Psychological Research
https://doi.org/10.1007/s00426-019-01217-y

ORIGINAL ARTICLE

Dual-memory retrieval efficiency after practice: effects of strategy manipulations

Franziska Heidemann[1,2] · Timothy C. Rickard[3] · Torsten Schubert[4] · Tilo Strobach[1]

Received: 11 October 2018 / Accepted: 12 June 2019

Abstract

The study investigated practice effects, instruction manipulations, and the associated cognitive architecture of dual-memory retrieval from a single cue. In two experiments, we tested predictions about the presence of learned parallelism in dual-memory retrieval within the framework of the set-cue bottleneck model. Both experiments included three experimental laboratory sessions and involved computerized assessments of dual-memory retrieval performance with strategy instruction manipulations. In Experiment 1, subjects were assigned to three distinct dual-task practice instruction groups: (1) a neutral instruction group without a specific direction on how to solve the task (i.e., neutral instruction), (2) an instruction to synchronize the responses (i.e., synchronize instruction), and (3) an instruction to use a sequential response style (i.e., immediate instruction). Results indicate that strategy instructions are able to effectively influence dual retrieval during practice. Mainly, the instruction to synchronize responses led to the presence of learned retrieval parallelism. Experiment 2 provided an assessment of the cognitive processing architecture of dual-memory retrieval. The results provide support for the presence of a structural bottleneck that cannot be eliminated by extensive practice and instruction manipulations. Further results are discussed with respect to the set-cue bottleneck model.

Introduction

In the context of advanced technologies such as smartphones and conversational systems, humans are increasingly engaging in multitasking (or dual-tasking) behaviors. During the last decades, a substantial body of research on dual-tasking has emerged from a variety of fields such as engineering, robotics, psychological and cognitive sciences, as well as medicine (Burgess, Veitch, de Lacy Costello, & Shallice, 2000; Chen & Yan, 2016; Jordan, Landau, & Iyengar, 2000). Even though all of these areas have distinct research foci, there is an emerging consensus that humans can engage in sequential (i.e., serial) as well as parallel response patterns in dual-task situations (e.g., Nino & Rickard, 2003; Ruthruff, Pashler, & Klaassen, 2001). A sequential response pattern in dual-task situations involves the execution of the first task (i.e., raising the left hand) before executing the second task (i.e., tapping with the right foot). In contrast, a parallel response pattern refers to the execution of at least parts of both tasks simultaneously. We refer to the execution of parts of the task, since it is still unknown which processes can occur in parallel and which cannot. For example, while a human being is well able to raise the left hand while tapping the right foot at the same time at a motoric level, we do not exactly know which mental processes are required to execute these tasks and if they can also operate simultaneously. This is especially interesting since dual-tasking has gained a lot of interest in various research areas, but there has been less focus on the investigation of the processes that are involved in dual-memory retrieval. Whereas dual-tasking in general refers to the global process of the execution of two tasks, dual retrieval refers to the retrieval process of two tasks in the context of various dual-tasking paradigms. This process evokes a search for mechanisms that might allow or prohibit different forms of dual-memory retrieval. One of the core matters in this research area is the investigation of the mechanisms that account for the emergence of different response patterns in dual-memory retrieval experiments. Therefore, the present paper investigates the strategic basis

✉ Franziska Heidemann
Franziska.Heidemann@ruhr-uni-bochum.de

1 Medical School Hamburg, Hamburg, Germany

2 Behavioral and Clinical Neuroscience, Ruhr-University Bochum, Massenbergstraße 9-13, 44787 Bochum, Germany

3 University of California, San Diego, USA

4 Martin-Luther University Halle-Wittenberg, Halle, Germany

of sequential and parallel response patterns in situations involving dual retrieval from long-term memory, with the aim of providing new insights into the underlying cognitive processes.

Dual-memory retrieval after practice: two retrieval patterns

A small number of studies assessed the mechanisms of practiced dual-memory retrieval from a single cue (Nino & Rickard, 2003; Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014). In these studies, subjects first learned an association between each of a set of color words, and both a vocal digit and a keypress response. In the vocal task, upon seeing a cue word (e.g., *red*), subjects had to speak a unique digit (e.g., "five") into a microphone. In the keypress task, they had to press a right- or left key on a keypress pad. After this learning, subjects practiced multiple triads, wherein each triad contained three blocks: (1) a single-retrieval vocal block (i.e., only the vocal response was practiced), (2) a single-retrieval keypress block (i.e., only the keypress response was practiced), and (3) a dual-retrieval block (i.e., subjects had to perform both responses to cue word presentation). In each block, each color word cue was presented once, with a total of ten (Nino & Rickard 2003) or 14 (Strobach, Schubert, Pashler, & Rickard, 2014) cues. Subjects received the same instruction for each block, namely to react as quickly and accurately as possible.

To assess different patterns in dual-memory retrieval, the analyses of these experiments involved comparisons of the observed data to predicted data. For these analyses, the observed reaction time patterns were compared to the quantitative predictions of a sequential retrieval model. We refer to this model as the *efficient sequential (ES) retrieval model* (please see Appendix A of Strobach, Schubert, Pashler, & Rickard (2014) for a thorough mathematical depiction of the ES model). The model predicts reaction time patterns that should be observed, when a sequential response pattern is adapted by an individual (i.e., the execution of one task after the other). This model incorporates three main assumptions: (1) there are independent and sequential perceptual, retrieval, and motor stages of processing, (2) a bottleneck exists exclusively during the memory retrieval stage of processing, such that the retrieval stage can occur for only one task at a time (whereas that stage can run in parallel with both the perceptual and motor stage), and (3) coordination of the three processing stages during dual retrieval is maximally efficient (i.e., has no or negligible coordination or task control delays). The ES model can be defined as an approximate lower bound RT estimate for sequential processing at the bottleneck-like retrieval stage. This means that observed reaction time patterns which are either similar or above the predicted data can be referred to as sequential response patterns and are in line with the hypotheses of the model. However, if the observed reaction times are falling below the ES prediction, we can conclude that the lower bound is systematically and substantially violated by the empirical data. This, in turn, implies that the sequential retrieval stage hypothesis can be considered false.

In previous work, the presence of two different types of response patterns could be assessed by comparing their data to the ES prediction (Nino & Rickard, 2003; Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014). This revealed two categories of subjects: Nongrouper subjects and grouper subjects, categorized according to differences in their inter-response intervals[1] (IRIs) on dual-retrieval trials. An IRI refers to the difference between *RT1* (the latency between cue presentation and the first executed response) and *RT2* (latency between cue presentation and the second executed response). *Nongrouper subjects* displayed large inter-response intervals. Large IRIs of nongrouper subjects led to the assumption that these subjects executed each response as soon as it was retrieved. This implies the presence of a sequential response pattern. While nongrouper subjects first displayed observed RTs above the ES prediction, over extensive practice, it was shown that both RT1 and RT2 for these nongrouper subjects converged on the quantitative predictions of the ES model. RTs and ES predictions converged on nearly the same values—for both the means and distributions quantiles—following that practice. This result indicated that nongrouper subjects indeed exhibited a sequential retrieval stage execution.

Contrary to the nongrouper subjects, the *grouper subjects* were characterized by small IRIs and reflected the synchronized execution of both responses (almost) at the same time. RT2 values for grouper subjects were initially above the ES retrieval prediction, suggesting that those subjects retrieved the two responses sequentially at the outset of dual-retrieval practice (just as nongrouper subjects did). However, their RT2 values fell several hundred milliseconds below the ES prediction by the end of practice. These results violate the ES prediction and implicate a form of *learned retrieval parallelism* for that subset of subjects. Learned retrieval parallelism refers to the occurrence of a parallel response pattern after practice.

The phenomenon of learned retrieval parallelism motivated a search for mechanisms that facilitate the onset of parallel retrieval for grouper subjects. According to the *set-cue bottleneck model* (Fig. 1 and Appendix), three

[1] In this line of research, we distinguish between long (IRI > 300 ms) and short (IRI < 300 ms) IRIs to assess potential differences in response styles (Nino and Rickard, 2003).

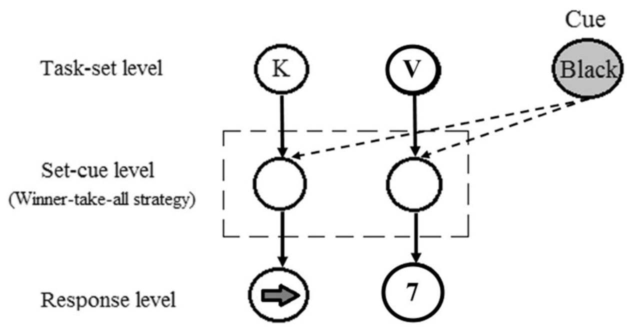


Fig. 1 Associated processing levels of the set-cue bottleneck model. The model presumes distinct depictions for the cue and the task set at the task-set level (i.e., either the keypress response (K) or the vocal response (V)). During learning, the set-cue level emerges which represents the connection of each cue and the associated response. This results in the connected response level for each cue–response pairing. The set-cue level incorporates the bottleneck, since only one node at the set-cue level can influence performance at time in a dual-retrieval condition. Individuals thus have to complete both responses sequentially

sequential, independent and additive processing stages exist: perceptual, central, and motor processing. Associations are formed only between activated nodes at neighboring levels of the hierarchy. At the input level, nodes represent the cue and the currently active task set(s). A task set refers to the goal to execute a response that is connected with a cue. When a set-cue node is selected, activation begins to flow from (only) one of the set-cue nodes to the response level. According to Nino & Rickard, (2003), activation streams can run in parallel from the earliest level of cue perception to the set-cue level. However, from the set-cue level to the response and motor levels, there is a winner-take-all competition at the set-cue level. In terms of this model, grouper subjects learned to chunk two responses independently for each cue, when both responses are concurrently active in working memory (please refer to Appendix for an in depth description of this process). This cue-level chunking process could enable retrieval of both responses in one pass through the same retrieval bottleneck that appears to govern performance at the beginning and throughout practice (Nino & Rickard, 2003). Taken together, this *cue-level account* implies that learned parallelism is only specific to practiced cues. In contrast, another possibility is that grouper subjects undergo a strategic and global switch from sequential to parallel retrieval following modest dual-retrieval practice (Meyer & Kieras, 1997; Oberauer & Bialkova, 2011; Oberauer & Kliegl, 2004). In this case, learned parallelism would follow a *task-level account* which implies that parallelism could occur for the task as a whole and would not be limited to specific cues. Previous examinations seemed to be in favor of the cue-level account and demonstrated that learned retrieval parallelism is a cue-level phenomenon and also indicated the presence of a cue-level chunking account in the context of the set-cue bottleneck model (Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014).

Response patterns: three factors of influence

In line with the findings discussed in the preceding section, the specific underlying mechanisms that account for either pattern (i.e., sequential in nongrouper subjects vs. parallel in grouper subjects) pose an important focus for dual-retrieval research. Why individuals reflect either of these response patterns could be accounted for by three different possibilities: (1) task-specific factors, (2) individual factors (personality, functional, and cognitive abilities) or (3) explicit task instructions.

Concerning task-specific factors, there might be certain task elements, such as low or high task demands, that drive the engagement in one of the (parallel versus sequential) strategies. For example, high task interference might not allow parallel execution (Ruthruff, Pashler, & Klaassen, 2001) and instead cause sequential task execution. In contrast, specific tasks could also facilitate the use of parallel response execution. A study by Orscheschek, Strobach, Schubert, & Rickard (2018) used a modification of the dual-retrieval task of Nino & Rickard (2003) and Strobach, Schubert, Pashler, & Rickard, (2014) to examine a new dual-retrieval context. Whereas previous results showed a distinction between grouper and nongrouper subjects (Nino & Rickard, 2003; Strobach, Schubert, Pashler, & Rickard, 2014) for two retrievals from episodic memory (i.e., the vocal and the keypress tasks), that distinction was not present for a dual-retrieval task that involved one retrieval from episodic memory and one retrieval from highly overlearned, and therefore automatized, semantic memory (Orscheschek, Strobach, Schubert, & Rickard, 2018). In line with other studies (Hazeltine, Teague, & Ivry, 2002; Strobach, Frensch, Mueller, & Schubert, 2012a, b), this study exhibited a preference for more parallel response strategies as well as evidence for learned retrieval parallelism across moderate to extensive practice.

Otherwise, there might be individual factors that could account for the adaption of either a nongrouping or grouping strategy. Individuals might have different functional abilities that could allow them to use parallel strategies, whereas other individuals might not have these abilities (e.g., increased vs. decreased working memory capacity; Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002). According to this view, past research suggested that there might be strategic choices in dual-tasking behavior. For example, individuals might have a preference for one specific strategy (Brüning & Manzey, 2018; Fischer & Plessow, 2015; Jansen, van Egmond, & de Ridder, 2016; Reissland & Manzey, 2016) that leads to an avoidance of other strategies. Further, individuals seem to prefer one task more than the other one, and it can by hypothesized that those preferences could possibly be determined by personality (Sanbonmatsu, Strayer, Medeiros-Ward, & Watson, 2013) or cognitive style (e.g., polychronicity vs. monochronicity; Ishizaka, Marshall, & Conte, 2001; Schell & Conte, 2008; Sternberg, Zhang, & Rayner, 2011).

However, it is still not entirely clear if the adaption of either response pattern is formed by (1) fundamental individual differences that only allow the use of one response pattern (i.e.; due to potential capacity limitations; Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002), or (2) by a strategy preference (e.g., response style) that can still be mediated by factors such as task demands and other factors (Jansen, van Egmond, & de Ridder, 2016; Sanbonmatsu, Strayer, Medeiros-Ward, & Watson, 2013). The difference between both accounts is that the latter approach would allow for individual flexibility of response patterns depending on task demands and instructions, whereas the first approach would not allow flexibility of the response patterns. Therefore, if response patterns were solely caused by fundamental individual differences, specific task and response instructions should not be able to influence response patterns. Nevertheless, other studies also showed that explicit instructions, for example priority instructions, are able to create specific response patterns (Levy & Pashler, 2008; Ruthruff, Pashler, & Hazeltine, 2003). A priority instruction refers to the instruction to prioritize one of the tasks in a dual-task situation. This form of instruction is often used in the psychological refractory period (PRP) paradigm. In a PRP paradigm, dual-tasking is tested with a subset of two different tasks which results in the presentation of two stimuli (i.e., S1 and S2) that are separated by a stimulus onset asynchrony (SOA) and require two distinct reactions (i.e., R1 and R2) (Fischer & Plessow, 2015; Pashler, 1994; Ruthruff, Van Selst, Johnston, & Remington, 2006). Studies that used priority instructions (i.e., "perform task A first") indeed showed that such instructions are effective to evoke sequential response patterns (Pashler & Johnston, 1998). Additionally, a small number of PRP studies have tried to instruct both, sequential and parallel response patterns. In these PRP studies that directly investigated the differences of instruction dependent response patterns, subjects had to follow intermixed sets of instructions: A priority instruction to evoke sequential response patterns and an instruction to distribute their capacity evenly on both tasks to evoke a parallel strategy (Lehle & Hübner, 2009; Lehle, & Hübner, 2009). In sum, their results showed a significant effect of both instructions that worked in single session experiments (Lehle, & Hübner, 2009) as well as in practice experiments (Lehle & Hübner, 2009).

Thus, instructions seem to be very effective in evoking the presence of specific response patterns (Fischer & Plessow, 2015; Jansen, van Egmond, & de Ridder, 2016). However, it has to be noted that individuals might not always want nor do they always have the capacity to follow priority instructions (Levy & Pashler, 2008; Miller & Durst, 2014). Moreover, other studies that used an applied dual-task paradigm (i.e. high-speed driving task vs. auditory memory task; Jansen, van Egmond, & de Ridder, 2016) also showed that the individual preference to complete one specific task first can be overruled by explicit task instructions, but that it takes extensive practice of dual-task situations to overrule these preferences. This is especially interesting, since the amount of practice might play an important role for the functionality and efficiency of strategy instructions, and might have different effects across different dual-task settings.

To sum up, even though there is a body of research on response strategies in dual-task scenarios, we are still lacking some important consensus about the underlying mechanisms that affect the adoption of response patterns. Especially, the query why some people seem to exhibit parallel patterns while others do not still needs proper examinations. To try to provide a deeper understanding of these mechanisms, we are going to investigate the presence of response patterns with the help of instruction manipulations. Despite the use of priority instructions and the instructions experiments in PRP and other dual-task situations, we are not aware of any experiments up to now that have specifically tested the possibility to use instruction manipulations to initiate serial as well as parallel response strategies during dual-retrieval practice with two retrievals from a single cue. Therefore, we are going to focus on the exploration of two specific aims in this paper: (1) A test of the impact of explicit task instructions on response patterns and strategies in practiced dual-memory retrieval (Experiment 1 and 2), and (2) an assessment of the processing architecture of this retrieval (Experiment 2).

Experiment 1

In Experiment 1, we assessed whether learned parallelism could be influenced by explicit task instructions. The general procedure of this experiment was developed in analogy to Experiment 2 by Nino and Rickard (2003). That is, we used the same cue–response combinations as well as the same number of triads. This resulted in a total of 30 triads, separated into 25 practice triads and five transfer triads. Next to that, we used three different instruction conditions in the dual-retrieval practice phase, manipulated between subjects: A *neutral* instruction, a *synchronize* instruction, and an *immediate* instruction. In all conditions, subjects were instructed to respond as quickly and accurately as possible. The *neutral* instruction reflected the same procedure that

Table 1 English translations of the specific instructions that were provided in each instruction group throughout practice

Instruction group	Specific dual-task instruction
Neutral Instruction	Please react as fast and as accurate as possible
Synchronize instruction	Wait until you have retrieved both reactions. Please give both reactions at the same time Please react as fast and as accurate as possible
Immediate instruction	Give each reaction as soon as you retrieve it. Please give each reaction after the other Please react as fast and as accurate as possible

The neutral instruction was used in all groups across transfer

was applied in previous experiments, i.e., subjects received no specific instruction under the dual-retrieval condition (Nino and Rickard, 2003; Strobach, Schubert, Pashler, & Rickard, 2014). In contrast, in the *synchronize* instruction condition, subjects were asked to wait until both responses were retrieved, and then to synchronize their response execution. Opposite to that, subjects in the *immediate* instruction condition were required to give the first response as soon as they retrieved it and give the second response afterwards (please refer to Table 1 for the exact strategy instructions in each group).

If explicit instructions have an impact on response strategies, then we should observe an increased proportion of grouper subjects in the synchronize instruction condition, relative to the neutral condition, and an increased proportion of non-grouper subjects in the immediate response condition, relative to the neutral condition. That outcome will be identified by a pattern of short IRIs and the onset of learned parallel retrieval in the synchronize instruction condition, and by performance consistent with the ES model in the immediate instruction condition. All of these results would exhibit that strategy manipulations could influence grouping behavior by increasing the proportion of grouper subjects in the synchronize instruction condition in contrast to the immediate instruction condition. If these empirical results would indeed be reflected in the data, we could assume that the adaption of each response style would not exclusively be affected by fundamental individual differences.

Furthermore, we want to investigate whether the strategy is robust or whether it flexibly adapts after giving a neutral strategy instruction at the end of practice (i.e., in the last five transfer triads). According to prior research, it can be hypothesized that individuals are likely to show strategy robustness even during the final neutral strategy instruction. This assumption is based on several findings from research on basic learning, cognition and memory principles (Dembo & Seli, 2004; Jansen, van Egmond, & de Ridder, 2016; Kramer, Larish, & Strayer, 1995). For example, the voluntary use of a new strategy after extensive practice of another one is unlikely, since individuals may need an opportunity to practice the new strategy to make it competitive with the old strategy (Dembo & Seli, 2004). Further, there might be some form of generalizable learned task coordination skills, which makes the adaption to a new strategy difficult, once another specific strategy has been extensively learned (Kramer, Larish, & Strayer, 1995). Finally, there may be little or no incentive to change strategies (Jansen, van Egmond, & de Ridder, 2016). In contrast, there could still be a strong impact of individual strategy preferences that might allow for a strategy change during the transfer triads (Brüning & Manzey, 2018; Fischer & Plessow, 2015; Jansen, van Egmond, & de Ridder, 2016; Reissland & Manzey, 2016). Since individuals might have been forced to adapt to a strategy that is different to their personally preferred strategy, they might switch their response strategy according to their personal preference during transfer.

Methods

Subjects

The experiment included a total of 72 subjects which were randomly divided across 3 different instruction groups (neutral instruction, synchronize instruction and immediate instruction). This resulted in 24 subjects in each group. All of the subjects were undergraduate students at the Medical School Hamburg, Germany. The total sample had a mean age of 23.3 years with a range from 18 to 30 years and N = 54 females. All of the subjects had normal or corrected to normal vision and were right handed. They either received credit points or a voucher with a monetary worth of 24 Euros; six took the voucher as compensation, whereas the remaining subjects used credit point compensation. All of the subjects were naïve to the research aim of the study. All subjects gave written informed consent prior to their inclusion and participation in the experiment.

An array of different cognitive assessments was used to investigate the possibility of individual differences between subjects. Each subject had to complete the D2 Test of Attention (Brickenkamp & Zillmer, 1998), three versions of the Digit-Span Test (Forward-Span, Backward-Span, Sequential-Span) as well as the Digit-Symbol Test taken from the Wechsler Adult Intelligence Scale (WAIS-IV; Wechsler, 2008), and a Vocabulary Test (WST; Schmidt & Metzler, 1992). Additionally, personality was assessed with the NEO-FFI multidimensional personality assessment (Costa & McCrae, 1989). These cognitive assessments were administered at the end of the computerized testing sessions

Table 2 Cognitive assessment outcomes for each instruction group

Assessment	Descriptives						ANOVA			
	Neutral instruction		Synchronize instruction		Immediate instruction		*df*	Mean square	*F*	*p*
	M	SD	*M*	SD	*M*	SD				
NEO-FFI (N)	20.8	8.04	21.7	8.1	19.1	7.5	2	29.49	.47	.63
NEO-FFI (E)	29.7	6.6	25.4	6.6	29.8	6.8	2	111.34	2.48	.09
NEO-FFI (O)	33.4	5.8	33.5	6.5	31.8	6.3	2	14.88	.38	.68
NEO-FFI (A)	33.8	6.1	34.4	4.6	34.3	4.4	2	1.95	.07	.93
NEO-FFI (C)	32.4	7.2	32.5	7.4	33.7	7.5	2	8.66	.16	.85
D2	210.8	26.3	242.8	102.2	198.2	34.8	2	9825.94	2.47	.09
Vocabulary Test	30.9	3.0	30.9	3.9	30	3.4	2	5.93	.49	.61
Digit-span (F)	9.8	1.8	9.6	2.0	10.2	1.8	2	1.68	.47	.63
Digit-span (B)	9	2.1	8.8	1.9	8.4	1.6	2	1.54	.42	.66
Digit-span (S)	8.7	2.1	8.9	2.1	9.2	2.1	2	.77	.17	.84
Digit-symbol	78.5	11.5	78.9	10.2	77.9	12.3	2	4.81	.04	.96

NEO-FFI: N = neuroticism, E = extraversion, O = openness, A = agreeableness, C = conscientiousness; D2 test = concentration performance; Digit-span F = forward, B = backward, S = sequential

Table 3 Overview of the general procedure in Experiment 1 and 2

Session	Duration	Experimental phases	Number of blocks/triads
1	One hour	Study phase Single-retrieval criterion phase Single-retrieval practice phase (Part 1)	10 blocks each task 5 blocks each task
2	One hour	Single-retrieval practice phase (Part 2) Single–dual practice phase (Part 1)	5 blocks each task 15 triads
3	One hour	Single–dual practice phase (Part 2) Single–dual transfer phase	10 triads 5 triads

A triad incorporates one block of each task: keypress task, vocal task, dual-retrieval task

(Session 1: D2 Test, NEO-FFI, Session 2: WST Test, Session 3: Digit-Span and Digit-Symbol Test) (Table 2).

Apparatus and cues

The experimental software package E-Prime software (Psychology Software Tools, Pittsburgh, PA) was used on IBM-compatible personal computers to assess each subject in an individual, sound-isolated lab cabin. Keypress as well as vocal responses were recorded with the E-Prime accompanying voice-key apparatus (Model 200A). This voice-key apparatus incorporates a serial response (SR) box which makes use of a 0-ms debounce period for keypress reactions (Psychology Software Tools, Pittsburgh, PA). The SR box automatically assessed keypress RT and coded the reactions for their accuracy. Additionally, a microphone was connected to the SR box to assess vocal responses. A list of the applied 10 cue words and the corresponding responses in the vocal and the keypress task are presented in Table 4. All of the cues were presented on a 19" CRT monitor which corresponds to a visual angel of 51.5° from a viewing distance of 50 cm. The cue words subtended up to 7 cm and letter

Table 4 Cue–response mappings in Experiment 1 and 2

Condition 1			Condition 2		
Color words	Vocal-digit response	Keypress response	Color words	Vocal-digit response	Keypress response
Red	5	←	Red	8	←
Green	4	←	Green	5	→
Blue	1	→	Blue	6	→
Yellow	3	→	Yellow	4	→
Purple	2	←	Purple	0	→
Brown	6	→	Brown	9	←
Black	7	←	Black	2	←
Orange	8	←	Orange	1	←
White	9	→	White	3	←
Pink	0	→	Pink	7	→
Gold	12	←	Gold	11	→
Silver	11	←	Silver	14	←
Gray	14	→	Gray	13	→
Olive	13	←	Olive	12	←

Experiment 1 used the first ten cues (red to pink). Experiment 2 used all 14 cues with a distinction between old cues (white) and new cues (gray) height was 1.7 cm. This resembles a visual angel of 8.0° and 1.9°, respectively. All of the cue words were displayed in black color on a white background.

Procedure and design

An overview of the exact study design can be found in Table 3. The experiment took place over a total of 3 sessions that had to be completed in a timeframe of 7 days. Each of the sessions lasted around 1 h. In the first session (Session 1), subjects were introduced to the 2 different retrieval tasks in the *study phase*. In each instruction group, one half of the subjects were introduced to the keypress task first, while the other group started with an introduction to the vocal task. This pattern was balanced across all instruction groups.

For the keypress task, subjects were instructed to memorize all of the 10 color cue words and the associated manual keypress responses (Table 4). Half of the subjects were presented with the cue–response mapping of condition 1 and the other half with the mappings of condition 2. The direction of the keypress response was indicated by the presentation of an arrow pointing in the associated keypress direction (i.e., either to the right or to the left). Subjects were specifically instructed to memorize the cue–response combination and to press the associated key on the keypress pad. First, the cue–arrow combination was presented for 5000 ms, followed by the appearance of a blank interval of 1000 ms and a fixation cross for 500 ms. In this situation, subjects were only required to focus on the cue–response combination to memorize it. Second, only the cue was presented and subjects were required to press the associated key. This cue presentation lasted until the subjects responded to the cue. During this *study phase*, each of the cue–response combinations was presented once and in randomized order across 2 study blocks.

Following the study blocks, subjects had to perform the *single-retrieval criterion phase*. In this phase, subjects were solely presented with the cue words and were required to retrieve the previously learned keypress direction and to press the associated key (e.g., either right or left). In this phase, a blank screen was presented for 1000 ms in each trial (e.g., each cue word) followed by the fixation cross for 500 ms, which was, in turn, followed by the cue word presentation. The cue word only disappeared from the screen when a response was given on the keypress pad. Subjects received a feedback response on screen after each incorrect trial, which displayed the correct response. The correct response was indicated by a right- or leftwards-pointing arrow and the whole feedback screen was presented for 2500 ms. The complete *single-retrieval criterion phase* incorporated 10 single-keypress blocks and each cue–response combination was practiced once per block in randomized order (e.g., 10 repetitions of all 10 trials per block).

For the vocal task, the procedure was similar to the keypress task. In the *study phase*, subjects were again required to memorize each cue–response combination, with half of the subjects in condition 1 and the other half in condition 2 (Table 4). Here, the correct response was indicated by displaying the correct digit number response together with the cue word. However, instead of giving a keypress response, subjects were required to speak the associated digit number into a microphone that was installed in front of them. In the *single-retrieval criterion phase*, the experimenter had to code the response accuracy of the vocal response. This required an additional interval of 2500 ms, which was included after the execution of the vocal response. The same feedback process that was used for the keypress task was applied for the vocal task as well.

At the end of session 1, subjects were given an additional practice. In the *single-retrieval practice phase*, they were given 5 blocks of the vocal and the keypress task, which resulted in a total of 10 blocks. The blocks were presented in alternating order, starting with the task that was introduced first in the study phase.

Session 2 started with the same additional *single-retrieval practice phase* of 5 blocks per task. This allows the subjects to re-familiarize with the task and prepares them for the upcoming phase. It further increases the response accuracy and reduces single-task reaction times (Strobach, Schubert, Pashler, & Rickard, 2014). After that, the *single–dual practice phase* started. This phase was formed by the practice of 25 triads (Practice Triads 1–25); however, only triads 1–15 were executed in Session 2. A triad is constituted by a set of 3 blocks: the keypress task (single-retrieval block), the vocal task (single-retrieval block) and the dual-retrieval task (dual-retrieval block). Single-retrieval trials in the single-retrieval blocks were identical to the *single-retrieval practice phase*. The dual-retrieval trials were similar to single-task trials; however, when the cue word was presented, subjects had to give both responses in one trial. Here, the cue word remained on the screen until both responses were executed.

As mentioned before and contrary to previous experiments, we introduced 3 different instruction manipulations in the dual-retrieval block (see Table 1): a *neutral* instruction, a *synchronize* instruction and an *immediate* instruction. Subjects in the *neutral* instruction group received no specific instruction on response strategy, whereas subjects in the *synchronize* instruction group had to synchronize their responses and subjects of the *immediate* instruction group had to follow a sequential response strategy. None of these strategy instructions involved a specific response order. The specific instructions were read to the subjects before the start of each dual-retrieval block.

The tasks in the *single–dual practice phase* were presented in 3 different orders: one third of the subjects in

Table 5 Error rates (in percent) from the start (Practice Triad 1) to the end of practice (Practice Triad 25) in Experiment 1

Instruction group	Keypress single retrieval		Vocal single retrieval		Keypress dual retrieval		Vocal dual retrieval		First dual-retrieval response		Second dual-retrieval response	
Practice triad	1	25	1	25	1	25	1	25	1	25	1	25
Neutral instruction	6.3	1.2	5.1	0.4	6.2	2.0	6.7	1.3	5.4	1.3	7.5	2.1
Synchronize instruction	3.9	0.8	4.3	0.8	7.5	2.0	7.0	2.5	7.9	1.2	6.6	3.8
Immediate instruction	7.9	1.2	3.6	0.8	8.3	0.8	2.9	0.8	4.1	1.2	7.0	1.2

First and second dual retrieval refer to the averaged first and second response in the dual-retrieval blocks, independent of the type of response that came first (i.e., keypress or vocal)

each instruction group performed the keypress task in the first single-retrieval block, the vocal task in the second single-retrieval block and the dual retrievals in the third block. Another third performed the dual-retrieval task first, then the keypress single-retrieval and lastly the vocal single retrieval. The last third performed the vocal single retrieval first, then the dual retrieval, and finished each triad with the keypress single retrieval.

Session 3 started with the rest of the *single–dual practice phase* (Triads 16–25). After that, subjects had to perform a *single–dual transfer phase* that involved 5 additional triads (Transfer triads 1–5). In this transfer phase, subjects in each of the 3 different instruction groups received the neutral instruction. This means that each subject was only instructed to solve the dual-retrieval task as fast and as accurately as possible.

Statistical approach

In all analyses across all groups, we excluded trials with RTs below 200 ms. To ensure an elaborated statistical approach to our data at the end of the practice phase (triads 21–25) and during transfer, we combined classical null hypothesis testing (NHST) with additional computations of Bayes' factors. The motivation to use Bayesian statistics is rooted in the nature of the approach, which is able to eliminate some of the concerns associated with NHST methods. Bayesian statistics provide an asset, since they can account for unbiased statistical information and do not need approximation assumptions for the homogeneity of variance (e.g., Kruschke, 2013). Further, it aids the assessment and creation of informative inferences since it is able to report combined probabilities of associated parameter values. The NHST analyses were conducted using IBM Statistical Package for Social Sciences (SPSS) version 22. To assess Bayes factors, we used JASP (JASP Version 0.8.6; JASP Team, 2018). JASP is making use of Bayesian analyses that are partly administered from Morey and Roudes Bayes Factor package in R.

Throughout our analyses, we are providing Bayes factors (BF_{10}) for paired sample *t* tests. A Bayes factor (BF_{10}) refers to the likelihood ratio of the data comparing the alternative hypothesis against the null hypothesis. Larger Bayes factors ($BF_{10} > 1$) are considered supportive of the alternative hypothesis (Kruschke, 2013). As a prior, we used the informed Cauchy prior (0.707), since we were not able to obtain effect size estimations from previous experiments that used the same methodology. The informed Cauchy prior is a wide prior, with a similar distribution as the t-distribution. For our analyses, we centered the prior on zero. In general, centering on zero can be classified as a standard practice in Bayes statistics (Hoijtink, Klugkist, & Boelen, 2008).

Results

Accuracy results

For a dual-retrieval trial to be categorized as correct, both responses had to be correct. Descriptively, error rates decreased across all instruction groups from the beginning to the end of practice. The decrease was evident for single-retrieval trials, as well as for dual-retrieval trials. Table 5 shows a detailed overview of the error rates across all instruction groups and all retrieval conditions.

RT results

Figure 2 displays the RTs averaged over all subjects in the correctly preformed single- and dual-retrieval trials (i.e., single-retrieval trials: keypress task, vocal task; dual-retrieval trials: RT1, RT2) for each of the instruction groups. There was a decrease in RTs from Triad 1 to Triad 25 and mainly for RT1 and RT2, with a slowing of response times at the beginning of Session 3 (Practice Triad 16). Additionally, there was an increase in RTs at the beginning of the transfer phase in the synchronize and immediate instruction groups.

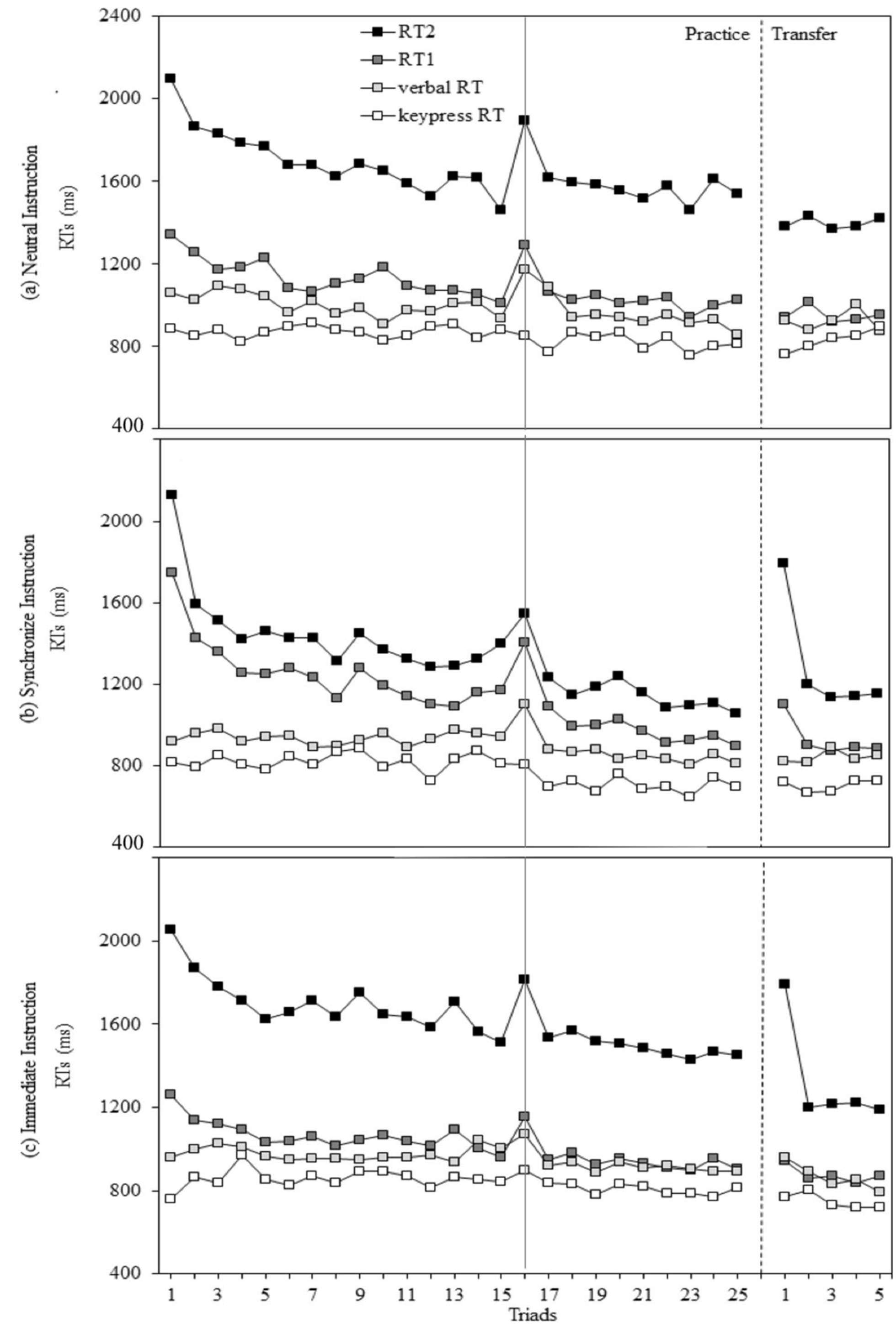


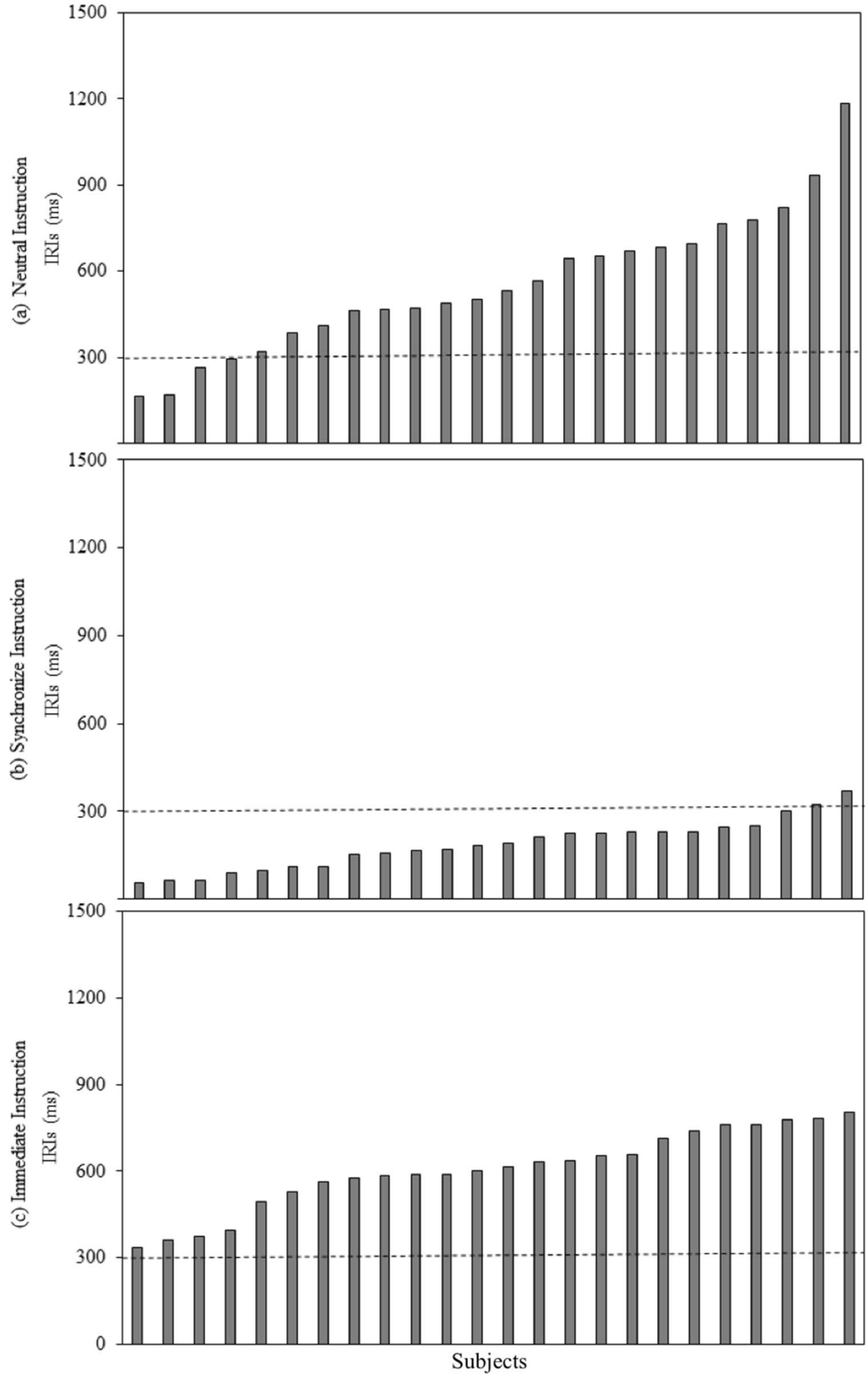
Fig. 2 Observed reaction times (RTs) in single-retrieval blocks of the keypress task and the vocal task as well as observed RTs in dual-retrieval blocks (i.e., RT1 and RT2) in the overall dataset during the 25 practice triads and 5 transfer triads in the neutral instruction group (**a**), synchronize instruction group (**b**), and immediate instruction group (**c**) of Experiment 1

IRI analysis

According to, and following, previous analyses by Nino and Rickard (2003) and Strobach, Schubert, Pashler, & Rickard, (2014), individual mean IRIs were computed across all practice-phase dual-retrieval trials for each subject. Figure 3 shows these results for each instruction group, rank ordered from short to long IRIs. We computed independent sample *t* tests to compare the IRIs in each instruction group.[2] The comparison of IRIs across the neutral and the synchronize instruction condition revealed a significant difference between both groups ($t(46) = 7.019$, $p < .001$) as well as extreme evidence for H1

[2] Bonferroni correction was applied to ensure alpha-correction for the three relevant analyses (alpha = .017).

Fig. 3 Inter-response intervals (IRIs) across the practice phase of individual subjects in the neutral instruction group (**a**), synchronize instruction group (**b**), and in the immediate instruction group (**c**) of Experiment 1



($BF_{10} = 1.017e+6$). The same result was observed for the comparison of IRIs in the synchronize and immediate instruction groups ($t(46) = -12.736$, $p < .001$; $BF_{10} = 3.100e+13$). In contrast, IRIs across the neutral instruction group versus the immediate instruction group revealed no significant difference between groups ($t(46) = -.873$, $p > .1$) and moderate evidence for H0 ($BF_{10} = 0.392$).

Since the previous analyses by Strobach, Schubert, Pashler, & Rickard (2014) and Nino & Rickard (2003) indicated that IRIs of below and above 300 ms mark distinct response patterns, subjects with an IRI below 300 ms were previously identified as *grouper subjects*, whereas subjects with an IRI above 300 ms were identified as *nongrouper subjects*. We applied the 300-ms distinction to our data and used it as a cut-off value to analyze the two sets of subjects. Note that we do not interpret the 300-ms distinction as a theoretical-driven criterion to divide both groups. However, we use it as an empirically driven criterion to divide two distinct groups of response patterns. Based on the IRI values, we categorized four subjects in the *neutral* instruction group (Fig. 3a) as groupers ($M = 221$ ms, $SD = 65$ ms) and 20 subjects as nongroupers ($M = 621$ ms, $SD = 207$ ms). In the *synchronize* instruction group (Fig. 3, Panel b), it can be observed that 22 subjects had mean IRI values of below 300 ms ($M = 166$ ms, $SD = 71$ ms). The remaining two subjects had mean IRIs above 300 ms ($M = 345$ ms, $SD = 32$ ms). Since all of the subjects in this subgroup were instructed to synchronize their responses, these results show that the synchronize instruction worked for most of the subjects. In contrast, IRI values across practice in the *immediate* instruction group (Fig. 3, Panel c) were all above the 300-ms cut-off value ($M = 605$, $SD = 138$). Therefore, all of these subjects are classified as nongroupers. As for the synchronize instruction group, these results show that our instruction to use a sequential response style worked as well. To assess if grouper and nongrouper subjects display coherent and equally systematic response patterns in the instructed groups (i.e., synchronize and immediate instruction groups) as subjects in the self-chosen strategy group (i.e., neutral instruction), two additional assessments were made: (1) A comparison between the IRIs of grouper subjects in the neutral and in the synchronize group indicated no difference between the IRIs of both grouper populations ($t(23) = 1.582$, $p = .127$). (2) A comparison of nongrouper subjects in the neutral and in the immediate instruction group revealed no difference between nongrouper subjects in both groups ($t(42) = 0.319$, $p = .751$). These results indicate that both subject groups reflect coherent response patterns and that they follow an equally systematic response pattern with self-chosen as with instructed strategies.

A Fischer´s exact test (FET) to compare the number of grouper subjects in the neutral instruction condition ($N = 4$) and in the synchronize instruction condition ($N = 22$) revealed a significant difference between both groups ($p < .001$, FET). The same results could be observed for the comparison of grouper subjects in the synchronize instruction condition ($N = 22$) versus the immediate instruction condition ($N = 0$) ($p < .001$, FET). An additional Fischer´s exact test to compare the number of nongrouper subjects in the neutral instruction condition ($N = 20$) and the immediate instruction condition ($N = 24$) did not indicate a significant difference ($p > .5$). These results are in line with our predictions that strategy manipulations will increase the proportion of grouper subjects under synchronize instructions, in comparison to the other instruction groups.

IRIs across the transfer phase are shown in Fig. 4. In the *neutral* instruction group (Fig. 4a), the pattern is similar to Fig. 3a. However, we observe generally shorter IRIs which are assumed to be caused by extensive practice which facilitated faster response patterns in general, as illustrated in Fig. 2 (i.e., RTs decreased over the course of training). In the transfer phase, grouper subjects ($N = 7$) had a mean IRI of 159 ms and a standard deviation of 48 ms. For nongrouper subjects ($N = 17$), those values were 561 ms and 158 ms, respectively. Figure 4b illustrates the *synchronize* instruction IRIs in the transfer phase. These results are of major interest, since subjects were presented with a neutral instruction during the transfer phase. The IRIs in Fig. 4b are no longer only reflecting short IRIs below 300 ms; instead, we can observe a wide range of short and long IRIs. The results show a subgroup of subjects ($N = 15$) with IRIs below 300 ms, with a mean IRI value of 162 ms ($SD = 88$ ms), as well as a subgroup of subjects ($N = 9$) with IRIs above 300 ms ($M = 633$ ms, $SD = 360$ ms). These IRI values indicate that the strategy instructions worked in the practice phase, but that they do not seem to be completely robust after practice when subjects are allowed to choose any strategy during transfer. Similar results occurred for the nongroupers in the *immediate* instruction group ($N = 16$) displayed in Fig. 4c ($M = 450$ ms, $SD = 310$ ms). Eight subjects displayed an IRI below 300 ms ($M = 189$ ms, $SD = 32$ ms), which indicates that they might have changed to a grouping strategy during the transfer phase. These results also suggest that the instructed strategies are not completely robust when subjects are no longer required to use them (see Fig. 4).

Practice-phase dual-retrieval RTs and ES model predictions

Mean RT data for both dual retrievals and the ES prediction for the *nongrouper subjects* across the neutral instruction and immediate response instruction conditions are shown in Fig. 5. If RT2 is either similar or above the predicted ES data, we observe sequential response patterns that are in line with the hypotheses of the model. When RT2 is falling below the ES prediction, we can conclude that the lower bound is systematically and substantially violated by the empirical data. This, in turn, implies that the sequential retrieval stage hypothesis can be considered false.

In the *neutral* instruction, RT1 was above the ES prediction across the whole practice phase (Triads 1–25; Fig. 5a). RT2 remained above the ES prediction until triad 6, after that RT2 remained essentially equivalent to the ES prediction. However, RT2 was significantly above the ES prediction on triad 25 ($t(19) = 2.527$, $p < .05$). Bayes statistics displayed that the nongrouper subjects showed moderate evidence for H0 in triads 21–24 ($BF_{10} > 0.244$). Further, Triad 25 showed anecdotal evidence for H1 ($BF_{10} = 2.818$)

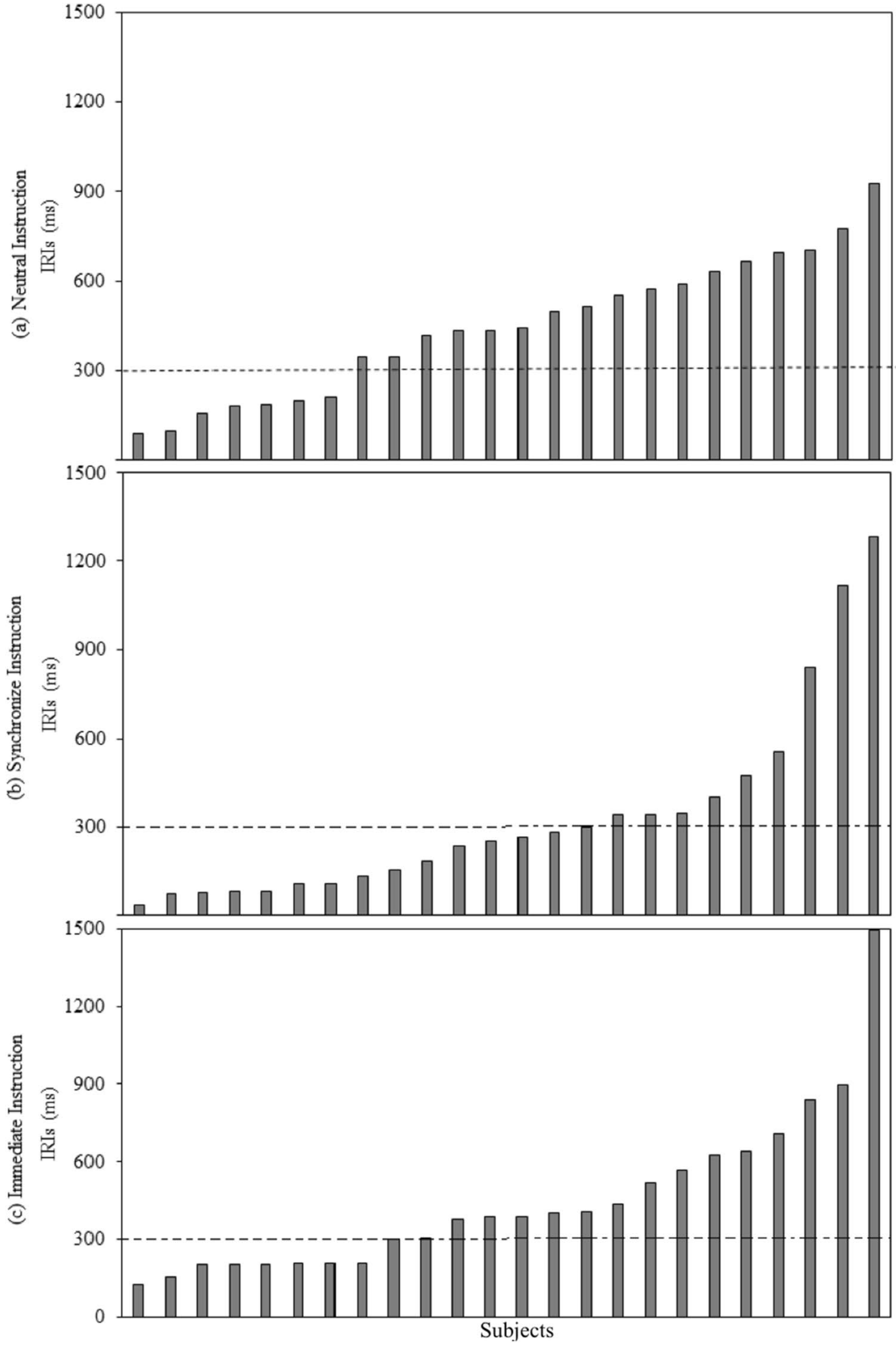


Fig. 4 Inter-response intervals (IRIs) across the transfer phase of individual subjects in the neutral instruction group (**a**), synchronize instruction group (**b**), and in the immediate instruction group (**c**) of Experiment 1

which further suggests that RT2 was significantly above the ES prediction on this triad. Taken together, these results match the assumption that nongrouper subjects are characterized by a sequential response pattern. In the synchronize instruction condition, only two subjects had IRIs longer than 300 ms which would classify them as nongroupers. In turn, no analyses of practice-phase data for those subjects were performed.

All subjects in the *immediat*e instruction group were classified as nongroupers. As illustrated in Fig. 5b, RT1 remained above the ES prediction across the whole practice phase. Further, RT2 averaged over all practice triads was equal to the ES prediction throughout the whole practice phase ($t(23) = 1.561$, $p > .1$; Fig. 5b). Further, Bayes factor analyses showed anecdotal evidence for H0 averaged over all practice triads ($BF_{10} = 0.621$). However, the evidence for

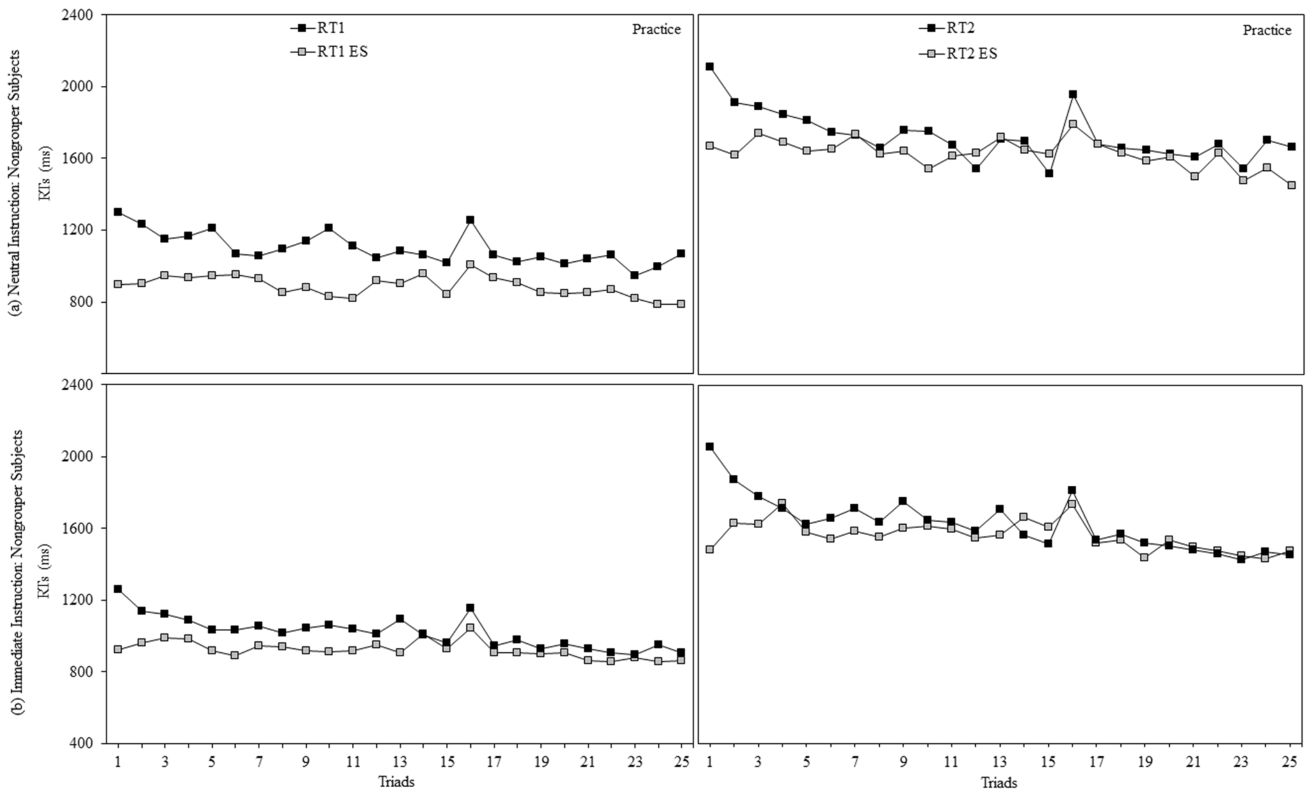


Fig. 5 Observed reaction times (i.e., RT1, RT2) as well the predictions of the efficient sequential (ES) model, sorted by the response style in the practice phase, during the single–dual practice for nongrouper subjects in the neutral instruction group (**a**), and in the immediate instruction group (**b**) of Experiment 1

H0 increased to a moderate level for RT2 averaged over the last nine practice triads ($BF_{10} = 0.216$) which indicates that subjects followed a sequential response style that was still mirroring the ES prediction across extensive practice.

Figure 6 shows the dual-retrieval data for RT1, RT2, and the ES predictions for the subgroups of *grouper subjects*. As can be observed in Fig. 6a, for grouper subjects in the *neutral* instruction condition, RT1 remained above the ES prediction, whereas RT2 was falling below the ES prediction from triad 2 on and remained below the ES prediction averaged throughout practice ($t(3) < -3.586$, $p < .05$). Averaged across practice, these grouper subjects showed moderate ($BF_{10} = 2.914$) evidence for H1. Averaged across the last nine practice triads (i.e., after substantial practice), these subjects displayed strong evidence for H1 ($BF_{10} = 9.093$). Taken together, this supports our previous assumptions that these grouper subjects' RTs significantly violate the ES prediction, showing evidence for learned parallelism for these subjects (Nino & Rickard, 2003; Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014). Figure 6b illustrates grouper subjects under the *synchronize* instruction condition. One subject had to be excluded from further analyses, due to severely increased RT values on 17 triads of RT1 and 15 triads of RT2 in the dual-retrieval condition, as well as on 4 triads in the keypress single task and 2 triads in the vocal single task. Across these triads, the mean of this subject was in between 2.8 and 3.6 standard deviations above the general mean, which would have accounted for confounded results, leaving 21 grouper subjects in this subgroup. RT1 remained above the ES prediction across the whole practice phase. In contrast, analyses showed that RT2 averaged across practice fell significantly below the ES prediction ($t(20) < -3.341$, $p < .01$) which indicates learned parallelism for these grouper subjects. Further, these subjects showed strong evidence for H1 averaged across the whole practice phase ($BF_{10} = 13.039$), as well as extreme evidence for H1 averaged across the last nine practice triads ($BF_{10} = 113.969$).

These results are supportive of our previous conclusions that the instruction to synchronize responses leads to learned parallelism. Finally, no grouper subjects were present throughout practice in the *nongrouping* instruction condition. In total, this implies that synchronized and parallel execution of two retrievals can be demonstrated

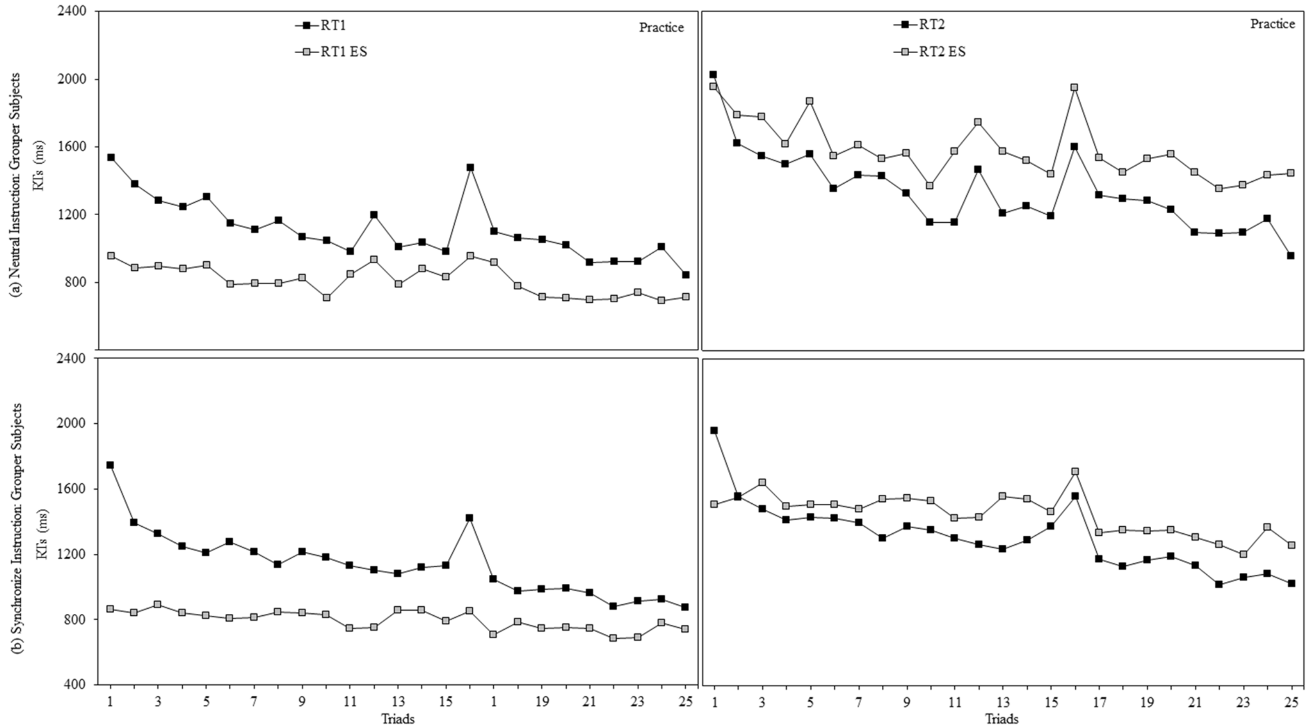


Fig. 6 Observed reaction times (i.e., RT1, Panel a and RT2, Panel b) as well as the predictions of the efficient sequential (ES) model, sorted by the response style in the practice phase, for grouper subjects in the neutral instruction group (**a**), and in the synchronize instruction group (**b**) of Experiment 1

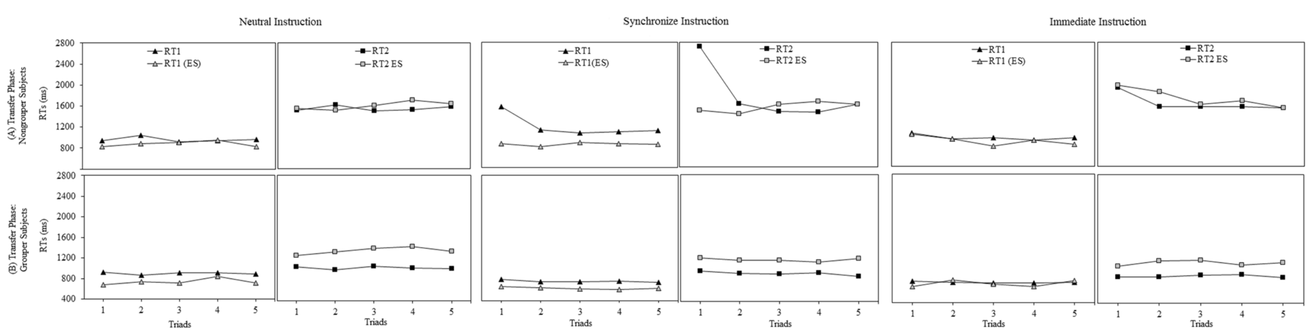


Fig. 7 Observed reaction times (i.e., RT1 and RT2) as well as the predictions of the efficient sequential (ES) model, sorted by the response style in the neutrally instructed transfer phase (**a**: nongrouper subjects; **b**: grouper subjects) of Experiment 1

for nearly all subjects under the instructed synchronize condition.

Transfer-phase dual-retrieval RTs

The transfer phase is of special interest for our analyses, since all subjects across all instruction groups received the same neutral strategy instruction and showed a number of strategy switches. To analyze these data, the response patterns across the transfer phase need to be assessed. An overview of RT1, RT2 and the associated ES prediction can be found in Fig. 7. Data of *nongrouper subjects* in the *neutral instruction* group (Fig. 7a) indicated that RT1 remained above or equal to the ES prediction; averaged RT2 values for these nongrouper subjects remained equal to the ES prediction across transfer ($t(16) = -1.138$, ps > .1). These subjects further showed moderate evidence for H0 averaged across transfer ($BF_{10} < 0.435$). In the *synchronize instruction* group, RT1 remained above the ES prediction. A *t* test of these nongrouper subjects indicated that RT2 across transfer was significantly below the ES prediction on practice triad 25 ($t(8) = -3.467$, $p < .01$), but above the ES prediction on the first transfer triad ($t(8) = 3.797$, $p < .01$). Averaged RT2 throughout the rest of transfer (transfer triads 2–5) remained significantly below the ES prediction ($t(8) < 3.519$, $p < .01$). These results are further supported by strong evidence for H1 averaged across transfer ($BF_{10} = 12.390$). These findings indicate a shift to sequential retrieval processing for the first transfer triad after practice for a subset of nine subjects. In the *immediate instruction* group, RT1 for *nongrouper subjects* was almost equal to the ES prediction. A *t* test of the averaged RT2 and the ES prediction across the nongrouper subjects in the transfer phase showed that nongrouper subjects displayed sequential retrieval throughout transfer ($t(15) = -0.051$, $p > .1$). We were further able to observe moderate evidence for H0 across the whole transfer phase ($BF_{10} = 0.256$) in this nongrouper subgroup.

Analyses of *grouper subjects* in the three subsamples across transfer (Fig. 7b) indicated that RT1 remained significantly above the ES prediction in the neutral and synchronize group. In contrast, averaged RT2 was significantly below the ES prediction throughout transfer across each instruction group (*neutral instruction group*: $t(6) = -9.583$, $p < .001$, $BF_{10} = 334.868$; *synchronize instruction group*: $t(13) = -6.578$, $p < .001$, $BF_{10} = 1320$; *immediate instruction* group: $t(7) = -3.104$, $p < .05$, $BF_{10} = 4.274$). These results support the assumption that the stable grouper subjects further exhibited learned parallelism and stuck with the grouping strategy. Results of the immediate instruction group are especially interesting, since the grouper subjects in the immediate instruction transfer phase showed sequential retrieval in the preceding practice phase (e.g., averaged RTs over Practice Triads 20–25: $t(7) = -0.450$, $p > .1$). According to previous findings (Kramer, Larish, & Strayer, 1995), individuals should not be able to exhibit forms of parallel response execution when they had no prior chance to practice this response strategy. However, the switch to a new strategy might be an indicator for an underlying personal preference that accounts for a switch to either grouping or nongrouping patterns.

Explorative analyses on strategy switches

As the previous analyses showed, we were able to observe a switch in response strategies from practice to transfer in several subjects in the synchronize instruction group as well as in the immediate instruction group. In the synchronize instruction group, fifteen subjects did not switch their grouping strategy during transfer, while eight subjects switched to a nongrouping strategy. Transfer-phase IRIs of individuals who switched from grouper subjects to nongrouper subjects ($M = 551$, $SD = 282$) increased the IRIs of those who did not switch ($M = 237$, $SD = 301$) their strategy. A *t* test showed that both groups differ significantly from each other ($t(21) = -2.427$, $p < .05$). Further, to assess potential differences on a practice-phase RT level, we computed the overall difference between the ES prediction and RT2 for each subject. A *t* test indicated no significant difference between the ES prediction and RT2 for switcher subjects ($M = 174$, $SD = 278$) versus strategy stable subjects ($M = 106$, $SD = 159$), ($t(21) = -.748$, $p > .5$).

For subjects in the immediate instruction group, we found sixteen subjects that remained nongrouper ($M = 580$, $SD = 303$) during transfer and eight subjects that switched to a grouping strategy ($M = 189$, $SD = 31$). Again, we compared the IRIs of both subsets of subjects during the transfer phase and found a significant difference in the IRI values ($t(22) = 3.599$, $p < .01$). The comparison of the overall difference between the ES prediction and RT2 in the practice-phase values showed no difference between switcher subjects ($M = 29$, $SD = 175$) to strategy stable subjects ($M = 86$, $SD = 230$), ($t(22) = -.616$, $p > .5$).

To assess if switcher subjects display a difference in the coherence and adaption of the response patterns in the transfer phase in comparison to their response patterns in the practice phase, we compared the IRIs of switcher and stable subjects across the last five practice triads with the five transfer triads. For the synchronize instruction group, switcher subjects exhibited different IRIs ($t(9) = -2.776$, $p < .05$) in contrast to stable subjects ($t(12) = -1.325$ $p > .1$). However, the IRIs of switcher subjects were not different from the IRIs of stable subjects across the last five practice triads ($t(21) = -1.220$ $p > .1$). In the immediate instruction group, we found the same results for the comparison of transfer and practice as in the synchronize instruction (switcher: $t(7) = 3.333$, $p < .05$; stable: $t(15) = 0.160$, $p > .5$). Nevertheless, we observed a difference between switcher and stable subjects across the last five practice triads ($t(22) = 2.260$ $p < .05$). Subjects that switched to a grouping strategy during transfer already exhibited shorter IRIs during the end of practice ($M = 423$ ms, SD=216) compared to stable subjects ($M = 596$ ms, $SD = 154$).

To assess for further differences between switcher and stable subjects, we looked into potential differences between these two subject groups across all of our additional cognitive assessments (D2 Test of Attention, Digit-Span Test, Digit-Symbol Test, WST, NEO-FFI). An overview of the mean scores for each test and subtest for each group can be found in Table 6. We did not observe any differences between switcher subjects and stable subjects in the synchronize instruction group (NEO-FFI N: $t(21) = -0.108$, $p = .915$; NEO-FFI E: $t(21) = -1.122$, $p = .279$; NEO-FFI O: $t(21) = -0.990$, $p = .338$; NEO-FFI A: $t(21) = -0.889$, $p = .388$; NEO-FFI C: $t(21) = -2.041$, $p = .059$; D2 Concentration: $t(21) = -1.095$, $p = .289$; Digit-Span Forward: $t(21) = 1.051$, $p = .308$; Digit-Span Backward: $t(21) = 0.126$, $p = .901$; Digit-Span Sequential: $t(21) = 0.138$, $p = .892$; Digit-Symbol Test: $t(21) = 0.284$, $p = .780$; WST: $t(21) = -.664$, $p = .516$). Likewise, these results were

Table 6 Cognitive assessment outcomes in each group

Strategy	Instruction	NEO-FFI										D2 test (CP)		Digit-span test						Digit-symbol test		Vocabulary test	
		N		E		O		A		C				Forward		Backward		Sequential					
		M	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD
Stable	Neutral	20.2	7.7	30.7	6.4	34.2	5.8	33.8	5.9	31.1	6.7	207.7	25.2	9.7	1.9	8.9	2.1	8.7	2.4	78.6	12.2	31.1	3.2
	Synchronize	21.5	7.1	24.1	6.2	32.4	7.5	33.7	5.2	30.1	7.9	223.3	47.7	10.0	2.4	8.8	2.2	9.0	2.1	79.4	11.6	30.5	4.5
	Immediate	19.3	9.8	30.4	8.3	30.6	6.3	34.4	4.7	36.7	7.9	206.0	26.5	10.2	1.6	8.8	1.5	9.8	1.4	78.8	9.8	29.7	3.7
Switcher	Neutral	24.7	10.7	24	5	29.0	4.6	34.0	8.5	40.3	4.9	231.3	29.5	10.3	2.1	9.7	2.5	9.0	1.1	78.3	8.1	30.0	2.6
	Synchronize	22.0	10.2	27.8	7.3	35.7	4.8	35.8	3.1	37.1	4.1	276.2	86.2	9.0	0.6	8.7	1.5	8.7	2.3	78.0	8.2	31.8	2.6
	Immediate	18.9	2.7	29.0	4.8	33.6	6.6	34.1	4.3	29.4	4.7	183.5	45.7	10.3	2.5	7.7	1.6	8.0	2.6	76.0	17.5	30.6	3.0

NEO-FFI: N = neuroticism, E = extraversion, O = openness, A = agreeableness, C = conscientiousness; D2 Test CP = concentration performance

mirrored for the comparison of switcher and stable subjects in the immediate instruction (NEO-FFI N: $t(21) = 0.115$, $p = .910$; NEO-FFI E: $t(21) = 0.407$, $p = .690$; NEO-FFI O: $t(21) = -0.931$, $p = .367$; NEO-FFI A: $t(21) = -0.131$, $p = .897$; NEO-FFI C: $t(21) = 2.160$, $p = .057$; D2 Concentration: $t(21) = 1.299$, $p = .213$; Digit-Span Forward: $t(21) = -0.174$, $p = .864$; Digit-Span Backward: $t(21) = 1.495$, $p = .154$; Digit-Span Sequential: $t(21) = 1.804$, $p = .090$; Digit-Symbol Test: $t(21) = 0.447$, $p = .661$; WST: $t(21) = -.543$, $p = .594$). Moreover, when we looked into a comparison of the cognitive assessment over all grouper versus nongrouper subjects across transfer, we found a significant difference between nongrouper and grouper subjects on the NEO-FFI Extraversion scale ($t(69) = 2.588$, $p < .05$). This result indicated that nongrouper subjects exhibit a higher level of extraversion ($M = 30.5$; SD = 6.5) than grouper subjects ($M = 25.8$; SD = 6.6). However, the remaining cognitive assessments indicated no differences between grouper and nongrouper subjects.

Discussion

The results of Experiment 1 indicated that (1) subjects can generally follow instructions regarding response execution strategy, (2) under instructions to synchronize responses, virtually all subjects exhibit learned parallelism, and (3) the instructions did not evoke robust and stable strategies for all subjects, which was marked by strategy switches for a number of subjects during the final transfer including a neutral instruction. Our data also support the hypothesis that the adaption of each response style (i.e., grouping vs. nongrouping) seems to be not exclusively affected by fundamentally distinct dissimilarities in cognitive capacities or personality. Rather, there seem to be strategy preferences for response styles, but these response styles are susceptible to a flexible adaption to task demands such as task instructions.

Additionally, the results are consistent with the set-cue bottleneck model, which assumes implicitly that all subjects can synchronize responses (though it does not speak to whether, or why, they will spontaneously use that strategy), and it specifies that response synchronization, or grouping, will eventually always lead to learned parallelism. The proposed model is also consistent with the empirically highly successful ES model of sequential retrieval stage processing when subjects were instructed to not synchronize responses.

Table 7 Error rates (in percent) from the start (Practice Triad 1) to the end of practice (Practice Triad 25) in Experiment 2

Instruction group	Keypress single retrieval		Vocal single retrieval		Keypress dual retrieval		Vocal dual retrieval		First dual-retrieval response		Second dual-retrieval response	
Practice triad	1	25	1	25	1	25	1	25	1	25	1	25
Synchronize instruction	3.7	2.5	4.6	1.4	4.4	2.3	2.3	1.7	3.9	2.3	2.3	1.7
Immediate instruction	5.6	1.8	3.2	2.0	3.6	1.7	1.2	0.3	1.8	1.5	2.9	2.4

Experiment 2

In Experiment 1, we were able to observe that instruction manipulations are an effective means of evoking a specific response strategy in dual-retrieval practice situations (i.e., grouping strategy and nongrouping strategy). Experiment 2 focused on the processing architecture of dual retrieval during practice under different response strategy conditions. Previous experiments showed that learned parallelism seems to be associated with cue-specific response chunking (Nino & Rickard, 2003; Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014). Cue-specific chunking only occurred after dual-retrieval practice when subjects grouped responses and it did not transfer to cues that had never been practiced in a dual-retrieval situation (i.e., the chunking process only occurs for specific, practiced, cues). As noted earlier, these results are in line with the *set-cue bottleneck model* of dual retrieval.

However, since the instruction to synchronize responses effectively evoked learned parallelism in Experiment 1, it needs to be further assessed if this instruction might yield transfer of learned parallelism to new cues (i.e., unpracticed cues) that are not practiced under dual-retrieval conditions. This is a possibility that would clearly speak against the assumption of a bottleneck stage of dual retrieval in the context of the set-cue bottleneck model. With this experiment, we, thus, perform an explicit test of the assumptions of the set-cue bottleneck model. Since we did not control for potential influences of dual-retrieval strategies on the dual-task processing architecture in earlier studies (Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014), this experiment reaches for a deeper understanding of strategies and their influence on retrieval processes.

As a primary change to Experiment 1, we used a randomly mixed combination of old (i.e., practiced) and new (i.e., unpracticed) cues as well as a continuing response strategy manipulation (i.e., synchronize versus immediate instructions) in the dual-retrieval task across practice and transfer. The cue-level chunking account predicts that we should observe no form of learned parallelism during the first exposure to new cues on the transfer test (e.g., RT2 should not fall below the ES prediction on the first transfer triad). Alternatively, a purely *instruction account,* equal to the task-level account, predicts task-level parallelism, which should allow the presence of learned parallelism on new cues (e.g., RT2 should fall below the ES prediction on the first transfer triad). In this scenario, the instruction to synchronize responses eliminates the retrieval stage processing bottleneck for new cues. In contrast, results for the immediate strategy instruction are assumed to reflect no impact on dual-retrieval response strategy toward parallelism. Subjects in this instruction manipulation group should reflect sequential retrieval, regardless of the cue type (i.e., old versus new cues).

Methods

Subjects

The experiment included a total of 48 subjects which were randomly divided across 2 different instruction groups (synchronize instruction and immediate instruction); we did not conduct a neutral instruction group since this group's experimental design would be similar to the design in Strobach, Schubert, Pashler, & Rickard (2014, Experiment 2) as well as in Orscheschek, Strobach, Schubert, & Rickard (2018, Experiment 2 and 3). As in Experiment 1, this division resulted in 24 subjects in each group. All of the subjects were undergraduate students at the Medical School Hamburg, Germany. All subjects gave written informed consent prior to their inclusion and participation in the experiment. The total sample had a mean age of 22.6 years with a range from 18 to 29 years. A total of 40 of the subjects were females. All of the subjects had normal or corrected to normal vision and were right handed. They were able to receive credit points or a voucher with a monetary worth of 24 Euros; only 4 subjects took the voucher as compensation. Alike Experiment 1, all of the subjects were naïve to the research aim of the study.

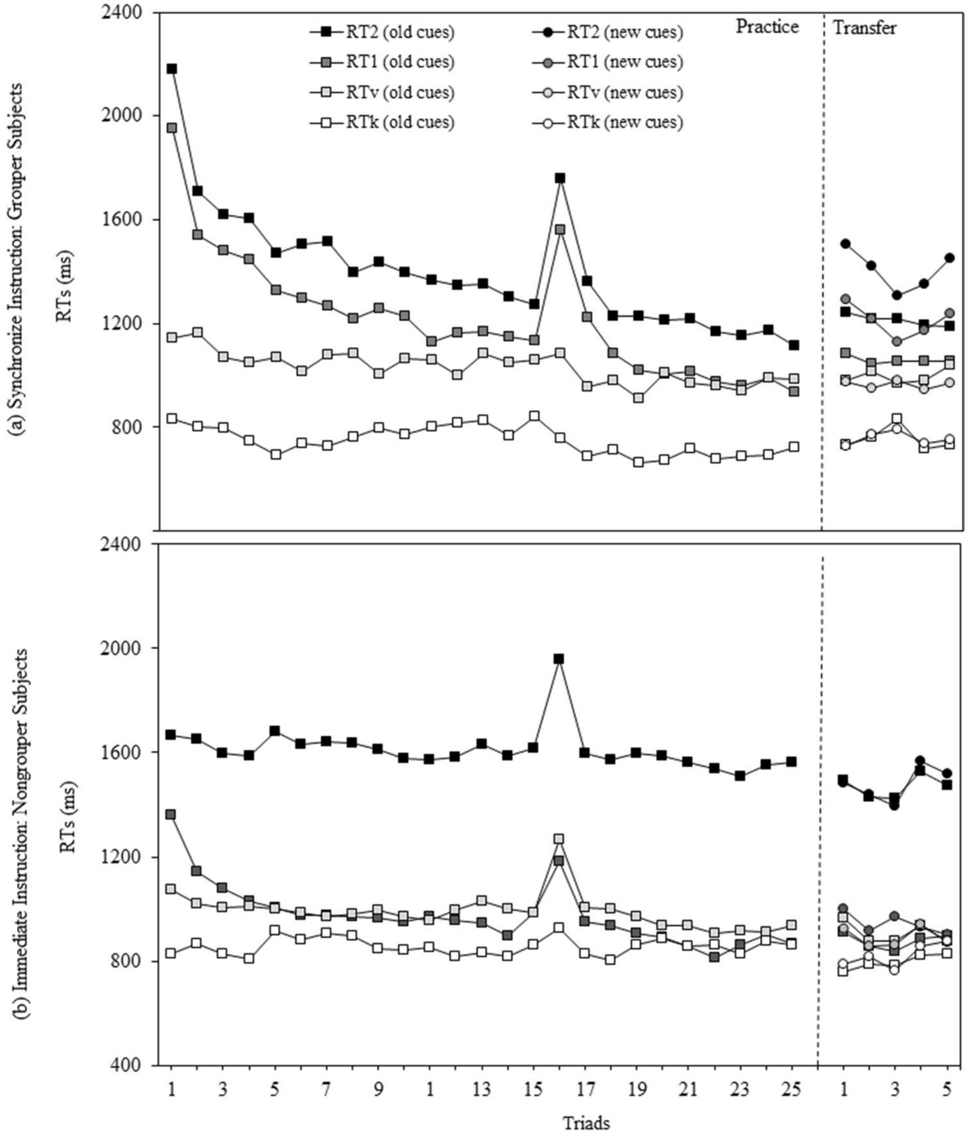


Fig. 8 Observed reaction times (RTs) in single-retrieval blocks of the keypress task and the vocal task as well as observed RTs in dual-retrieval blocks (i.e., RT1 and RT2) in the overall dataset during the 25 practice triads and 5 transfer triads in the synchronize instruction (**a**) and the immediate instruction (**b**) group of Experiment 2

Apparatus, cues, design, procedure, and statistical approach

Except the following changes, all elements were equal to Experiment 1. Only 2 instruction manipulation groups were used: the *synchronize instruction* and the *immediate instruction;* subjects in each group were presented with the same instruction throughout practice and transfer. Further, we increased the number of cues to a total of 14 cues (Table 7). All of these 14 cues were presented across the single-task trials in all phases (i.e., all 14 cues in the single-keypress task and in the single-vocal task). However, only 7 cues (i.e., old cues) were presented in the dual-retrieval trials across the single–dual practice phase, while all 14 cues were presented in the dual-retrieval trials (i.e., 7 old cues and 7 new cues) of the transfer phase.

Results

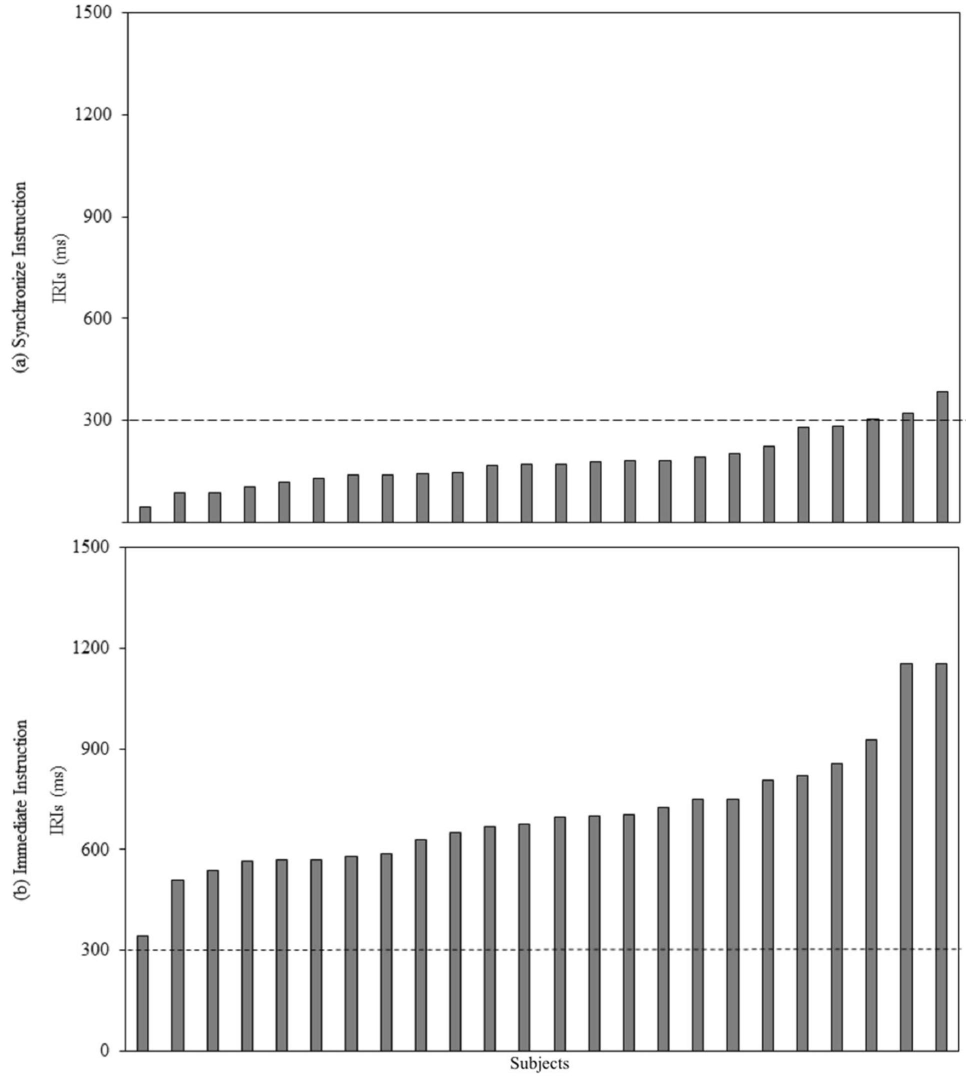
Accuracy results

Likewise to Experiment 1, we excluded trials with RTs below 200 ms. The error rates decreased from the beginning to the end of practice across both instruction groups. Further, the decrease affected both single- and dual-retrieval trials (see Table 7 for a detailed overview of the error rates).

RT results

Figure 8 displays the RTs averaged over all subjects in the correctly performed single- and dual-retrieval trials (i.e., single-retrieval trials: keypress task, vocal task; dual-retrieval trials: RT1, RT2) for each of the two instruction groups. We

Fig. 9 Inter-response intervals (IRIs) across the practice phase of individual subjects in the synchronize instruction group (**a**) and the immediate instruction group (**b**) of Experiment 2



observed a decrease in RTs across triads 1–25. This decrease was mainly evident for RT1 and RT2. Additionally, there was a slowing of response times at the beginning of Session 3 (Practice triad 16). Further, we detected an increase in RTs at the beginning of the transfer phase for new cues in the grouping and nongrouping instruction groups.

IRI analysis

Mean IRIs were computed across all practice-phase dual-retrieval trials for each subject. Figure 9 shows these results for both instruction groups, rank ordered from short to long IRIs. As in Experiment 1, we categorized subjects with an IRI below 300 ms as *grouper subjects* and subjects with an IRI above 300 ms as *nongrouper subjects* (Nino and Rickard, 2003; Strobach, Schubert, Pashler, & Rickard, 2014). According to this cut-off value, twenty-one subjects in the synchronize instruction condition were identified as response grouper ($M = 160$ ms, SD = 58 ms) and three subjects as response nongrouper ($M = 329$ ms, SD = 37 ms). In the immediate instruction condition, all subjects displayed an IRI above 300 ms and were classified as nongrouper subjects ($M = 705$ ms, SD = 185 ms).

Practice-phase dual-retrieval RTs

As can be observed in Fig. 10a, for grouper subjects in the *synchronize* instruction condition, RT1 remained above the ES prediction, whereas RT2 was falling below the ES prediction from triad 2 but remained close to the ES prediction until triad 6. From triad 7 on, RT2 fell below the ES prediction. By the end of practice (Triads 21–25), the averaged RT2 was significantly below the ES prediction ($t(20) = -5.437$, $p < .001$). Moreover, these grouper subjects displayed extreme ($BF_{10} = 946$) evidence for H1. This supports our previous findings and indicates that these subjects indeed showed learned parallelism (Nino & Rickard, 2003;

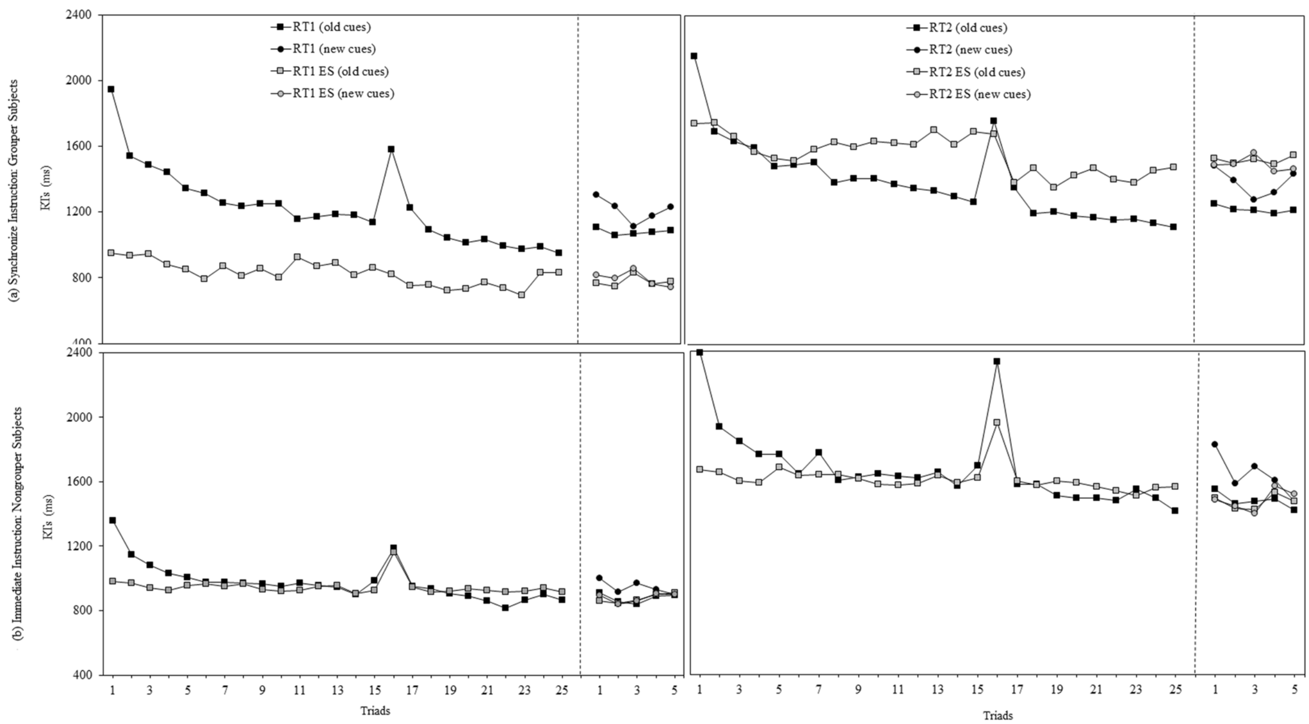


Fig. 10 Observed reaction times (i.e., RT1, RT2) as well the predictions of the efficient sequential (ES) model during the single–dual practice and transfer phases for grouper subjects in the synchronize instruction group (**a**) and the immediate instruction group (**b**) of Experiment 2

Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014).

In the *immediate* instruction group, we were able to observe 24 subjects that can be classified as nongroupers. RT1 was close to the ES prediction, whereas RT2 was first above but then remained close to the ES prediction (Fig. 10b). By the end of practice (Triads 21–25), averaged RT2 remained equal to the ES prediction ($t(23) = -1.317$, $p > .1$). Additionally, these nongrouper subjects displayed anecdotal ($BF_{10} = 0.462$) evidence for H0. Taken together, these results are further supportive of our previous observations in Experiment 1 regarding the effectiveness of strategy instructions. We were also able to replicate the finding of Experiment 1 that the instruction to group responses is an effective approach to evoke RT2 patterns that reflect learned parallelism.

Transfer-phase dual-retrieval RTs

The first transfer triad is of prior importance for the interpretation of Experiment 2 (Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014). For *grouper subjects* in the synchronize instruction condition (Fig. 10a) an ANOVA of Cue type (old cues vs. new cues) and Dataset (RT2 vs. ES prediction) displayed a main effect for Dataset ($F(1,20) = 5.959$, $p < .05$), as well as a significant interaction effect ($F(1,29) = 15.835$, $p < .001$). Further paired *t* tests indicated a significant difference for RT2 and the ES prediction for old cues ($t(20) = -4.107$, $p < .001$), but no such difference for new cues ($t(20) = -.113$, $p > .5$). The additional Bayes *t* tests on the first transfer triad revealed very strong evidence for H1 for old cues ($BF_{10} = 61.239$) and moderate evidence for H0 for new cues ($BF_{10} = 0.229$).

For the immediate instruction group (Fig. 10a), an ANOVA of Cue type (old cues vs. new cues) and Dataset (RT2 vs. ES prediction) showed no significant interaction effect ($F(1,23) = 4.043$, $p = .056$) but a main effect for Cue type ($F(1,23) = 4.988$, $p = .036$) as well as Dataset ($F(1,23) = 5.093$, $p = .034$). Paired *t* tests showed no significant difference between RT2 and the ES prediction on old cues on the first transfer triad ($t(23) = .815$, $p > .1$). Bayes tests provided further support by indicating moderate evidence for H0 ($BF_{10} = 0.290$). For new cues, RT2 was significantly above the ES prediction ($t(23) = 2.313$, $p < .05$), which shows no violation of the ES prediction for new cues. Further, this result was reflected by anecdotal evidence for H1 ($BF_{10} = 1.956$).

Discussion

The results of Experiment 2 replicated and extended the results of Experiment 1, indicating that the instruction to synchronize responses also led to the presence of learned retrieval parallelism with 14 and seven cues in single- and dual-retrieval trials during practice, respectively, and 14 cues in all trial types during transfer. Additionally, we can conclude that the instruction to synchronize responses was not able to provide data suggesting the elimination of the retrieval stage processing bottleneck. These new results extended the set-cue bottleneck model, which proposes a cue-specific chunking mechanism for learned parallelism even under the condition of synchronize instructions (Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014).

General discussion

This paper had two specific research aims: (1) investigation of the impact of explicit instructions on response patterns during dual-memory retrieval practice (Experiments 1 and 2), and (2) assessment of the processing architecture of dual-memory retrieval after practice (Experiment 2). The results of Experiments 1 and 2 showed that instruction manipulations were able to effectively influence response patterns and to induce an effective grouping strategy which was marked by RT2 patterns that revealed the presence of learned parallelism for subjects in the synchronize instruction condition. Further, the instruction to use a sequential response pattern in the immediate instruction group was effective as well. Subjects in this instruction group showed strict sequential response patterns as predicted by the ES model. In line with previous experiments (Nino & Rickard, 2003; Strobach, Schubert, Pashler, & Rickard, 2014), subjects in the neutral instruction group reflected RT2 patterns that indicated the presence of learned retrieval parallelism as well as sequential response patterns. The data of Experiment 1 also showed that when providing a neutral instruction after dual-retrieval practice, some subjects remained with the practiced synchronized or immediate instruction while others switched the strategies. In Experiment 2, it was hypothesized that, according to a cue-level chunking account, no evidence for learned parallelism should be found during the first exposure to new cues even though subjects were explicitly instructed to practice parallel response patterns. Next to that, the immediate instruction group was hypothesized to show no evidence for learned parallelism across practice as well as transfer.

Impact of instructions on response patterns during dual-memory retrieval practice

The results of Experiments 1 and 2 can be discussed with respect to the influence of explicit task instructions. It can be concluded that explicit instruction manipulations are able to effectively influence and direct dual-retrieval patterns. Whereas previous dual-retrieval experiments that used two responses from a single cue were only able to observe and analyze response patterns which were naturally adapted by the subjects (i.e., grouper subjects vs. nongrouper subjects; Nino & Rickard, 2003; Orscheschek, Strobach, Schubert, & Rickard, 2018; Strobach, Schubert, Pashler, & Rickard, 2014), this study showed that both, grouping and nongrouping, response patterns can also be created through instructions. Interestingly, the results of subjects who naturally chose one of the strategies (i.e., either grouper or nongrouper subjects in the neutral instruction group) mirrored the results in either instruction manipulation group (i.e. grouper subjects in the synchronize instruction and nongrouper subjects in the immediate instruction group). This shows not only that the instructions worked, but also that the instructed response patterns are highly comparable to the naturally adapted patterns. Our initial question was, why individuals reflect grouping or nongrouping patterns and if the adaption of either response pattern is formed by (1) fundamental individual differences that only allow the use of one response pattern (Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002), or (2) by a strategy preference (e.g., response style) that can still be mediated by factors such as task demands and other features (Jansen, van Egmond, & de Ridder, 2016). As we mentioned before, we concluded that, if response patterns were exclusively caused by fundamental individual differences, specific task and response instructions should not affect response patterns. However, our results indeed showed that strategy manipulations through instructions were able to influence grouping and nongrouping behavior by increasing the proportion of grouper subjects in the synchronize instruction condition in contrast to the immediate instruction condition and vice versa. The results further showed that response grouping was able to induce parallel retrieval for all, or nearly all, subjects that naturally chose to synchronize their responses as well as subjects that were instructed to do so in the synchronize instruction condition. Looking at all of our results from the present study, we can assume that the adaption of each response style (i.e., grouping vs. nongrouping) seems to be not exclusively affected by fundamental individual differences such as cognitive capacities or personality. Rather, there seem to be strategy preferences for response styles, but these response styles are susceptible to a flexible adaption to task demands such as task instructions.

Our results extend our knowledge on explicit instruction manipulations to a dual-retrieval context from PRP training and non-training studies. In most PRP training studies on instructions, subjects have to follow a specific priority instruction (i.e., "respond to task A as fast as possible"). Such studies were mostly able to observe that subjects followed the instruction and exhibited sequential response patterns. Even though this sequential pattern could be accounted for by the presence of dual-task costs (i.e., response slowing for the second response in a dual-task situation) (Han & Marois, 2013), priority instructions still worked out even after extensive practice which is known to reduce dual-task costs (Maquestiaux, Laguë-Beauvais, Bherer, & Ruthruff, 2008; Strobach & Schubert, 2017). This means, that even though there was a reduction of the reaction time for the second response after practice, subjects still followed the priority instruction and completed the first task prior to the second one. In addition, several training and PRP non-training experiments that instructed the subjects to equally emphasize both tasks were able to observe nongrouping as well as grouping patterns, with some experiments even showing a complete elimination of dual-task costs after practice (Strobach, Frensch, Mueller, & Schubert, 2012a, b). Ruthruff, Pashler, & Klaassen (2001) introduced a dual-task situation in which both tasks did not share output modalities, were given equal importance through an instruction and were further presented at the same time (i.e., SOA = 0 ms). The authors were able to observe a grouping strategy with a mean IRI of 64 ms for subjects that were explicitly instructed to produce both responses at the same time and to accentuate both tasks evenly. These results are in line with our observations of the synchronize strategy instruction manipulation in Experiments 1 and 2. Another group of subjects in Ruthruff et al. study was instructed to place equal importance on each task and to complete the task as fast as possible (e.g., no explicit grouping strategy). In this instruction group, they observed grouper as well as nongrouper subjects similar to the results in our neutral instruction manipulation group. Taken together, these results are in line with our present findings and demonstrate that specific task instructions are able to evoke sequential as well as parallel response patterns in dual-task situations.

Further, the significant influence of explicit task instructions can also be observed in numerous cognitive training situations. For example, Laine, Fellman, Waris, & Nyman, (2018) investigated the relationship between strategy usage and working memory-training effects in an *n*-back task. The authors differentiated between an externally provided strategy by the experimenter and internal self-chosen strategies (e.g., rehearsal, updating, grouping), measured by self-report; such external and internal strategies are equivalent to our grouping and non-grouping instruction groups versus the neutral instruction group. Their results showed a clear benefit of strategy use, regardless of an internally chosen or externally provided strategy, in comparison to subjects that reportedly applied no strategy. It was further concluded that the level of detail in a strategy is accountable for higher performance in working memory training. This conclusion was based on the finding that subjects, which reported a detailed internal strategy, reflected high performance levels like the subjects that were provided with the detailed external strategy. Similar to Laine et al. we did observe that explicit strategy instructions are effective means to increase specific strategy use. Further, the instructed response patterns in our study (i.e., synchronize instruction and immediate instruction) cannot be interpreted as having increased efficacy or significantly heightened performance than the self-chosen strategies in the neutral instruction group. In general, comparing our findings of successful strategy instructions with other studies on instructions in cognitive training and applied psychology (i.e., studies on spatial rotation training (Meneghetti, Cardillo, Mammarella, Caviola, & Borella, 2017) and math difficulties (Swanson, Orosco, & Lussier, 2014), we are able to see a broad picture of the effectivity of strategy instructions.

As outlined before, strategies seem to be able to make cognitive training more effective and help to get the best performance out of existing cognitive resources (Laine, Fellman, Waris, & Nyman, 2018; McNamara & Scott, 2001; Meneghetti, Cardillo, Mammarella, Caviola, & Borella, 2017). Additionally, previous research on instruction manipulations found that subjects are well able to follow and adapt to explicit task instruction (Jansen, van Egmond, & de Ridder, 2016). These results are in line with our present findings as well, since a large number of subjects appear to have still used the previously instructed strategy during the transfer phase of Experiment 1 in which they received a neutral task instruction and could have applied any response strategy. Further, the other group of subjects, classified as switcher subjects, might have tried to explicitly follow the new neutral instruction that generally evokes the presence of grouper and nongrouper subjects in other experiments as well (Nino & Rickard, 2003; Strobach, Schubert, Pashler, & Rickard, 2014). In contrast, it could also be argued that individual strategy preferences could have determined strategy switches. While all subjects were well able to follow the instructed patterns, the moment they received a chance to use a different strategy might have caused switcher subjects to perform another strategy that they prefer more. This assumption would be in line with the findings of some authors who have examined the possibility of individual preferences for specific retrieval strategies (Brüning & Manzey, 2018; Fischer & Plessow, 2015; Reissland & Manzey, 2016). However, our data only partially indicated that such preferences could be previously determined by different response patterns during practice: While the IRIs of switcher and stable subjects did not differ during the practice phase in the synchronize instruction group, in the immediate instruction group, switcher subjects already indicated significantly lower IRIs during the practice phase. Therefore, while our results seem to support the assumption that the preference for response strategies, or styles, is, or can be, determined by task demands and instructions, we also found evidence that suggests an influence of individual preference on strategy choice in dual-memory retrieval. Further, despite a group of authors that assumed that such strategy preferences are caused by differential functional abilities such as high or low working memory capacity (Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002) or by person inherent characteristics such as personality factors (Ishizaka, Marshall, & Conte, 2001; Sanbonmatsu, Strayer, Medeiros-Ward, & Watson, 2013), we cannot provide further evidence for these claims. Our cognitive assessments did not display any differences between grouper and nongrouper subjects with regard to processing speed, attention or working memory. We only established to see a difference in the level of extraversion on the NEO-FFI. However, since there were no differences across the other big-five personality dimension, definite conclusions on an influence of extraversion require further in-depth investigations. Additionally, we did not find any divergence between switcher and stable subjects in terms of individual differences regarding cognitive domains or personality factors based on our additional measures.

Processing architecture of dual-memory retrieval after practice

Looking into our assessment of the processing architecture of dual-memory retrieval after practice in Experiment 2, our results showed no evidence for learned parallelism for new cues in the transfer phase. Instead, subjects did reflect sequential processing as soon as new cues were introduced on the dual-retrieval task. This reflects a response pattern that is in line with the assumption of cue-specific response chunking in the context of the set-cue bottleneck model. With the present experiment, we were further able to argue against previous concerns which stated that subjects might engage in a trial-wise parallel-versus-sequential strategic decision on old cues and new cues, respectively, during the transfer test (Strobach, Schubert, Pashler, & Rickard, 2014). Due to the effectiveness of explicit task instructions, we further conclude it unlikely that subjects made a strategic choice about which pattern to use during the dual-retrieval trials (i.e., parallel vs. sequential). The finding further supports the notion of learned retrieval parallelism in a cue-level chunking account.

Moreover, the results also point towards the presence of a structural bottleneck during practiced dual-memory retrieval. Early dual-task studies have assumed that the bottleneck observed in dual-task situations might be due to a result of priority instructions (Meyer & Kieras, 1997). This would imply that subjects would indeed possess the ability to perform processes in parallel, but that the instruction to prioritize one task would lead to a voluntary adaption of a bottleneck strategy. Additionally, some authors further suggested that the bottleneck could be eliminated after extensive task practice with equal task priority (Schumacher et al. 2001); however, these studies provided no direct empirical evidence for this suggestion. In contrast, an array of different studies opposed this prediction, since a bottleneck was further observed in designs that instructed equal emphasis on both tasks (Ruthruff, Pashler, & Klaassen, 2001; Ruthruff, Pashler, & Hazeltine, 2003; Tombu & Jolicoeur 2000) as well as in studies that used neutral instructions and extensive practice (Nino & Rickard, 2003; Strobach, Schubert, Pashler, & Rickard, 2014). In this experiment, we were able to provide an even stronger assessment of the possibility that the bottleneck could be eliminated under pressure to engage in parallel processing. Our results show that the subjects in Experiment 2 were not able to perform parallel response patterns for new cues, even though they were explicitly instructed to do so and exclusively showed evidence for instructed learned parallelism on old cues during extensive practice. This provides strong evidence for a structural base of the bottleneck and supports the assumptions of the set-cue bottleneck model of dual-memory retrieval.

Looking at limitations, we have to focus on the distinction between grouper and nongrouper subjects according to the absolute IRI threshold of 300 ms. While previous studies (Nino & Rickard, 2003; Strobach, Schubert, Pashler, & Rickard, 2014) were able to observe a substantial gap at 300 ms, this gap was not evident in our data. Despite, the threshold of 300 ms was still applied throughout our analyses. Nevertheless, our results show, that both groups of grouper and nongrouper subjects indeed differ with regard to their relation with the ES prediction. We are aware that there might be a potential risk of including subjects in one of the subsamples that are just above or beyond the threshold and which data patterns might point into a different direction. Therefore, we strongly encourage the search for more theoretically implied as well as methodologically stronger assessment methods that can account for a comprehensive and less empirical discrimination between grouper and nongrouper subjects.

To conclude, the results of the present study extend previous studies on dual-memory retrieval from a single cue, and to advancing our understanding of the effects and the role of instruction manipulations in dual-task research. This study was the first to analyze instructed forms of learned retrieval parallelism of dual-memory retrieval. The results further support the set-cue bottleneck model as a candidate cognitive processing architecture for dual-memory retrieval.

Acknowledgements We would like to thank Anja Skoglund, Merle Schüler and Cerly Teymourian for their assistance with data collection. The study and data collection have been performed in accordance with Standard 8 of the American Psychological Association's Ethical Principles of Psychologist and Code of Conduct. The manuscript does not contain clinical studies or clinical patient data. Informed consent was obtained from all individual participants included in the study. The author(s) declared no potential conflicts of interest with respect to the research, authorship, and /or publication of the article. This study was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under grant number STR 1223/1.

Funding The study was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under grant number STR 1223/1.

Compliance with ethical standards

Conflict of interest The author(s) declared no potential conflicts of interest with respect to the research, authorship, and /or publication of the article.

Ethical approval The study and data collection have been performed in accordance with Standard 8 of the American Psychological Association's Ethical Principles of Psychologist and Code of Conduct. The manuscript does not contain clinical studies or clinical patient data. All procedures performed in studies were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies performed on animals.

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

Appendix

The set-cue bottleneck model of dual-memory retrieval

The set-cue bottleneck model assumes three sequential, independent, and additive processing stages: perceptual, central, and motor processing. Figure 11 shows how these three stages map onto the set-cue architecture to be described in detail next. The perceptual stage includes all activation flow from the moment of cue presentation up to threshold-level activation at the input level. The retrieval stage spans from the moment of activation gating (i.e., the moment of threshold activation) from input level to threshold activation at the response level. The motor stage includes the gated activation flow from the response level to completion of the motor response. These three stages can be mapped directly to the perceptual, retrieval, and motor stages of the set-cue bottleneck model.

The set-cue bottleneck model assumes that associations are formed only between activated nodes at adjacent levels

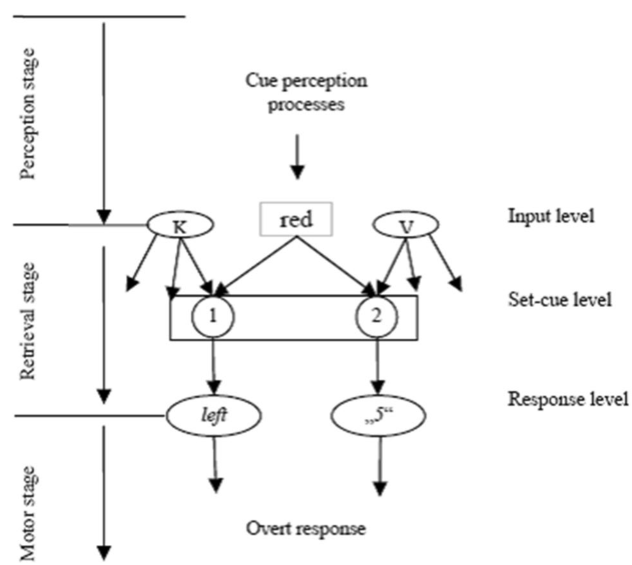


Fig. 11 Basic architecture of the set-cue bottleneck model including perceptual, retrieval, and motor stages. Nodes at the input level represent the presented cue and the currently active task set(s); each node at the set-cue level represents a particular combination of presented cue and task set; each task set node in turn can have an association of one or more nodes at the response level. K = task set of keypress task; V = task set of vocal task

of the hierarchy. Nodes at the input level represent the presented cue and the currently active task set(s). A task set will be defined for current purposes as a goal to execute one of a set of responses in a given modality. For the *keypress task*, the task set node (*K*) represents the goal to execute a correct key press (left or right) when the cue is presented. The *vocal-digit task* set node (*V*) represents the goal to execute the correct vocal-digit response when the cue is presented. Activation of task set nodes is, thus, assumed to be under strategic control. Note that for a given task (keypress or vocal) there is only one task set node.

Each node at the second level of the set-cue bottleneck model (the set-cue level) represents a particular combination of presented cue and task set. Each task set node in turn can have an association with one or more nodes at the response level. Nino and Rickard (2003) proposed that, whereas activation streams can flow in parallel from the earliest level of cue perception to the set-cue level, and from the set-cue level to the response and motor levels, there is a winner-take-all competition at the set-cue level. After a winning set-cue node is selected, activation begins to flow from (only) the winning set-cue node to the response level. In the model as developed here, the entire *retrieval stage* of processing—from the moment at which activation flow is gated forward from the input level to the moment of threshold activation at the response level—must be completed before activation flow for a second retrieval event can be initiated from the input level.

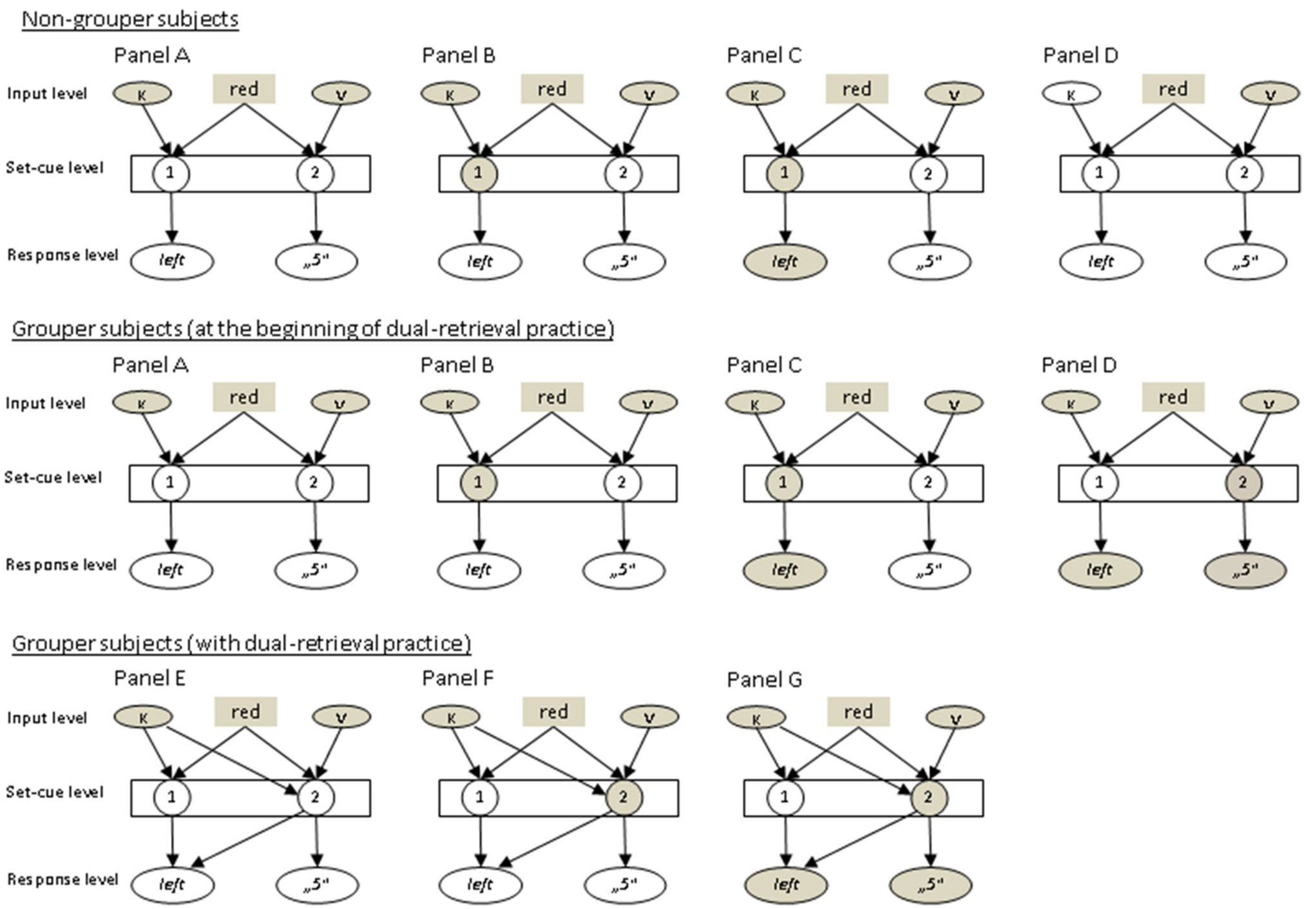


Fig. 12 Snapshot activations of dual retrieval in the set-cue bottleneck model for nongrouper subjects throughout dual-retrieval practice (first row) and grouper subjects (i.e., groupers) at the beginning of dual-retrieval practice (second row) and with dual-retrieval practice (third row)

To specify mechanisms of learned parallelism in grouper subjects, snapshot activation states of the set-cue bottleneck model during dual retrieval are shown for two cases in Fig. 12. The first case depicts dual retrieval for groupers on the first dual-retrieval trial. Panel A depicts the node activation state just after cue presentation and just prior to activation gating to the set-cue level. Because both responses are required on dual-retrieval trials, both task sets are assumed to be activated at the outset of each trial. Activation then flows in parallel to all associated nodes at the set-cue level. The example case in which the set-cue node corresponding to the keypress task (labeled as node 1) wins the competition (i.e., reaches an activation threshold first, resulting in suppression of activation in all other nodes at that level), is shown in Panel B. Factors that could determine the winning set-cue node include noise, differential strength of the associations formed during single-retrieval phase of learning, or strategic scheduling that gives preference to one category of retrieval (i.e., keypress or vocal-digit) for first execution. Strategic scheduling could be implemented within this framework as enhanced initial activation of one of the task set nodes.

After the winning set-cue node reaches an activation threshold, activation then flows to the response level (Panel C). Because groupers are synchronizing their response execution, however, they do not immediately execute the first retrieved response, but rather keep that response active by a working memory mechanism as they execute the second retrieval. It is also reasonable to assume that the task set corresponding to the first retrieved response also remains activated until that response is executed. For set-cue node 2 to win the second competition, some mechanism that biases the competition in favor of node 2 must be assumed. Given such a mechanism, set-cue node 2 would win the second competition, and activation would flow from that set-cue node to the vocal-digit response (Panel D). At that point, both task sets and both responses are activated and there is synchronized keypress and vocal-digit response execution.

Important for the present context, we assume that a necessary and sufficient condition for associations between two nodes to be formed or strengthened is their joint activation. It follows that, during the activation state depicted in Panel D, there will be strengthening of the existing associations as well as formation of new associations between a) the keypress task set and the active set-cue node 2, and b) set-cue node 2 and the active keypress response node, as depicted in Panel E under the heading "Grouper subjects (with dual-retrieval practice)". In the set-cue bottleneck model, then, there is no task-level learning associated with response chunking or learned retrieval parallelism. Rather, response chunking occurs independently for each cue (i.e., cue-specific).

After one or more grouped response trials for a cue, the associations represented by the dashed arrows in Panel E become strong enough that they can support retrieval of both the vocal and keypress response via set-cue node 2. On subsequent dual-retrieval trials, set-cue node 2 is likely to win the initial activation competition at the set-cue level even though it was not the initial winner on previous dual-retrieval trials for that cue (in this example). This outcome would be expected because, at that point during training, both task sets would send activation to set-cue node 2, whereas only the keypress task set would send activation to set-cue node 1. Furthermore, because set-cue node 2 is linked with both response nodes, it activates both a keypress and a vocal-digit response. Hence, subjects can learn to retrieve both responses while making only one pass through the set-cue bottleneck stage, i.e., subjects can learn to chunk responses. The sequence of activation states for this case is shown in Fig. 12f, g.

References

Brickenkamp, R., & Zillmer, E. (1998). *The d2 test of attention*. Seattle, WA: Hogrefe & Huber.

Brüning, J., & Manzey, D. (2018). Flexibility of individual multitasking strategies in task-switching with preview: are preferences for serial versus overlapping task processing dependent on between-task conflict? *Psychological Research, 82*(1), 92–108. https://doi.org/10.1007/s00426-017-0924-0.

Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia, 38*(6), 848–863.

Chen, Q., & Yan, Z. (2016). Does multitasking with mobile phones affect learning? A review. *Computers in Human Behavior, 54,* 34–42.

Costa, P., & McCrae, R. (1989). *NEO five-factor inventory (NEO-FFI)*. Odessa, FL: Psychological Assessment Resources.

Dembo, M. H., & Seli, H. P. (2004). Students' resistance to change in learning strategies courses. *Journal of Developmental Education, 27*(3), 2–11.

Fischer, R., & Plessow, F. (2015). Efficient multitasking: parallel versus serial processing of multiple tasks. *Frontiers in Psychology, 6,* 1–11. https://doi.org/10.3389/fpsyg.2015.01366.

Han, S. W., & Marois, R. (2013). The source of dual-task limitations: Serial or parallel processing of multiple response selections? *Attention, Perception, & Psychophysics, 75*(7), 1395–1405. https://doi.org/10.3758/s13414-013-0513-2.

Hazeltine, E., Teague, D., & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance, 28*(3), 527–545.

Hoijtink, H., Klugkist, I., & Boelen, P. (2008). *Bayesian evaluation of informative hypotheses*. New York: Springer Science & Business Media.

Ishizaka, K., Marshall, S. P., & Conte, J. M. (2001). Individual differences in attentional strategies in multitasking situations. *Human Performance, 14*(4), 339–358. https://doi.org/10.1207/S15327043hup1404_4.

Jansen, R. J., van Egmond, R., & de Ridder, H. (2016). Task prioritization in dual-tasking: Instructions versus preferences. *PLoS ONE, 11*(7), 1–27. https://doi.org/10.1371/journal.pone.0158511.

JASP Team (2018). JASP (Version 0.8.6) [Computer software].

Jordan, J. D., Landau, E. M., & Iyengar, R. (2000). Signaling networks: the origins of cellular multitasking. *Cell, 103*(2), 193–200.

Kramer, A. F., Larish, J. F., & Strayer, D. L. (1995). Training for attentional control in dual task settings: a comparison of young and old adults. *Journal of Experimental Psychology: Applied, 1*(1), 50–76. https://doi.org/10.1037/1076-898x.1.1.50.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General, 142*(2), 573–603. https://doi.org/10.1037/a0029146.

Laine, M., Fellman, D., Waris, O., & Nyman, T. J. (2018). The early effects of external and internal strategies on working memory updating training. *Scientific Reports, 8*(1), 4045–4056. https://doi.org/10.1038/s41598-018-22396-5.

Lehle, C., & Hübner, R. (2009). Strategic capacity sharing between two tasks: Evidence from tasks with the same and with different task sets. *Psychological Research, 73*(5), 707–726.

Lehle, C., Steinhauser, M., & Hübner, R. (2009). Serial or parallel processing in dual tasks: What is more effortful? *Psychophysiology, 46*(3), 502–509. https://doi.org/10.1111/j.1469-8986.2009.00806.x.

Levy, J., & Pashler, H. (2008). Task prioritisation in multitasking during driving: Opportunity to abort a concurrent task does not insulate braking responses from dual-task slowing. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 22*(4), 507–525. https://doi.org/10.1002/acp.1378.

Maquestiaux, F., Laguë-Beauvais, M., Bherer, L., & Ruthruff, E. (2008). Bypassing the central bottleneck after single-task practice in the psychological refractory period paradigm: Evidence for task automatization and greedy resource recruitment. *Memory & Cognition, 36*(7), 1262–1282. https://doi.org/10.3758/MC.36.7.1262.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition, 29*(1), 10–17.

Meneghetti, C., Cardillo, R., Mammarella, I. C., Caviola, S., & Borella, E. (2017). The role of practice and strategy in mental rotation training: transfer and maintenance effects. *Psychological Research, 81*(2), 415–431. https://doi.org/10.1007/s00426-016-0749-2.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I. Basic mechanisms. *Psychological Review, 104*(1), 3–65.

Miller, J., & Durst, M. (2014). "Just do it when you get a chance": the effects of a background task on primary task performance. *Attention, Perception, & Psychophysics, 76*(8), 2560–2574.

Nino, R. S., & Rickard, T. C. (2003). Practice effects on two memory retrievals from a single cue. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 29*(3), 373–388.

Oberauer, K., & Bialkova, S. (2011). Serial and parallel processes in working memory after practice. *Journal of Experimental Psychology: Human Perception and Performance, 37*(2), 606–614. https://doi.org/10.1037/a0020986.

Oberauer, K., & Kliegl, R. (2004). Simultaneous cognitive operations in working memory after dual-task practice. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 689–707. https://doi.org/10.1037/0096-1523.30.4.689.

Orscheschek, F., Strobach, T., Schubert, T., & Rickard, T. (2018). Two retrievals from a single cue: A bottleneck persists across episodic and semantic memory. *Quarterly Journal of Experimental Psychology*, *72*(5), 1005–1028. https://doi.org/10.1177/1747021818776818.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological Bulletin, 116*(2), 220–244.

Pashler, H., & Johnston, J. C. (1998). Attentional limitations in dual-task performance. In H. Pashler (Ed.), *Attention* (pp. 155–189). Hove: Psychology Press.

Reissland, J., & Manzey, D. (2016). Serial or overlapping processing in multitasking as individual preference: Effects of stimulus preview on task switching and concurrent dual-task performance. *Acta Psychologica, 168,* 27–40. https://doi.org/10.1016/j.actpsy.2016.04.010.

Ruthruff, E., Pashler, H. E., & Hazeltine, E. (2003). Dual-task interference with equal task emphasis: Graded capacity sharing or central postponement? *Perception & Psychophysics, 65*(5), 801–816.

Ruthruff, E., Pashler, H. E., & Klaassen, A. (2001). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement? *Psychonomic Bulletin & Review, 8*(1), 73–80.

Ruthruff, E., Van Selst, M., Johnston, J. C., & Remington, R. (2006). How does practice reduce dual-task interference: Integration, automatization, or just stage-shortening? *Psychological Research, 70*(2), 125–142. https://doi.org/10.1007/s00426-004-0192-7.

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PLoS ONE, 8*(1), e54402. https://doi.org/10.1371/journal.pone.0054402.

Schell, K. L., & Conte, J. M. (2008). Associations among polychronicity, goal orientation, and error orientation. *Personality and Individual Differences, 44*(1), 288–298. https://doi.org/10.1016/j.paid.2007.08.009.

Schmidt, K., & Metzler, P. (1992). *Wortschatztest (WST)("German vocabulary test for crystalline intelligence estimation")*. Weinheim: Beltz Test GmbH.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological science, 12*(2), 101–108.

Sternberg, R. J., Zhang, L.-F., & Rayner, S. (2011). *Handbook of intellectual styles: Preferences in cognition, learning, and thinking*. Berlin: Springer Publishing Company.

Strobach, T., Frensch, P., Mueller, H. J., & Schubert, T. (2012a). Testing the limits of optimizing dual-task performance in younger and older adults. *Frontiers in Human Neuroscience, 6,* 39–50. https://doi.org/10.3389/fnhum.2012.00039.

Strobach, T., Frensch, P., Müller, H., & Schubert, T. (2012b). Age-and practice-related influences on dual-task costs and compensation mechanisms under optimal conditions of dual-task performance. *Aging Neuropsychology, and Cognition, 19*(1–2), 222–247. https://doi.org/10.1080/13825585.2011.630973.

Strobach, T., & Schubert, T. (2017). Mechanisms of practice-related reductions of dual-task interference with simple tasks: data and theory. *Advances in cognitive psychology, 13*(1), 28–41. https://doi.org/10.5709/acp-0204-7.

Strobach, T., Schubert, T., Pashler, H., & Rickard, T. (2014). The specificity of learned parallelism in dual-memory retrieval. *Memory & Cognition, 42*(4), 552–569. https://doi.org/10.3758/s13421-013-0382-x.

Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence, 30*(3), 261–288.

Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2014). The effects of mathematics strategy instruction for children with serious problem-solving difficulties. *Exceptional Children, 80*(2), 149–168. https://doi.org/10.1177/001440291408000202.

Tombu, M., & Jolicoeur, P. (2000). *Is the PRP effect due to a strategic or structural bottleneck*. New Orleans: In Poster presented at the annual meeting of the Psychonomic Society.

Wechsler, D. (2008). *Wechsler adult intelligence scale-fourth edition (WAIS–IV): San Antonio*. San Antonio, TX: The Psychological Corporation.