Working memory capacity and dual-task interference in picture naming

Vitória Piai [a,b,*], Ardi Roelofs [a]

[a] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands
[b] International Max Planck Research School for Language Sciences, Wundtlaan 1 6525 XD, Nijmegen, The Netherlands

ARTICLE INFO

ABSTRACT

Researchers have found no agreement on whether dual-task interference in language performance, such as dual-task interference from tone discrimination on picture naming, reflects passive queuing or active scheduling of processes for each task. According to a passive-queuing account, while a central response-selection bottleneck is occupied by the tone discrimination task, picture naming is held in a passive queue until the bottleneck is freed. In contrast, according to an active-scheduling account, participants determine the order in which the tasks proceed, monitor progress on the tasks, suspend picture naming and hold it in working memory, and determine when to resume picture naming. Here, we report a study that assessed the relative merits of the queuing and scheduling accounts by examining whether the magnitude of dual-task interference in picture naming is associated with individual differences in the capacity of monitoring and updating of working memory representations, as assessed by the operation-span task. We observed that the updating/monitoring ability correlated with the speed of picture naming and with the magnitude of the interference from tone discrimination on picture naming. These results lend support to the active-scheduling account of dual-task interference in picture naming.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Speakers are typically able to quickly and accurately access words in long-term memory. Lexical access in spoken word production has been extensively studied using the picture–word interference (PWI) paradigm, in which speakers name pictures while trying to ignore spoken or written distractor words. For example, they say "dog" to a pictured dog combined with the written word *cat* (i.e., a word from the same semantic category, here animals; the semantic condition) or the word *tree* (the unrelated condition). Previous research has shown that mean response time (RT) is longer for semantically related picture–word stimuli relative to unrelated stimuli, an effect called *semantic interference* (e.g., Damian & Martin, 1999; Glaser & Düngelhoff, 1984; Glaser & Glaser, 1989). The semantic interference effect is only obtained when speakers have to select a word to name the picture, but not when a manual response to the picture is required (Schriefers, Meyer, & Levelt, 1990), which suggests that the effect arises during lexical selection for word production. Given that the semantic effect is one of interference rather than facilitation, lexical selection has been taken to be a competitive process (e.g., Abdel Rahman & Melinger, 2009; Hantsch, Jescheniak, & Schriefers, 2005; Levelt, Roelofs, & Meyer, 1999; Roelofs,

1992, 2003; Starreveld & La Heij, 1996). This account of semantic interference has been computationally implemented in a number of models of word production, including the model of Starreveld and La Heij (1996) and WEAVER++ (Levelt et al., 1999; Piai, Roelofs, & Schriefers, 2011; Roelofs, 1992, 2003, 2007, 2008a,b,c).

The assumption of a lexical response-selection locus of the semantic interference effect was recently challenged by two dual-task studies (Ayora et al., 2011; Dell'Acqua, Job, Peressotti, & Pascali, 2007). These studies revealed that the semantic interference effect may be absent when participants perform picture naming in the context of a tone discrimination task requiring a manual response. According to Dell'Acqua et al. and Ayora et al., response selection constitutes a structural processing bottleneck in dual-task performance (cf. Ferreira & Pashler, 2002). Therefore, response selection in the tone discrimination task precludes concurrent response selection in picture naming. While the central response-selection bottleneck is occupied by the tone discrimination task, picture naming is held in a passive queue until the bottleneck is freed, henceforth the *passive queuing account* (cf. Adcock, Constable, Gore, & Goldman-Rakic, 2000; Bunge, Klingberg, Jacobsen, & Gabrieli, 2000; Dux et al., 2009; Jiang, Saxe, & Kanwisher, 2004; Pashler, 1984, 1994, 1998; Pashler & Johnston, 1989). If the semantic interference effect arises before response selection in picture naming (e.g., during perceptual and conceptual processing of the picture), the effect may be absorbed while response selection in the tone discrimination task is taking place and picture naming waits in queue. Consequently, the semantic interference effect will disappear in concurrent task performance, as empirically observed by Dell'Acqua et al. and

* Corresponding author at: Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands. Tel./fax: +31 24 3612635.
E-mail addresses: v.piai@donders.ru.nl (V. Piai), a.roelofs@donders.ru.nl (A. Roelofs).

Ayora et al. In contrast, if the semantic interference effect arises in lexical response selection (e.g., Abdel Rahman & Melinger, 2009; Hantsch et al., 2005; Levelt et al., 1999; Roelofs, 1992, 2003; Starreveld & La Heij, 1996), the effect should not have been absorbed according to the response-selection bottleneck account (Pashler, 1994, 1998; Pashler & Johnston, 1989), contrary to what Dell'Acqua et al. and Ayora et al. observed.

However, although very influential, the assumption of a structural response-selection bottleneck is debatable (e.g., Fan et al., 2012; Hübner & Lehle, 2007; Israel & Cohen, 2011; Kahneman, 1973; Karlin & Kestenbaum, 1968; Lehle & Hübner, 2009; Leonhard & Ulrich, 2011; Meyer & Kieras, 1997a,b; Miller, Ulrich, & Rolke, 2009; Navon & Miller, 2002; Pannebakker et al., 2011; Schumacher et al., 1999, 2001; Schvaneveldt, 1969; Szameitat, Lepsien, von Cramon, Sterr, & Schubert, 2006; Szameitat, Schubert, Müller, & von Cramon, 2002; Tombu & Jolicœur, 2003). According to alternative accounts of dual-task performance, response selection for two tasks may occur in parallel depending on the strategic choice of participants concerning the amount of overlap between tasks (cf. Logan & Gordon, 2001; Meyer & Kieras, 1997a,b; Navon & Miller, 2002; Tombu & Jolicœur, 2003). Under one such strategic account, henceforth the *active scheduling account* (cf. Kondo et al., 2004; Logan & Gordon, 2001; Luria & Meiran, 2003; Meyer & Kieras, 1997a,b; Szameitat et al., 2006, 2002), participants determine the order of the tasks, monitor progress on the tasks, suspend picture naming before or after response selection (depending on the amount of overlap allowed between tasks) and hold it in working memory, and determine when to resume picture naming (Piai et al., 2011; Roelofs, 2007, 2008a; Roelofs & Piai, 2011). If response selection in the tone discrimination task temporally overlaps with response selection in picture naming, semantic interference arising during response selection in picture naming may be resolved while response selection in the tone discrimination task is in progress. As a consequence, the semantic interference effect will disappear in concurrent task performance, as empirically observed by Dell'Acqua et al. and Ayora et al.

An important difference between the passive queuing and active scheduling accounts concerns the ability of monitoring and updating of working memory, often referred to as *working memory capacity* (e.g., Conway et al., 2005; Kane & Engle, 2002; Miyake et al., 2000). This ability is central to the active scheduling account (Kondo et al., 2004; Logan & Gordon, 2001; Meyer & Kieras, 1997a,b; Szameitat et al., 2006, 2002) but not to the passive queuing account (Adcock et al., 2000; Bunge et al., 2000; Dux et al., 2009; Jiang et al., 2004; Pashler, 1984, 1994, 1998; Pashler & Johnston, 1989). This implies that individual differences in working memory capacity should affect dual-task performance under the active scheduling account, but not under the passive queuing account. In the present article, we report an experiment that assessed whether dual-task interference from tone discrimination on picture naming depends on individual differences in working memory capacity.

The remainder of the article is organised as follows. First, we describe the findings of Dell'Acqua et al. (2007) and Ayora et al. (2011), and the queuing and scheduling accounts in more detail. Next, we derive predictions concerning the influence of working memory capacity on dual-task performance and report a dual-task experiment, involving tone discrimination and PWI tasks, designed to test these predictions.

1.1. Passive queuing and active scheduling accounts

In their studies of picture naming during dual-task performance, Dell'Acqua et al. (2007) and Ayora et al. (2011) employed the widely used psychological refractory period (PRP) procedure, in which participants have to make two different responses to corresponding stimuli (Tasks 1 and 2), with the common instruction that the Task 1 response should precede the Task 2 response. The stimuli are presented with certain stimulus onset asynchronies (SOAs), which

usually range from 0 to 1000 ms. A typical finding is that Task 2 response time (RT2) is longer at short (e.g., 0 ms) than at long SOAs (e.g., 1000 ms), reflecting dual-task interference. In the studies of Dell'Acqua et al. and Ayora et al., participants performed a manual tone discrimination task (Task 1) and a PWI task (Task 2, picture naming while ignoring written distractor words) with varying SOAs. In the experiment of Dell'Acqua et al., the tones preceded the picture–word stimuli by SOAs of 100, 350 or 1000 ms. Dell'Acqua et al. observed that the semantic interference effect was much smaller at the 350-ms SOA than at the 1000-ms SOA and that the effect was absent at the 100-ms SOA. These findings were replicated by Ayora et al. using the SOAs of 100 and 1000 ms. These results suggest that, at the short SOA of 100 ms, the semantic interference from the picture–word stimuli was resolved while performing the tone discrimination task.

Following (Pashler, 1984, 1994, 1998; Ferreira & Pashler, 2002; Pashler & Johnston, 1989), Dell'Acqua et al. and Ayora et al. assumed that dual-task interference arises from a structural response-selection bottleneck. According to this passive queuing account, response selection in the tone discrimination task precludes concurrent response selection in picture naming. Dell'Acqua et al. argued that, if semantic interference in picture naming arises in response selection, then the semantic interference and SOA effects should be additive, that is, the magnitude of the semantic interference effect should not differ between long and short SOAs, contrary to what they empirically observed. Therefore, under the assumption of a response-selection bottleneck, the finding that the semantic interference effect disappears at short SOAs in dual-task performance suggests that the locus of the semantic interference effect is prior to the response selection stage.

Other accounts of PRP performance assume no such structural response-selection bottleneck, but maintain that simultaneous selection of two responses in dual-task performance is possible depending on a strategic choice regarding the amount of parallelism allowed between Task 1 and Task 2 (cf. Kahneman, 1973; Kondo et al., 2004; Logan & Gordon, 2001; Meyer & Kieras, 1997a,b; Navon & Miller, 2002; Tombu & Jolicœur, 2003). The assumption that dual-task performance is not constrained by a structural response-selection bottleneck is supported by several findings (e.g., Fan et al., 2012; Hübner & Lehle, 2007; Israel & Cohen, 2011; Kahneman, 1973; Karlin & Kestenbaum, 1968; Lehle & Hübner, 2009; Leonhard & Ulrich, 2011; Logan & Gordon, 2001; Meyer & Kieras, 1997a,b; Navon & Miller, 2002; Pannebakker et al., 2011; Schvaneveldt, 1969; Szameitat et al., 2006, 2002; Tombu & Jolicœur, 2003). For example, there is evidence that the magnitude of dual-task interference can be eliminated with practice (e.g., Schumacher et al., 1999, 2001, but see Ruthruff, Van Selst, Johnston, & Remington, 2006; Van Selst, Ruthruff, & Johnston, 1999) or by increasing the number of short SOAs in an experimental block of trials (Miller et al., 2009). The malleability of dual-task interference should not be observed under a structural response-selection bottleneck.

Various studies have suggested that an executive control system operates during dual-task performance in order to coordinate or supervise processing (e.g., Fan et al., 2012; Kondo et al., 2004; Leonhard & Ulrich, 2011; Logan & Gordon, 2001; Luria & Meiran, 2003; Meyer & Kieras, 1997a,b; Schumacher et al., 1999, 2001; Szameitat et al., 2006, 2002). Such an account "assumes that PRP phenomena result from scheduling and control strategies enacted by a central executive, whereas RSB [response-selection bottleneck] theory says nothing about such strategies and explains PRP phenomena without them" (Logan & Gordon, 2001, p. 395). Under an active scheduling account of concurrent tone discrimination and picture naming, participants strategically postpone particular stages of picture–word processing until particular stages of tone processing have been finished (e.g., Lamers & Roelofs, 2011; Piai et al., 2011; Roelofs, 2007, 2008a; Roelofs & Piai, 2011). To this end, participants set a point during Task 2 performance at which processing is strategically suspended, which is typically before or after response selection in Task 2. Moreover, they set a point during Task 1 performance at which Task 2 processing is resumed, which is

typically after response selection in Task 1. The suspension and resumption points serve to maintain acceptable levels of speed and accuracy, to minimise resource consumption and crosstalk between tasks, and to satisfy instructions about task priorities (i.e., the Task 1 response should precede the Task 2 response). To determine when the suspension and resumption points are reached and working memory has to be updated, participants monitor progress on Tasks 1 and 2 throughout the trial. If the semantic interference effect in picture naming arises in response selection and participants postpone post-selection processes in PWI until response selection for the tone task has been completed, then the semantic interference and SOA effects should be underadditive. The semantic interference effect should be maximal at long SOAs and decrease at shorter SOAs. This is what Dell'Acqua et al. (2007) and Ayora et al. (2011) observed. However, if participants are more conservative and postpone response selection in the PWI task until response selection for the tone task has been completed, then the semantic interference and SOA effects should be additive. That is, the magnitude of the semantic interference effect should not differ between short and long SOAs. This is what Schnur and Martin (2012) and Piai et al. (submitted for publication-a) observed. The utility of the active scheduling account of dual-task performance was demonstrated in computer simulations with the WEAVER++ model of word planning (e.g., Levelt et al., 1999; Roelofs, 1992, 2003) extended by assumptions about active scheduling from Meyer and Kieras (1997a,b). We refer to Piai et al. (2011), Roelofs (2007, 2008a), and Roelofs and Piai (2011) for reports on these simulations.

1.2. Role of monitoring and updating of working memory

An important difference between passive queuing and active scheduling accounts of dual-task performance concerns the role of working memory, which is an important component of the executive functions required for complex cognition (e.g., Miyake et al., 2000). According to passive queuing accounts, executive functions in general, or monitoring and updating of working memory in particular, do not contribute to dual-task interference (e.g., Adcock et al., 2000; Bunge et al., 2000; Dux et al., 2009; Jiang et al., 2004; Pashler, 1984, 1994, 1998; Pashler & Johnston, 1989). For example, Adcock et al. argued that their fMRI data "provide no evidence for a neural locus for executive function, in the prefrontal cortex or elsewhere, selectively activated by dual-task performance". Instead, their findings were taken to support the view that "various specialized information-processing systems in the human brain may, by their interplay, accomplish the regulation of complex operations such as multitasking" (p. 3572). Similarly, Jiang et al. argued that "the amount of increased central executive control during the short-SOA condition is relatively small and that PRP queuing is primarily passive" (p. 395). According to this passive queuing account, when the central response-selection mechanism is busy selecting a response for the tone task (Task 1), response selection for picture naming (Task 2) simply waits until the central response-selection mechanism is finished, causing dual-task interference at short SOAs. Fig. 1A illustrates this passive queuing account.

In contrast, according to the active scheduling account, executive functions in general, and working memory in particular, play an important role in dual-task performance (cf. Kane & Engle, 2002; Kondo et al., 2004; Logan & Gordon, 2001; Szameitat et al., 2006, 2002). According to Meyer and Kieras (1997b), active scheduling of tasks requires "monitoring the contents of working memory and inserting or deleting task goals and other control items at appropriate points along the way" (p. 751). The monitoring and updating is assumed to be achieved by condition-action rules (Meyer & Kieras, 1997a,b). Fig. 1B illustrates the timing of some of the major updating and monitoring operations at the 0-ms SOA, assuming a pre-selection suspension point for picture naming (Task 2) and resumption of picture naming after response selection in tone discrimination (Task 1). At the beginning of each trial, Tasks 1 and 2 are enabled by placing the

corresponding task goals in working memory. Next, progression of Tasks 1 and 2 is monitored. If pre-selection (i.e., perceptual and conceptual) processes are completed for Task 2, progression on this task is temporarily suspended. If the monitoring process assesses that response selection for Task 1 is completed, processing for Task 2 is resumed (the latter will take some time, hence the partial temporal overlap between resumption and Task 1 post-selection processes). Finally, both Tasks 1 and 2 run to completion. Under this account, the magnitude of dual-task interference from Task 1 on Task 2 is not only determined by the duration of Task 1 response selection, but also by the duration of updating and monitoring operations. According to Meyer and Kieras (1997b), "working memory for storing procedural information (e.g., task goals and sequential control notes) is conceivably limited" (p. 786). Therefore, under the active scheduling account, individual differences in working memory capacity may influence dual-task interference, whereas this is not expected under the passive queuing account.

According to the WEAVER++ model (Levelt et al., 1999; Roelofs, 1992, 2003, 2007, 2008a,b,c), updating/monitoring ability should not only affect dual-task performance, but also the speed of picture naming per se. Word planning in this model is achieved by spreading activation through a lexical network while condition-action rules determine what is done with the activated lexical information depending on the task goal in working memory (e.g., to name a picture). The task goal is achieved by sequentially updating subgoals in the course of the word planning process. In picture naming, these successive subgoals concern selection of a concept for the perceptual information, selection of a lemma for a selected concept (lemmas specify grammatical information, such as word class), selection of one or more morphemes for the selected lemma, selection of phonemes for the selected morphemes, syllabification of the selected phonemes and assignment of stress, and selection of syllable motor programmes for the syllabified phonemes. Under this account, the ability to monitor and update goals, subgoals, and intermediate representations in planning should affect the speed of picture naming (Roelofs & Piai, 2011). This prediction was recently confirmed in a study by Shao, Roelofs, and Meyer (2012), who observed correlations between updating/monitoring ability and picture naming RT. In particular, participants with higher updating/monitoring ability named the pictures faster. In sum, under the WEAVER++ model, updating/monitoring ability should not only affect the magnitude of dual-task interference, but also picture naming RTs per se.

1.3. The present study

The PRP experiment reported in the present article assessed whether individual differences in working memory capacity are related to Task 2 picture naming RTs and to the magnitude of dual-task interference. Participants performed a manual tone discrimination task (Task 1) and a PWI task (Task 2) with short and long SOAs. Dual-task interference may be reflected in two performance measures. First, in dual-task performance, RT2 is typically longer at short than at long SOAs, whereas Task 1 RT (RT1) is not affected. We refer to this difference in RTs between long and short SOAs as the *SOA effect* (cf. Pashler & Johnston, 1989). Second, picture naming RTs are longer than tone discrimination RTs, and more so at short than at long SOAs (e.g., Dell'Acqua et al., 2007; Piai et al., submitted for publication-a; Schnur & Martin, 2012). We refer to the difference between RT2 and RT1 (at each SOA) as the *task-difference effect*. In PRP performance, RT2 typically increases with increasing RT1, and more so at short than long SOAs, which is taken to reflect dual-task interference (cf. Dell'Acqua et al., 2007; Pashler, 1994). If the updating/monitoring ability affects both picture naming RTs and dual-task interference, there should be correlations between scores for updating/monitoring ability, on the one hand, and picture naming RT (RT2) and the task-difference effect (RT2 minus RT1), on the other hand, especially at the 0-ms SOA. At this short SOA, performance strongly depends on the updating/monitoring ability because of the temporal overlap between Task 1 and Task 2 processes. At the 1000-ms SOA,
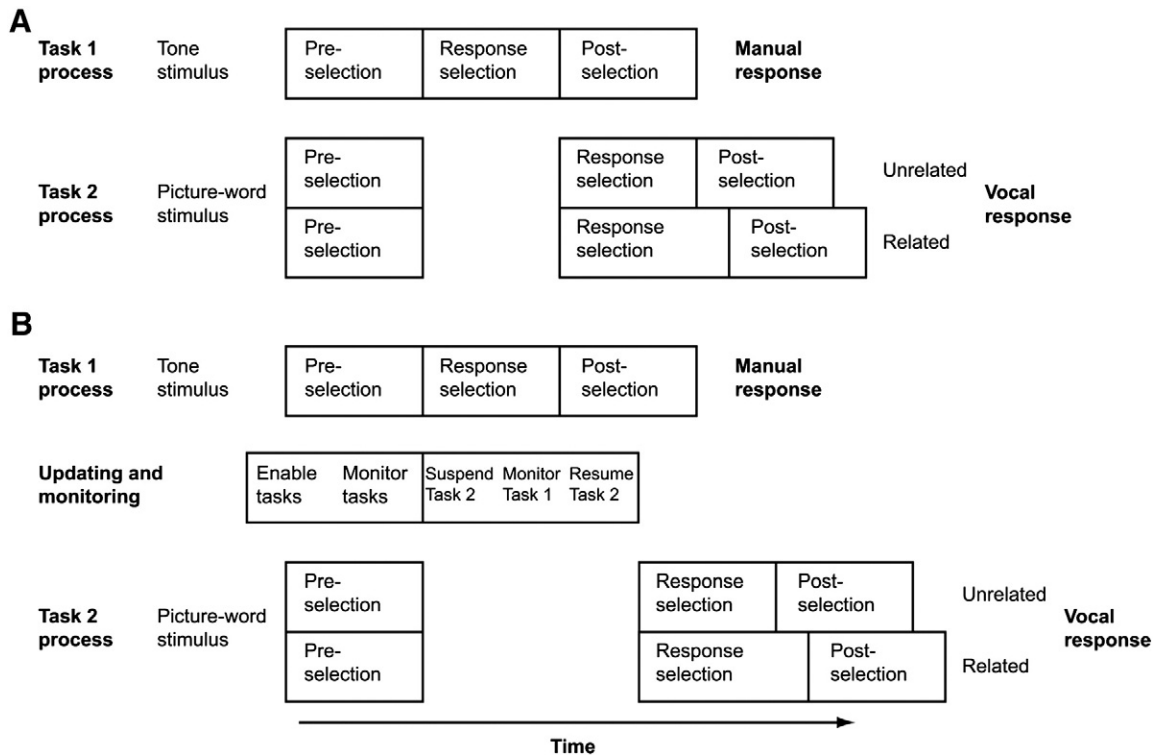
Fig. 1. Time course of processing during dual-task performance in the psychological refractory period paradigm at the 0-ms stimulus onset asynchrony (SOA). Processing for Tasks 1 and 2 begins with the presentation of the tone and picture–word stimuli, followed by pre-selection processes (i.e., perceptual and conceptual encoding), response selection, and post-selection processes (i.e., encoding of motor programmes), and ending with overt responses. Under the passive queuing account (A), dual-task interference reflects how long picture naming has to wait before response selection for tone discrimination is done. Under the active scheduling account (B), dual-task interference reflects the strategic choice concerning the amount of overlap permitted between the tasks and the duration of updating and monitoring operations.

performance is less dependent on the updating/monitoring ability because Task 1 and Task 2 processes overlap less or not at all. Note that the updating/monitoring ability affects RT2 not only at the short SOA but also to some extent at the long SOA. Consequently, the contribution of this ability may be substantially subtracted out in computing the SOA effect for RT2 (i.e., the difference between RTs at short and long SOAs). Therefore, the task-difference effect at the 0-ms SOA should be the most sensitive measure of the involvement of updating/monitoring abilities in dual-task performance. Consequently, we expect a correlation between updating/monitoring ability and the task-difference effect at the 0-ms SOA, but a weaker or no correlation at the 1000-ms SOA, and also a weaker or no correlation between updating/monitoring ability and the SOA effect. Finally, we did not expect the updating/monitoring ability to correlate with the magnitude of the semantic and Stroop-like effects. As can be seen in Fig. 1B, the updating and monitoring processes do not affect response selection in picture naming per se, which occurs after Task 2 has been resumed (and the critical monitoring and updating operations have taken place). Thus, the magnitude of the distractor effects should be independent of the monitoring and updating ability. These predictions were tested in the study reported below.[1]

---

[1] Note that two issues have to be carefully separated. The first issue concerns the source of dual-task interference: a passive response-selection bottleneck or active scheduling. The passive bottleneck account does not predict an effect of working memory capacity on the magnitude of dual-task interference, whereas an active scheduling account does predict such an effect. The second issue concerns the variability of semantic effects in the PRP paradigm. We argued that under an active scheduling account, semantic effects arising during lemma retrieval may be additive or underadditive with SOA depending on the strategic decision about the amount of overlap allowed between Tasks 1 and 2. This may explain the variability of findings in the literature (i.e., the underadditivity observed by Dell'Acqua et al. (2007) and the additivity observed by Schnur and Martin (2012) and by ourselves).

In our study, individual differences in updating/monitoring ability were assessed by the operation-span (Ospan) task (Conway et al., 2005; Miyake et al., 2000; Turner & Engle, 1989). In this task, participants verified series of simple mathematical operations (e.g., Is $3 \times 5 - 1 = 16$?) interleaved with unrelated words (e.g., dog) to memorise. Participants saw one operation–word pair at a time, and each series of operation–word pairs ranged from two to six items in length. For each pair, the participants read the operation aloud, indicated with a button press whether the operation was correct, and then read the word aloud. After a random number of operation–word pairs, a cue appeared indicating that the participants had to verbally recall all the words from the current series. Thus, the Ospan task requires participants to monitor and update working memory representations. Performance on this task measures working memory capacity, that is, the updating/monitoring ability (Miyake et al., 2000).

The SOAs between the tone discrimination task and the PWI task were 0 and 1000 ms. Following Piai, Roelofs, and Van der Meij (2012), we used three picture–word conditions. Picture and word were semantically related, unrelated, or identical. In the identical condition, the distractor word was the name of the picture (e.g., the word dog combined with a picture of a dog). Whereas the difference in performance between semantically related and unrelated trials reflects a semantic effect, the difference between semantic and identical trials indexes a Stroop-like effect, analogous to the difference between incongruent and congruent conditions in the classic colour–word Stroop task (e.g., the word red presented in incongruent green ink versus in congruent red ink). Whereas Dell'Acqua et al. (2007) and Ayora et al. (2011) observed underadditive effects of semantic relatedness and SOA in the PWI task, Fagot and Pashler (1992) observed additive effects of Stroop condition (incongruent vs. congruent) and SOA. Given our earlier results (Piai et al., submitted for publication-a), we expected not only to find the Stroop-like effect to be additive with SOA (replicating Fagot and Pashler), but also to obtain additive effects of semantic

relatedness and SOA (replicating our earlier findings and those of Schnur & Martin, 2012).

2. Method

2.1. Participants

Thirty native speakers of Dutch (5 male, mean age 20.1 years) from the participant pool of Radboud University Nijmegen participated in the experiment for monetary compensation. All participants reported having normal or corrected-to-normal vision, and normal hearing.

2.2. Picture–word interference task

2.2.1. Materials and design

Thirty-two pictures of common objects were selected from the picture data base of the Max Planck Institute for Psycholinguistics, Nijmegen. The objects belonged to eight different semantic categories. There were three distractor types: Distractors from the same semantic category as the pictures (semantic condition), semantically unrelated distractors (unrelated condition), and distractors that were the basic-level names of the pictures (identical condition). The semantically related and unrelated picture–word stimuli were created by re-combining the words from the identical condition with other pictures. Thus, all distractors belonged to the response set. A list of the materials can be found in Appendix A.

The distractor words occupied on average $2.8° \times 0.9°$ and the pictures $5.7° \times 5.7°$ of visual angle at a viewing distance of approximately 60 cm. The stimuli for the tone task were two pure tones of 300 Hz (low tone) and of 800 Hz (high tone) with a duration of 300 ms. The SOA values used were 0 ms and 1000 ms, which varied randomly from trial to trial. Each picture–word stimulus appeared once with each tone at each SOA, totalling 384 trials. Trials were randomised using Mix (Van Casteren & Davis, 2006) with one unique list per participant. The randomisation was constrained to contain at most three consecutive trials with the same tone, the same distractor type, or the same SOA.

2.2.2. Procedure and apparatus

The presentation of stimuli and the recording of responses were controlled by Presentation Software (Neurobehavioral Systems, Albany, CA). The vocal responses were measured with a voice key and the tones were presented via closed headphones. The button box was designed with force-sensitive resistors so that button presses were silent and did not trigger the voice key. Participants were instructed to press the left or right button to respond to the tones (left button for the low tone and right button for the high tone), and to name the pictures and to try to ignore the distractor words. Priority to the manual task was emphasised, as well as speed and accuracy for both manual and naming responses. Participants were familiarised with the pictures and their names using a booklet. Next, they were familiarised with the tones. A practice block of twelve trials followed, with different materials from the experimental ones. Then, the experiment proper was administered.

At 0-ms SOA, a trial began with the picture–word stimulus and the tone being presented with the same onset. At 1000-ms SOA, the tone was presented first, followed by the picture–word stimulus. The picture–word stimulus always remained on the screen for 1250 ms, followed by a black screen for 1750 ms. Manual responses were measured from tone onset and naming responses from picture–word stimulus onset.

2.2.3. Analysis

First, we discarded all trials in which participants named the picture before classifying the tone, trials with manual RTs shorter than 100 ms or with naming RTs shorter than 200 ms, and trials in which the voice key was triggered by a non-response sound. Next, trials with incorrect tone classification and trials with naming responses containing a disfluency or a wrong response word were coded as errors and discarded

from all the subsequent RT analyses. Manual RTs were submitted to by-participant (F_1) repeated-measures ANOVAs with distractor type (semantic, unrelated, and identical) and SOA (0 and 1000 ms) as within-participant and within-item variables. Naming RTs were additionally submitted to by-item (F_2) repeated-measures ANOVAs. Planned contrasts were analysed with two-tailed dependent t-tests. Errors were submitted to logistic regression analyses with distractor type and SOA as predictors for each task separately. For completeness, Cohen's d is provided for the Stroop-like (semantic vs. identical) and semantic interference (semantic vs. unrelated) effects in the RTs, and 95% confidence intervals are provided additionally for the effects in the naming RTs.

2.3. Operation-span task

2.3.1. Design, procedure and analysis

The Ospan task used was similar to the version reported by Engle, Cantor, and Carullo (1992), with 60 mathematical operations paired with 60 Dutch words. On each trial, our participants read aloud a mathematical operation (e.g., Is $3 \times 5{-}1 = 16$?) presented in the centre of the screen and indicated its correctness by means of a button press (2 buttons: correct, incorrect). Then they read aloud a word, which was presented below the operation. After a randomly varying number of trials (between 2 and 6), a recall prompt appeared. Participants recalled the words orally and the experimenter wrote down their responses. There was one randomisation for each participant. For more details about the procedure, see Engle et al. (1992) and Conway et al. (2005). The Ospan task was administered after the PWI session.

Partial-credit unit scoring was used for establishing participants' Ospan scores (cf. Conway et al., 2005). That is, we calculated the mean proportion of correctly recalled words for each recall block using the same load weight regardless of the block length (i.e., recalling one word in a two-word block weighs as much as recalling three words out of six).

2.4. Analyses of individual differences

We investigated individual differences in dual-task performance by relating Ospan scores to RTs. For that purpose, we computed Pearson's product–moment correlations to examine the association between Ospan scores and mean RT1 and RT2, averaged over distractor types, and between Ospan scores and dual-task interference. As outlined previously, dual-task interference was operationalised in two ways. Firstly, we computed participants' mean RTs, averaged over distractor types, at the 0-ms SOA minus RTs at the 1000-ms SOA for each task separately, which is the SOA effect. Usually, RT1 does not differ much between short and long SOAs. However, RT2 tends to be much longer at short than at long SOAs, which is the SOA effect for Task 2. Secondly, we computed the task-difference effect for each participant by subtracting RT1 from RT2 on a trial-by-trial basis, averaged over distractor types. Task 2 picture-naming RTs are usually longer than Task 1 tone-discrimination RTs, and this difference between tasks is larger at short than at long SOAs (see, e.g., Fig. 2 of Dell'Acqua et al., 2007, and Fig. 5 of Pashler, 1994). As previous research has shown (e.g., Dell'Acqua et al., 2007; Pashler, 1994), the difference between RT2 and RT1 at short SOAs (e.g., 0-ms SOA) not only reflects the difference in RT between the tasks per se (as long SOAs such as the 1000-ms SOA mainly do), but the task-difference at short SOAs also reflects the extent to which Task 1 delays the performance of Task 2. Since the task-difference effect is calculated for each SOA separately, it is more likely to be sensitive to variations in updating/monitoring ability than the SOA effect. The latter is calculated by comparing a condition sensitive to the updating/monitoring ability (the 0-ms SOA) and a condition also sensitive to it but presumably somewhat less so (the 1000-ms SOA), so that the influence of updating/monitoring ability is likely to be subtracted out substantially. Finally, we examined the association between participants' Ospan scores and the magnitude of the semantic interference

and Stroop-like effects. Since multiple correlations were tested, we used Bonferroni corrected p-values. For each correlation between two dependent variables (RTs and Ospan scores), corrections were based on the number of comparisons involving an independent variable (SOA and distractor type).

To test whether Ospan scores contribute to dual-task performance beyond general processing speed (reflected in absolute RTs), we conducted linear regression analyses and compared the resulting models using ANOVAs. For each SOA separately, a simple model was fit with only RT2 as a predictor of the task-difference effect and a second model was fit to which Ospan scores were added as a predictor. The models were fit for the task-difference effect only because this measure allows us to assess dual-task interference for each SOA separately.

3. Results

Fig. 2 shows the RTs and Table 1 shows the error rates for the manual (Task 1) and naming (Task 2) responses as a function of SOA and distractor type. One participant was excluded from all analyses because his/her accuracy in the verification of the operations in the Ospan task was below 85%.

3.1. Manual responses (tone task)

The log-odds of incorrect manual responses increased with a factor of 1.1 for the 0-ms SOA relative to the 1000-ms SOA, β coefficient = -0.001, $S.E. = 0.001$, $Wald\ Z = -8.42$, $p < .001$. For the RTs, there was a main effect of SOA, $F_1(1,28) = 53.08$, $p < .001$, indicating that responses were slower at the 0-ms than at the 1000-ms SOA. There was also a main effect of distractor type, $F_1(2,56) = 9.38$, $p < .001$. Distractor type and SOA interacted, $F_1(2,56) = 4.29$, $p = .018$, revealing that the magnitude of the interference effects differed between SOAs. The Stroop-like effect (i.e., semantic vs. identical) was significant at the 0-ms SOA, $t(28) = 2.10$, $p = .045$, $d = .04$; and at the 1000-ms SOA, $t(28) = 3.99$,
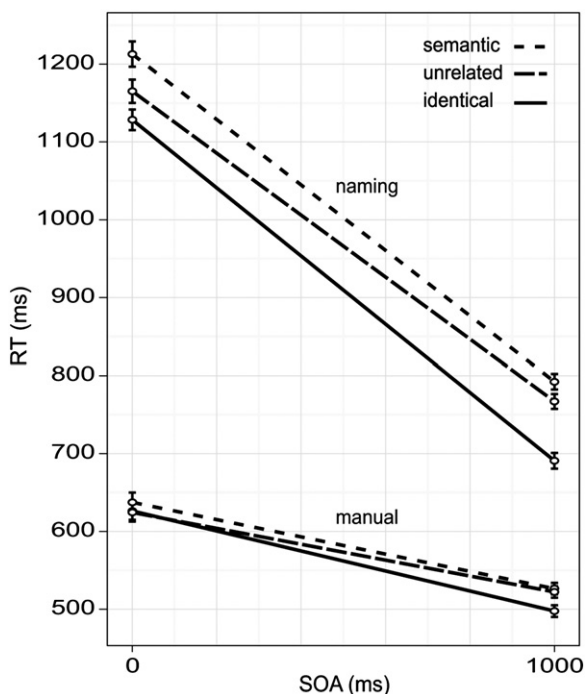

Fig. 2. Manual (Task 1) and naming (Task 2) response times (RTs) as a function of SOA and distractor type. Error bars indicate 95% confidence intervals. SOA = stimulus onset asynchrony. RT = response time.

Table 1

Error rates for the tone discrimination (Task 1) and picture naming (Task 2) responses as a function of SOA and distractor type.

| Distractor/SOA	Task 1		Task 2	
	0 ms	1000 ms	0 ms	1000 ms
Semantic	4.5	1.2	4.5	3.6
Unrelated	4.2	1.2	3.1	2.4
Identical	4.9	1.3	0.4	1.5

$p < .001$, $d = .20$; whereas the semantic interference effect (i.e., semantic vs. unrelated) was significant at the 0-ms SOA, $t(28) = 2.27$, $p = .031$, $d = .05$; but not at the 1000-ms SOA, $t(28) < 1$, $d = .02$. Thus, distractor effects from the picture–word stimuli for the picture naming task were reflected in the tone discrimination RTs.

3.2. Naming responses (PWI)

The log-odds of an incorrect response in the semantic condition were 1.5 times higher than in the unrelated condition, β coefficient = 0.41, $S.E. = 0.14$, $Wald\ Z = 2.96$, $p = .003$; and 5 times higher than in the identical condition, β coefficient = 1.63, $S.E. = 0.22$, $Wald\ Z = 7.42$, $p < .001$. For the RTs, there was a main effect of SOA, $F_1(1,28) = 311.55$, $p < .001$, $F_2(1,31) = 5644.50$, $p < .001$, confirming that responses were slower at the 0-ms than at the 1000-ms SOA; and there was a main effect of distractor type, $F_1(2,56) = 64.00$, $p < .001$, $F_2(2,62) = 67.89$, $p < .001$. Distractor type and SOA interacted, $F_1(2,56) = 3.35$, $p = .042$, $F_2(2,62) = 6.05$, $p = .004$, indicating that the magnitude of the interference effects was different between SOAs. The Stroop-like effect was significant at the 0-ms SOA, $t_1(28) = 8.30$, $p < .001$, $t_2(31) = 7.60$, $p < .001$, 95% CI [64, 110], $d = .28$; and at the 1000-ms SOA, $t_1(28) = 8.20$, $p < .001$, $t_2(31) = 9.78$, $p < .001$, 95% CI [69, 114], $d = .56$. Also, the semantic interference effect was significant at the 0-ms SOA, $t_1(28) = 6.24$, $p < .001$, $t_2(31) = 5.133$, $p < .001$, 95% CI [32, 64], $d = .15$; and at the 1000-ms SOA, $t_1(28) = 3.68$, $p < .001$, $t_2(31) = 2.66$, $p = .012$, 95% CI [11, 38], $d = .14$. However, whereas the magnitude of the Stroop-like interference did not differ between SOAs, $t_1(28) < 1$, $t_2(31) = -1.52$, $p = .138$, the semantic interference effect was larger at the 0-ms SOA than at the 1000-ms SOA, $t_1(28) = 2.30$, $p = .029$, $t_2(31) = 2.04$, $p = .050$.

3.3. Individual differences

Ospan scores ranged between .54 and .94 (mean = .76). Fig. 3 shows the correlations between Ospan scores and RT2 at the 0-ms SOA (upper left) and 1000-ms SOA (upper right). Ospan scores correlated with RT2 at the 0-ms SOA, $r = -.436$, $p = .036$; and at the 1000-ms SOA, $r = -.515$, $p = .008$ (p-values corrected for two comparisons). Thus, the higher the participants' Ospan scores, the faster they named the pictures. Ospan scores did not correlate with RT1 at any SOA, $rs < -.292$, $ps > .248$ (corrected for two comparisons). Ospan scores did not correlate with the SOA effect either for RT1, $r = -.065$, $p = .738$; nor for RT2, $r = -.262$, $p = .338$. Fig. 3 also shows the correlation between Ospan scores and the task-difference effect at the 0-ms (lower left) and 1000-ms SOAs (lower right). The task-difference effect only marginally correlated with Ospan scores at the 1000-ms SOA, $r = -.377$, $p = .088$, but it did correlate significantly at the 0-ms SOA, $r = -.479$, $p = .018$ (p-values corrected for two comparisons). So, the delay in performing Task 2 in the context of Task 1 was greater with decreasing Ospan scores at the 0-ms SOA but less so at the 1000-ms SOA. The task-difference effect did not correlate with RT1 at any SOA, $rs < .250$, $ps > .191$. Moreover, the task-difference effect did not correlate with RT2 (i.e., the absolute picture naming RT) at the 0-ms SOA, $r = .333$, $p = .154$, although it correlated with RT2 at the 1000-ms
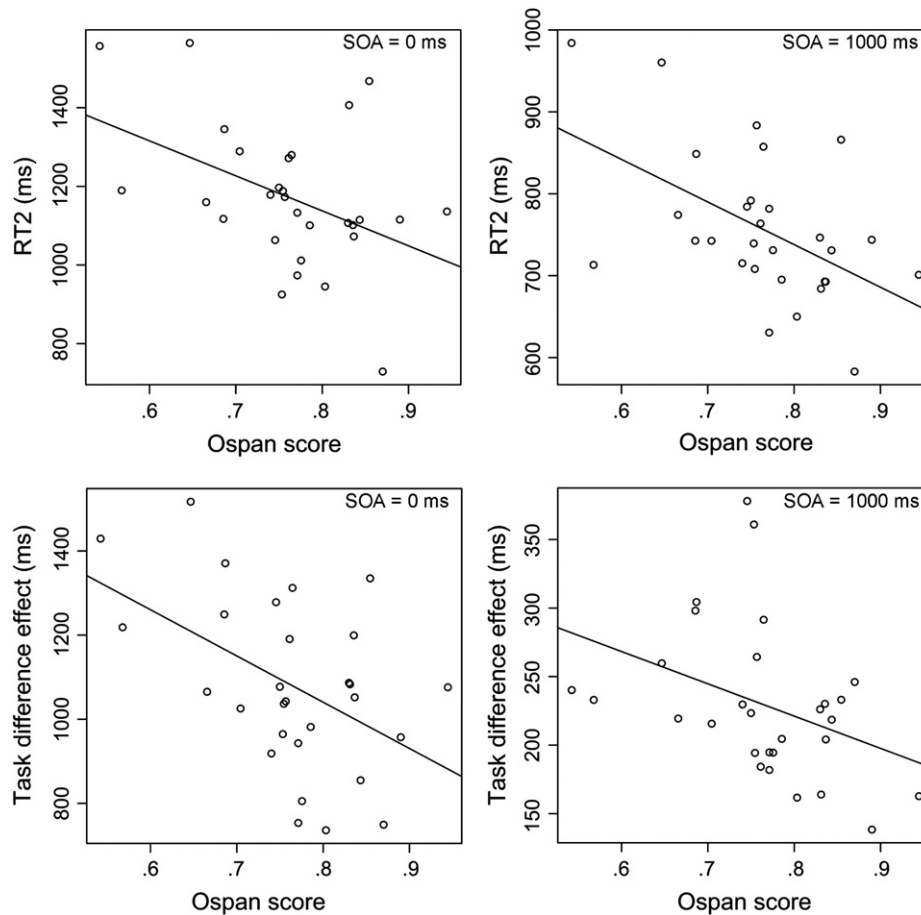
Fig. 3. Scatter plot of the relationship between Ospan scores and RT2 at the 0-ms SOA (upper left) and 1000-ms SOA (upper right) and between Ospan scores and the task-difference effect (RT2 − RT1) at the 0-ms SOA (lower left) and 1000-ms SOA (lower right). Each point represents one participant. The regression lines were derived from fitting a linear model for RT2 (upper row) or the task-difference effect (lower row) as a function of Ospan scores. Ospan = operation span; SOA = stimulus onset asynchrony.

SOA, $r = .482$, $p = .016$ (p-values corrected for two comparisons). This suggests that the task-difference effect at the 0-ms SOA is capturing an aspect of performance that is more than the general speed of responding.

We also assessed whether Ospan scores correlated with the magnitude of the semantic interference and Stroop-like effects. The results showed that Ospan scores did not correlate with any effect at any SOA: Stroop-like effect at the 0-ms SOA, $r = -.406$, $p = .116$; Stroop-like effect at the 1000-ms SOA, $r = -.236$, $p = .872$; semantic interference effect at the 0-ms SOA, $r = .001$, $p = .996$; semantic interference effect at the 1000-ms SOA, $r = .025$, $p = .896$ (p-values corrected for four comparisons).

Since the task-difference effect at the 0-ms SOA shows the highest association with individual differences in Ospan, we wanted to further ensure that Ospan scores were adding explanatory value to this dual-task measure beyond general speed of responding. The task-difference measure and RT2 were normally distributed. The collinearity between Ospan scores and RT2 was 24 at both the 0-ms and the 1000-ms SOAs. To remove this collinearity, we residualised RT2 against Ospan scores by fitting a regression model to RT2 as a function of Ospan scores for each SOA separately. The residuals (the distance between the observed and the fitted values) of these models were taken as the picture naming RT (RT2) in the subsequent linear regression models. The residualised RT2 still correlated strongly with the original RT2 at both the 0-ms SOA ($r = .90$) and the 1000-ms SOA ($r = .86$), whereas it no longer correlated with Ospan scores at the 0-ms and 1000-ms SOAs (both $rs < .01$).

Linear regression analyses were conducted for each SOA separately. A simple model was fit with only the residualised RT2 as a predictor of the task-difference effect and a second model was fit to which Ospan scores were added as a predictor. Table 2 shows the results of the regression analyses for the two models for each SOA separately, one based on the residualised RT2 (RT2 model) and a second model to which Ospan scores were added (Ospan model). At the 0-ms SOA, the Ospan model had superior explanatory power, $F(1,27) = 14.69$, $p < .001$, showing that Ospan scores contribute to the estimation of participants' task-difference effect at the 0-ms SOA beyond general naming speed. Whereas general naming speed accounts for 34.2% of the variance in the task-difference effect, adding Ospan scores to the model increases its explanatory power by 22.15%, accounting for 56.3% of the total variance. At the 1000-ms SOA, the Ospan model had superior explanatory power, $F(1,27) = 4.51$, $p = .043$, although the total explained variance was small: General naming speed accounted for 0.5% of the variance and Ospan scores added 11.5% of explained variance, but the model was still only able to explain 12% of the variance of the task-difference effect.

To summarise, we expected working memory capacity to correlate with picture naming RTs as well as with the task-difference effect (especially at the 0-ms SOA), but not or less so with the SOA effect (because in this effect, the contribution of working memory capacity is likely to be subtracted out). These predictions were confirmed by our data. Moreover, we expected no correlation between working memory capacity and the magnitude of the semantic and Stroop-like effects, again as observed.

Table 2
Linear regression models for the task-difference effect (RT2 − RT1) at the 0-ms (top) and 1000-ms (bottom) SOAs as a function of RT2 (RT2 model) and as a function of RT2 and Ospan scores (Ospan model). RT = response time. SOA = stimulus onset asynchrony. Adj. = Adjusted. RSS = Residual sum of squares. Resid. = Residualised.

Model	Unstandardised β	t-value	p-value	Adj. R²	RSS	df
RT2 model 0-ms SOA				.342	83.31	1, 27
Intercept	539.81	34.89	<.001			
Resid. RT2	.378	3.94	<.001			
Ospan model 0-ms SOA				.563	67.86	2, 26
Intercept	960.01	8.70	<.001			
Resid. RT2	.378	4.84	<.001			
Ospan	−549.74	−3.83	<.001			
RT2 model 1000-ms SOA				.005	55.73	1, 27
Intercept	229.55	22.18	<.001			
Resid. RT2	.145	1.07	=.294			
Ospan model 1000-ms SOA				.120	52.42	2, 26
Intercept	409.51	4.80	<.001			
Resid. RT2	.145	1.14	=.266			
Ospan	−235.43	−2.13	=.043			

4. Semantic interference in dual-task performance

Previous studies observed that the semantic interference effect was smaller at short than at long SOAs (Ayora et al., 2011; Dell'Acqua et al., 2007), or equally large at short and long SOAs (Piai et al., submitted for publication-a; Schnur & Martin, 2012). Van Maanen, Van Rijn, and Taatgen (2012) observed that semantic interference decreased with decreasing SOA in one experiment, but not in another. In the present experiment, the semantic interference effect was somewhat larger at the short than at the long SOA, whereas the magnitude of the Stroop-like effect did not differ between SOAs (cf. Fagot & Pashler, 1992). It is unclear why the magnitude of the semantic interference effect was somewhat larger at the short than at the long SOA in the present experiment. In a series of six experiments in our lab (including experiments using the same materials as the present experiment), the magnitude of the semantic interference and Stroop-like effects did not differ between short and long SOAs (Piai et al., submitted for publication-a). This corresponds to two experiments in another lab (using different materials but a procedure similar to the one used in the present study), in which the semantic interference effect was also additive with the SOA effect (Schnur & Martin, 2012).

To confirm that the present finding of a larger semantic interference effect at short than at long SOAs is exceptional, we conducted a new experiment with 14 novel participants, who performed the picture–word interference task as reported in the present study (see Section 2.2 for details) but with semantic and unrelated distractors only. For the RTs of the manual responses to the tones, no main effect of distractor type, SOA, or interaction were observed, all ps > .185. For the picture naming RTs, a main effect of SOA was obtained, $F_1(1,13) = 72.43$, $p < .001$, $F_2(1,31) = 1306$, $p < .001$; as well as a main effect of distractor type, $F_1(1,13) = 10.14$, $p = .007$, $F_2(1,31) = 23.79$, $p < .001$. Crucially, the semantic interference effect did not differ between the SOAs of 0 ms (56 ms) and 1000 ms (45 ms), as indicated by the lack of an interaction between distractor type and SOA, $F_1(1,13) = 2.47$, $p = .140$, $F_2(1,31) < 1$.

Altogether, these new observations plus those from the six experiments by Piai et al. (submitted for publication-a) and the two experiments by Schnur and Martin (2012) suggest that the present finding of a larger semantic effect at short than at long SOAs is an exceptional one. Other studies found less rather than more semantic interference at short than at long SOAs (Ayora et al., 2011; Dell'Acqua et al., 2007; Van Maanen et al., 2012). Importantly, working memory capacity did not correlate with the magnitude of the semantic and Stroop-like effects in the present study. Thus, the unusual finding of a larger semantic

effect at short than at long SOAs does not affect our main conclusion that dual-task interference depends on working memory capacity.

5. Discussion

In the present study, we examined how individual differences in updating/monitoring ability, as assessed by the Ospan task (Conway et al., 2005; Turner & Engle, 1989), relate to the magnitude of dual-task interference in the PRP procedure. We observed that the Ospan scores not only correlated with picture naming RTs (cf. Shao et al., 2012), but also with the difference in RTs between the tone-discrimination and picture-naming tasks (the task-difference effect), beyond general speed of responding (i.e., RT1 and RT2). The correlation between Ospan scores and the task-difference effect was especially prominent at the 0-ms SOA, which suggests that performance more strongly depends on the updating/monitoring ability at the 0-ms than at the 1000-ms SOA. Ospan scores did not correlate with the magnitude of the distractor-type effects at any SOA. In the remainder of this article, we discuss the theoretical implications of the present findings for the issue of whether dual-task interference from tone discrimination on picture naming reflects passive queuing or active scheduling. Moreover, we discuss whether the present findings are better explained by active scheduling or central capacity sharing, which are both strategic accounts that predict effects of working memory capacity.

5.1. Passive queuing versus active scheduling

As outlined previously, the dual-task interference from tone discrimination on picture naming has received different explanations in the literature. According to a passive-queuing account (e.g., Dell'Acqua et al., 2007), picture naming is passively queued as long as the central response-selection bottleneck is occupied by the tone-discrimination process (e.g., Pashler, 1984, 1994, 1998; Pashler & Johnston, 1989). Proponents of this account advocate that executive functions, such as monitoring and updating of working memory, do not contribute to dual-task interference (e.g., Adcock et al., 2000; Dux et al., 2009; Jiang et al., 2004). As Logan and Gordon (2001) stated, "Of course, RSB [response-selection bottleneck] theorists are aware of executive effects in the PRP (e.g., Pashler, 1994), but the theorists' awareness is not the issue. Rather, the issue is the constructs that the theorists include in their theories, and the RSB theories do not include executive scheduling and control strategies in their explanations of PRP phenomena" (p. 395). Under an active-scheduling account (e.g., Kondo et al., 2004; Lamers & Roelofs, 2011; Logan & Gordon, 2001; Meyer & Kieras, 1997a; Piai et al., 2011; Roelofs, 2007, 2008a; Roelofs & Piai, 2011; Szameitat et al., 2006, 2002), an executive control process actively coordinates Task 1 and Task 2 processes by monitoring progress on the tasks, suspending picture naming and holding it in working memory, and determining when to resume picture naming. This latter account predicts a relation between individual differences in updating/monitoring ability and dual-task interference, as observed in the present experiment. Thus, our results provide evidence for the active-scheduling account of dual-task interference in picture naming.

A number of observations support our claim that the correlation between Ospan scores and dual-task interference reflects more than a correlation with general speed of responding. First of all, adding Ospan scores to the regression model for the task-difference effect increases the adjusted R^2, especially at the 0-ms SOA, indicating that this measure of working memory capacity (together with RT2) is a better predictor of the task-difference effect than RT2 alone. Moreover, although the task-difference effect correlates with RT2 at the 1000-ms SOA, it does not correlate with RT2 at the 0-ms SOA. Thus, if the task-difference effect is reflecting speed of naming only, it should correlate most strongly with RT2 at the 0-ms SOA, but empirically, the strongest correlation is observed for the task-difference effect at the 1000-ms

SOA. These results suggest that there is a relation between Ospan scores and the ability to coordinate processing in the tone discrimination and PWI tasks (monitoring/updating) that goes beyond the speed of naming pictures only.

In the present study, Ospan scores correlated significantly with the mean picture naming RTs at both SOAs. In contrast, Shao et al. (2012) observed no significant correlation between Ospan sores and mean object naming RT (no distractor words presented), although such a correlation was obtained between Ospan scores and the tail of the RT distribution (reflecting slow responses). This finding suggests that the role of updating/monitoring ability is especially evident for difficult trials. In the present study, participants had to name objects in the context of a tone discrimination task, which made performance more difficult. Moreover, written distractor words were superimposed onto the pictured objects, making naming even more difficult. This may explain why Ospan scores correlated significantly with the mean object naming RTs in the present study, whereas Shao et al. observed correlations with the distribution tail only.

Further support for this difficulty-of-processing interpretation comes from an additional analysis of data from another experiment in our lab, which was conducted for different purposes and is reported elsewhere (Piai et al., submitted for publication-b). In this experiment, participants' Ospan scores were obtained in the exact same way as reported above. In that experiment, 40 participants decided on a trial-by-trial basis whether to name the picture of an object or read aloud the word of PWI stimuli depending on the pitch of a tone, which was presented simultaneously with or 1000 ms before picture–word onset. The task decision makes picture naming presumably more difficult than the naming of isolated pictures in the study of Shao et al. (2012). For the data set of Piai et al. (submitted for publication-b), a correlation was observed between participants' mean picture-naming RTs at the 0-ms SOA and their Ospan scores, $r = -.415$, $p = .008$, as shown in Fig. 4. These results suggest that the difficulty of the task situation determines to what extent the updating/monitoring ability is engaged and support flexible, rather than rigid, coordination of processes in dual-task performance (e.g., Miller et al., 2009).

To our knowledge, the present study is the first to examine whether a relation exists between the updating/monitoring ability and dual-task interference in the PRP procedure. Whereas we obtained a relation between updating/monitoring ability and dual-task performance, this was not obtained by Miyake et al. (2000). However, the dual-task procedure employed by Miyake et al. was one in which participants performed a spatial-scanning task and a word-generation task simultaneously for 3 min. Aggregate performance over the whole time period was assessed. This procedure is clearly different from the PRP procedure used in the present study. In particular, whereas Miyake et al. assessed continuous performance, we examined discrete performance. As argued extensively by Pashler (1998), studies of continuous performance provide little evidence about the time course of processing in the two tasks. This is because the aggregate measures of performance are simply too gross. In contrast, discrete performance, as assessed by the PRP procedure, allows for a detailed analysis of the temporal dimension of dual-task performance. So, the finding that dual-task performance did not relate to the updating/monitoring ability in Miyake et al.'s study does not challenge our conclusions concerning PRP performance.

5.2. Active scheduling versus central capacity sharing

In the present study, the semantic interference and Stroop-like effects were not only observed in the Task 2 naming RTs but also in the Task 1 tone discrimination RTs. The Stroop-like effect was present at both the short and long SOAs, whereas the semantic interference effect was present at the short SOA only. Moreover, the Task 1 RTs were longer at the short than at the long SOA. These SOA and distractor effects on Task 1 RTs are not expected under the passive queuing account (see Fig. 1A), but the effects may occur under the active scheduling account. The SOA effect on Task 1 RTs may be due to a greater dual-task overhead cost (cf. Roelofs, 2008a) or a more conservative Task 1 response criterion (cf. Meyer, Roelofs, & Levelt, 2003) at short compared to long SOAs. Alternatively, these RT1 effects may occur when central capacity is shared between the Task 1 and Task 2 processes (e.g., Kahneman, 1973; Lehle & Hübner, 2009; Navon & Miller, 2002; Tombu & Jolicœur, 2003).

A key prediction of central capacity sharing is that not only RT2 but also RT1 increases as SOA decreases, as observed in the present study. The amount of central capacity that is available for each participant may increase with increased working memory capacity. This would explain the present finding that dual-task performance depends on working memory capacity. However, if RT1 and RT2 increase with decreasing SOA because of central capacity sharing between Task 1 and Task 2 processes at short but not at long SOAs, working memory capacity should not only correlate with dual-task interference and RT2, as observed, but also with RT1. However, in the present study, no such correlation between working memory capacity and RT1 was obtained. This suggests that active scheduling provides a better account of the present findings than central capacity sharing.

6. Conclusions

To conclude, we observed that the ability to update and monitor information in working memory, as assessed by the operation-span task, was correlated with the speed of picture naming and with the magnitude of interference of tone discrimination on picture naming, beyond the speed of responding per se. These results lend support to active scheduling rather than passive queuing accounts of dual-task interference in language performance.
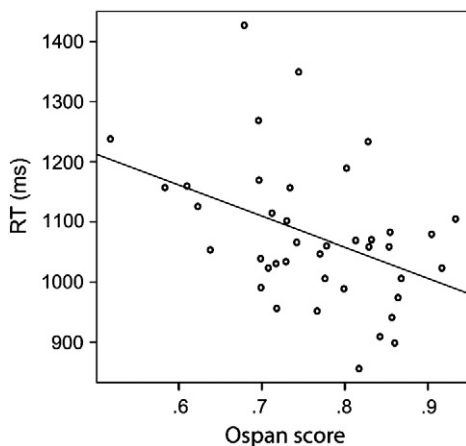
Acknowledgements

Fig. 4. Scatter plot of the relationship between Ospan scores and naming RTs at the 0-ms SOA for Experiments 1 and 2 reported in Piai et al. (manuscript submitted for publication-b). Each point represents one participant. The regression line was derived from fitting a linear model for naming RTs as a function of Ospan scores. Ospan = operation span; RT = response time; SOA = stimulus onset asynchrony.

Appendix A. Materials of the experiment (English translations are in-between parentheses)

	Picture name/identical distractor	Semantically related distractor	Unrelated distractor
Animals	hert (deer)	konijn	bureau
	konijn (rabbit)	hert	arm
	zwaan (swan)	geit	rok
	geit (goat)	zwaan	beker
Clothing	jas (jacket)	hemd	kasteel
	hemd (singlet)	jas	oor
	rok (skirt)	trui	zwaan
	trui (sweater)	rok	dolk
Transportation	auto (car)	bus	konijn
	bus (bus)	auto	glas
	trein (train)	fiets	kerk
	fiets (bicycle)	trein	kast
Buildings	kerk (church)	fabriek	been
	fabriek (factory)	kerk	neus
	molen (mill)	kasteel	kan
	kasteel (castle)	molen	jas
Weapons	dolk (dagger)	zwaard	trui
	zwaard (sword)	dolk	tafel
	kanon (cannon)	pistool	bord
	pistool (gun)	kanon	bed
Service	kan (pitcher)	beker	molen
	beker (cup)	kan	geit
	bord (plate)	glas	kanon
	glas (glass)	bord	bus
Furniture	bed (bed)	tafel	pistool
	tafel (table)	bed	zwaard
	bureau (desk)	kast	hert
	kast (wardrobe)	bureau	fiets
Body parts	neus (nose)	arm	fabriek
	arm (arm)	neus	trein
	been (leg)	oor	auto
	oor (ear)	been	hemd

References

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language & Cognitive Processes*, 24, 713–734.

Adcock, R. A., Constable, R. T., Gore, J. C., & Goldman-Rakic, P. S. (2000). Functional neuroanatomy of executive processes involved in dual-task performance. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 3567–3572.

Ayora, P., Peressotti, F., Alario, F. -X., Mulatti, C., Pluchino, P., Job, R., et al. (2011). What phonological facilitation tells about semantic interference: A dual-task study. *Frontiers in Psychology*, 2, 57.

Bunge, S. A., Klingberg, T., Jacobsen, R. B., & Gabrieli, J. D. E. (2000). A resource model of the neural basis of executive working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 3573–3578.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12, 769–786.

Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 345–361.

Dell'Acqua, R., Job, R., Peressotti, F., & Pascali, A. (2007). The picture–word interference effect is not a Stroop effect. *Psychonomic Bulletin & Review*, 14, 717–722.

Dux, P. E., Tombu, M. N., Harrison, S., Rogers, B. P., Tong, F., & Marois, R. (2009). Training improves multitasking performance by increasing the speed of information processing in human prefrontal cortex. *Neuron*, 63, 127–138.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 972–992.

Fagot, C., & Pashler, H. (1992). Making two responses to a single object: Implications for the central attentional bottleneck. *Journal of Experimental Psychology. Human Perception and Performance*, 18, 1058–1079.

Fan, Z., Singh, K., Muthukumaraswamy, S., Sigman, M., Dehaene, S., & Shapiro, K. (2012). The cost of serially chaining two cognitive operations. *Psychological Research*, 76, 566–578.

Ferreira, V., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 1187–1199.

Glaser, W. R., & Düngelhoff, F. -J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology. Human Perception and Performance*, 10, 640–654.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology. General*, 118, 13–42.

Hantsch, A., Jescheniak, J., & Schriefers, H. (2005). Semantic competition between hierarchically related words during speech planning. *Memory & Cognition*, 33, 984–1000.

Hübner, R., & Lehle, C. (2007). Strategies of flanker coprocessing in single and dual tasks. *Journal of Experimental Psychology. Human Perception and Performance*, 33, 103–123.

Israel, M., & Cohen, A. (2011). Involuntary strategy-dependent dual task performance. *Psychological Research*, 75, 513–524.

Jiang, Y., Saxe, R., & Kanwisher, N. (2004). Functional magnetic resonance imaging provides new constraints on theories of the psychological refractory period. *Psychological Science*, 15, 390–396.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review*, 9, 637–671.

Karlin, L., & Kestenbaum, R. (1968). Effects of number of alternatives on the psychological refractory period. *Quarterly Journal of Experimental Psychology*, 20, 167–178.

Kondo, H., Morishita, M., Osaka, N., Osaka, M., Fukuyama, H., & Shibasaki, H. (2004). Functional roles of the cingulo-frontal network in performance on working memory. *NeuroImage*, 21, 2–14.

Lamers, M., & Roelofs, A. (2011). Attention and gaze shifting in dual-task and go/no-go performance with vocal responding. *Acta Psychologica*, 137, 261–268.

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, 707–726.

Leonhard, T., & Ulrich, R. (2011). Determinants of central processing order in psychological refractory period paradigms: Central arrival times, detection times, or preparation? *Quarterly Journal of Experimental Psychology*, 64, 2012–2043.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences*, 22, 1–38.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, 108, 393–434.

Luria, R., & Meiran, N. (2003). Online order control in the psychological refractory period paradigm. *Journal of Experimental Psychology. Human Perception and Performance*, 29, 556–574.

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

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review*, 104, 749–791.

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, 48, 131–147.

Miller, J., Ulrich, R., & Rolke, B. (2009). On the optimality of serial and parallel processing in the psychological refractory period paradigm: Effects of the distribution of stimulus onset asynchronies. *Cognitive Psychology*, 58, 273–310.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, a H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100.

Navon, D., & Miller, J. (2002). Queuing or sharing? A critical evaluation of the single-bottleneck notion. *Cognitive Psychology*, 44, 193–251.

Pannebakker, M. M., Jolicœur, P., van Dam, W. O., Band, G. P. H., Ridderinkhof, K. R., & Hommel, B. (2011). Mental rotation impairs attention shifting and short-term memory encoding: Neurophysiological evidence against the response-selection bottleneck model of dual-task performance. *Neuropsychologia*, 49, 2985–2993.

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology. Human Perception and Performance*, 10, 358–377.

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

Pashler, H. (1998). *The psychology of attention.* Cambridge, MA: MIT Press.

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology*, 41, 19–45.

Piai, V., Roelofs, A., & Schriefers, H. (2011). Semantic interference in immediate and delayed naming and reading: Attention and task decisions. *Journal of Memory and Language*, 64, 404–423.

Piai, V., Roelofs, A., & Van der Meij, R. (2012). Event-related potentials and oscillatory brain responses associated with semantic and Stroop-like interference effects in overt naming. *Brain Research*, 1450, 87–101.

Piai, V., Roelofs, A., & Schriefers, H. (Manuscript submitted for publication-a). *Locus of semantic interference in picture naming: Evidence from dual-task performance*.

Piai, V., Roelofs, A., & Schriefers, H. (Manuscript submitted for publication-b). *Task decision and semantic interference in picture naming*.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, 110, 88–125.

Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory and Language*, 57, 232–251.

Roelofs, A. (2008a). Attention, gaze shifting, and dual-task interference from phonological encoding in spoken word planning. *Journal of Experimental Psychology. Human Perception and Performance*, 34, 1580–1598.

Roelofs, A. (2008b). Dynamics of the attentional control of word retrieval: Analyses of response time distributions. *Journal of Experimental Psychology. General*, 137, 303–323.

Roelofs, A. (2008c). Tracing attention and the activation flow in spoken word planning using eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 353–368.

Roelofs, A., & Piai, V. (2011). Attention demands of spoken word planning: A review. *Frontiers in Psychology*, *2*, 307.

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*, 125–142.

Schnur, T. T., & Martin, R. (2012). Semantic picture–word interference is a postperceptual effect. *Psychonomic Bulletin & Review*, *19*, 301–308.

Schriefers, H., Meyer, A., & Levelt, W. J. M. (1990). Exploring the time-course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language*, *29*, 86–102.

Schumacher, E. H., Lauber, E. J., Glass, J. M., Zurbriggen, E. L., Gmeindl, L., Kieras, D. E., et al. (1999). Current response-selection processes in dual-task performance: Evidence for adaptive executive control of task scheduling. *Journal of Experimental Psychology. Human Perception and Performance*, *25*, 791–814.

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

Schvaneveldt, R. (1969). Effects of complexity in simultaneous reaction time tasks. *Journal of Experimental Psychology*, *81*, 289–296.

Shao, Z., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *Quarterly Journal of Experimental Psychology*, *65*, 1927–1944.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 896–918.

Szameitat, A. J., Lepsien, J., von Cramon, D. Y., Sterr, A., & Schubert, T. (2006). Task-order coordination in dual-task performance and the lateral prefrontal cortex: An event-related fMRI study. *Psychological Research*, *70*, 541–552.

Szameitat, A. J., Schubert, T., Müller, K., & von Cramon, D. Y. (2002). Localization of executive functions in dual-task performance with fMRI. *Journal of Cognitive Neuroscience*, *14*, 1184–1199.

Tombu, M., & Jolicœur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology. Human Perception and Performance*, *29*, 3–18.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, *38*, 584–589.

Van Maanen, L., Van Rijn, H., & Taatgen, N. (2012). RACE/A: An architectural account of the interactions between learning, task control, and retrieval dynamics. *Cognitive Science*, *36*, 62–101.

Van Selst, M., Ruthruff, E., & Johnston, J. C. (1999). Can practice eliminate the psychological refractory period effect? *Journal of Experimental Psychology. Human Perception and Performance*, *25*, 1268–1283.