion.

The results support the idea that the SPT emphasize item-specific information in the same way as the pleasantness-rating task. Specifically, VT can use the item-specific and relational information equally, so they display a typical serial position curve under the standard instruction (Seiler and Engelkamp 2003). When VT are given to the pleasantness-rating instruction, they upset the balance of the use of item-specific and relational information, making VT involve too much item-specific information and hinder the use of relational information. The result of this situation is that pleasantness rating leads to an enhanced recency effect in VT. However, SPT mainly rely on item-specific processing in encoding (e.g., Steffens et al. 2009; Kubik et al. 2014a, b, 2016; Mulligan and Peterson 2015; Schult et al. 2014), which can provide optimal encoding so that subjects can achieve a high memory performance without the need for additional encoding strategies such as the conceptual strategy (Zimmer and Engelkamp 1999). Thus, additional item-specific tasks cannot change the shape of the serial position curve in SPT. Therefore, the encoding and retrieval processes in VT approximate those of SPT, which leads to parallel serial position curves in VT and SPT with no primacy effect, but with a clear recency effect (Seiler and Engelkamp 2003).

The "recall-recognition" results demonstrated that item-specific processing leads to poor output monitoring of intermediate items and recency items. The item-specific instructions not only change the serial position curve of the free recall, but also change the serial position curve of the "recall-recognition." So the "recall-recognition" performance of VT with item-specific instructions is as bad as that of SPT, and both are significantly worse than those of VT under the standard condition. As mentioned above, the "recall-recognition" test can accurately detect whether the subjects consciously monitor the retrieval process. Specifically, if the retrieval process relies on search recall, which involves intention, effort, and monitoring, happens over a long time, and has an obvious memory trace, the subjects are better at monitoring the free recall process and can clearly remember whether the words have been recalled already. Thus, the "recall-recognition" performance would be strong in this test. However, if the retrieval was automatic—the retrieval process of which does not need intention, effort, or monitoring, happens instantly, and involves a non-obvious memory trace—subjects would fail to remember whether the words have been recalled already. Thus, the "recall-recognition" performance would be poor. From the standard instruction to the item-specific processing instruction, the "recall-recognition" scores of VT changed from good to bad, which indicates that the retrieval process underwent a shift from dependence on search-based retrieval to dependence on automatic-based retrieval. On the contrary, SPT mainly rely on automatic retrieval, so additional item-specific processing cannot change the mode of retrieval that it depends on.

The findings of the present study can further support the information processing view of the SPT (Schatz et al. 2011; Zimmer et al. 2000). The present study has discovered that the output monitoring of the middle and recency items in SPT was significantly poorer than in VT, indicating that the retrieval process of these items in SPT needs less intention, effort, or monitoring and can therefore support the non-strategic processing theory from the retrieval phase. We suggest item-specific information leads to automatic retrieval. Further work to investigate the SPT effect from the perspective of both the encoding and retrieval process might allow us to clarify the differences between SPT and VT and may lead to further improvements in explaining the origins of the SPT effect from an information processing view.

The results can explain some phenomena in daily life. Although we have unforgettable memories of some of our experiences, it is easy to forget whether we have told these experiences to others; therefore, we sometimes tell people about certain experiences repeatedly. In order to avoid this phenomenon, we can interact with the listeners more when we tell people about these experiences. In this way, due to the interaction with the listener, the automatic retrieval strategy is turned into a clue recall strategy, and the output monitoring in the retrieval can be enhanced so that one can more clearly remember that the event has been previously described to others.

Conclusions

The results highlighted three key aspects. First, the study replicated the result that SPT lack a primacy effect, yet display an extended recency effect relative to VT and that the pleasantness-rating task changes the shape of the serial position of VT, which leads to parallel serial position curves in VT and SPT. This supports the idea that the SPT emphasize item-specific information in the same way as the pleasantness-rating task. Second, the SPT (and the VT with the pleasantness-rating task) produce very poor recognition of the correct recall of the last items (recency effect) and poorer recognition of the middle items relative to the VT alone. Third, this was taken as support for the idea that an emphasis on item-specific information leads to automatic retrieval and thus poor awareness of the prior correct recall.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Number 31371022), the National Natural Science Foundation of Jiangsu Province (Grant Number BK20170235), and the Philosophy and Social Science Fund in Colleges and Universities of Jiangsu Province (2016SJB190010)

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Human and animal rights The procedure performed in the study involving human participants was in accordance with the ethical standards of the Ethnic Committee of Jilin University.

Informed consent Informed consent was obtained from all individual participants included in the study prior to testing.

References

Cohen RL (1981) On the generality of some memory laws. Scand J Psychol 22(1):267–281
Cohen RL (1983) The effect of encoding variables on the free recall of words and action events. Mem Cognit 11(6):575–582
Cohen RL (1989) The effects of interference tasks on recency in the free recall of action events. Psychol Res 51(4):176–180
Engelkamp J (1998) Memory for actions. Psychology Press, Hove
Engelkamp J, Seiler KH, Zimmer HD (2005) Differential relational encoding of categorical information in memory for action events. Mem Cognit 33(3):371–379
Feyereisen P (2009) Enactment effects and integration processes in younger and older adults" memory for actions. Memory 17(4):374–385
Gardiner JM, Klee H (1976) Memory for remembered events: an assessment of output monitoring in free recall. J Verbal Learn Verbal Behav 15(2):227–233
Gardiner JM, Passmore C, Herriot P, Klee H (1977) Memory for remembered events: Effects of response mode and response-produced feedback. J Verbal Learn Verbal Behav 16:45–54
Hainselin M, Quinette P, Juskenaite A, Desgranges B, Martinaud O, De La Sayette V, Eustache F (2014) Just do it! How performing an action enhances remembering in transient global amnesia. Cortex 50:192–199
Helstrup T (1987) One, two, or three memories? A problem-solving approach to memory for performed acts. Acta Psychol 66(1):37–68
Kausler DH (1989) Impairment in normal memory aging: Implications of laboratory evidence. In: Gilmore GC, Whitehouse PJ, Wykle ML (eds) Memory aging and dementia. Springer, New York, pp 41–73
Klee H, Legge D (1980) Remembering the recall of words. Can J Psychol 34(1):86
Knopf, M. (2005). Memory for subject-performed action events—In search of a memory theory. Unpublished presentation in Stockholm
Koriat A, Ben-Zur H, Sheffer D (1988) Telling the same story twice: output monitoring and age. J Mem Lang 27(1):23–39
Kormi-Nouri R (1995) The nature of memory for action events: an episodic integration view. Eur J Cognit Psychol 7(4):337–363
Kormi-Nouri R, Nilsson LG, Bäckman L (1994) The dual-conception view reexamined: attentional demands and the encoding of verbal and physical information in action events. Psychol Res 57(1):42–46
Kubik V, Obermeyer S, Meier J, Knopf M (2014a) The enactment effect in a multi-trial free-recall paradigm. J Cognit Psychol 26(7):781–787
Kubik V, Söderlund H, Nilsson L-G, Jönsson FU (2014b) Individual and combined effects of enactment and testing on memory for action phrases. Exp Psychol 61(5):347–355
Kubik V, Olofsson JK, Nilsson L-G, Jönsson FU (2016) Putting action memory to the test: testing affects subsequent restudy but not long-term forgetting of action events. J Cognit Psychol 28(2):209–219
Mangels JA, Heinberg A (2006) Improved episodic integration through enactment: Implications for aging. J Gen Psychol 133(1):37–65
Mecklenbräuker S, Steffens MC, Jelenec P, Goergens NK (2011) Interactive context integration in children? Evidence from an action memory study. J Exp Child Psychol 108(4):747–761
Mulligan NW, Peterson DJ (2015) Negative and positive testing effects in terms of item-specific and relational information. J Exp Psychol Learn Mem Cognit 41:859–871
Nilsson L-G (2000) Remembering actions and words. In: Craik FIM, Tulving E (eds) The Oxford handbook of memory. Oxford University Press, Oxford, pp 137–148
Nilsson LG, Bäckman L (1989) Implicit memory and the enactment of verbal instructions. In: Lewandowsky S, Dunn JC, Kirsner K (eds) Implicit memory: theoretical issues. Erlbanm, Hillsdale, pp 173–183
Roediger HL, Zaromb FM (eds) (2010) Memory for actions: How different?. Psychology Press, New York
Sahakyan L (2010) Rapid communication environmental context change affects memory for performed actions. Q J Exp Psychol 63(3):425–433
Schatz TR, Spranger T, Knopf M (2010) Is there a memory profit after repeated learning of subject-performed actions? Comparing direct and long-term memory performance level as a function of age. Scand J Psychol 51(6):465–472
Schatz T, Spranger T, Kubik V, Knopf M (2011) Exploring the enactment effect from an information processing view: What can we learn from serial position analyses? Scand J Psychol 52(6):509–515
Schult J, von Stulpnagel R, Steffens MC (2014) Enactment versus observation: item-specific and relational processing in goal-directed action sequences (and lists of single actions). PLoS ONE 9(6):e99985

Seiler KH, Engelkamp J (2003) The role of item-specific information for the serial position curve in free recall. J Exp Psychol Learn, Mem Cognit 29(5):954–964

Spranger T, Schatz TR, Knopf M (2008) Does action make you faster? A retrieval-based approach to investigating the origins of the enactment effect. Scand J Psychol 49(6):487–495

Steffens MC, Jelenec P, Mecklenbräuker S (2009) Decomposing the memory processes contributing to enactment effects by multinomial modelling. Eur J Cognit Psychol 21(1):61–83

Zhao M-F, Zimmer HD, Zhou X, Fu X (2016) Enactment supports unitisation of action components and enhances the contribution of familiarity to associative recognition. J Cognit Psychol 28(8):932–947

Zimmer HD, Engelkamp J (1999) Levels-of-processing effects in subject-performed tasks. Mem Cognit 27(5):907–914

Zimmer HD, Helstrup T, Engelkamp J (2000) Pop-out into memory: a retrieval mechanism that is enhanced with the recall of subject-performed tasks. J Exp Psychol Learn, Mem, Cognit 26(3):658–670

Zimmer HD, Cohen RL, Guynn MJ, Engelkamp J, Kormi-Nouri R, Foley MA (2001) Memory for action: a distinct form of episodic memory?. Oxford University Press, New York