

# Disentangling cognitive from motor control: Influence of response modality on updating, inhibiting, and shifting

Marpessa Rietbergen<sup>a,b,c,\*</sup>, Ardi Roelofs<sup>a</sup>, Hanneke den Ouden<sup>a</sup>, Roshan Cools<sup>b</sup>

<sup>a</sup> Radboud University, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition, Montessorilaan 3, 6525 HR, Nijmegen, the Netherlands

<sup>b</sup> Radboud University, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognitive Neuroimaging, Kapittelweg 9, 6525 EN, Nijmegen, the Netherlands

<sup>c</sup> International Max Planck Research School for Language Sciences, Wundtlaan 1, 6525 XD, Nijmegen, the Netherlands

## ARTICLE INFO

### Keywords:

Cognitive control  
Motor control  
Response modality

## ABSTRACT

It is unclear whether cognitive and motor control are parallel and interactive or serial and independent processes. According to one view, cognitive control refers to a set of modality-nonspecific processes that act on supramodal representations and precede response modality-specific motor processes. An alternative view is that cognitive control represents a set of modality-specific operations that act directly on motor-related representations, implying dependence of cognitive control on motor control. Here, we examined the influence of response modality (vocal vs. manual) on three well-established subcomponent processes of cognitive control: shifting, inhibiting, and updating. We observed effects of all subcomponent processes in reaction times. The magnitude of these effects did not differ between response modalities for shifting and inhibiting, in line with a serial, supramodal view. However, the magnitude of the updating effect differed between modalities, in line with an interactive, modality-specific view. These results suggest that updating represents a modality-specific operation that depends on motor control, whereas shifting and inhibiting represent supramodal operations that act independently of motor control.

## 1. Introduction

An important question regarding human cognition concerns the extent to which higher cognitive functions are related to lower motor processes (e.g., Haggard, Rossetti, & Kawato, 2007). According to one view, cognitive functions concern supramodal representations and processes, which occur before modality-specific motor processes are engaged. Alternatively, cognitive processes operate directly on modality-specific motor-related representations, which implies a close relationship between cognitive functions and motor processes. The question of whether cognitive processes and representations are supramodal or modality-specific (and cognition is “embodied” or not, e.g., Pecher & Zwaan, 2005) has been investigated in several domains, including memory, language, and thinking. It has also been examined in the domain of cognitive control (Brass & von Cramon, 2007), which is the topic of the present article.

Cognitive control is an umbrella term covering several different abilities, and concerns the higher-level processes that regulate lower-level processes needed to remain goal-directed, especially in the face of distraction (e.g., Miller & Cohen, 2001). According to an influential

proposal by Miyake et al. (2000), cognitive control comprises the three separable functions of updating, inhibiting, and shifting (see also Friedman et al., 2008). Updating is the ability to maintain and actively manipulate the contents of working memory, inhibiting is the ability to lower the interference of unwanted stimuli or responses, and shifting is the ability to switch back and forth between tasks or mental sets. In contrast to cognitive control, motor control entails the planning and selection of motor responses. It should be noted that the updating, inhibiting, and shifting components of cognitive control are separable but also correlated, as evident from a large-scale latent variable analysis by Miyake et al. (2000). Thus, experimental tasks and manipulations are unlikely to index the three cognitive control abilities in a pure fashion. Importantly, however, the three components nevertheless predominantly contributed to experimental tasks that were considered to mainly tap each component ability. Likewise, Sikora, Roelofs, Hermans, and Knoors (2016) observed that certain experimental manipulations (i.e., length, distractor, and sequence, to be discussed later) uniquely correlated with the updating, inhibiting, and shifting abilities, as measured by standard tests. Thus, although the updating, inhibiting, and shifting components are unlikely to contribute in a pure fashion to

\* Corresponding author at: Radboud University, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition, Montessorilaan 3, 6525 HR, Nijmegen, the Netherlands.

E-mail address: [M.Rietbergen@donders.ru.nl](mailto:M.Rietbergen@donders.ru.nl) (M. Rietbergen).

<https://doi.org/10.1016/j.actpsy.2018.09.008>

Received 18 January 2018; Received in revised form 15 June 2018; Accepted 15 September 2018

0001-6918/© 2018 Elsevier B.V. All rights reserved.

experimental tasks and manipulations, their contributions can be separated.

We first discuss previous studies that investigated the influence of response modality on shifting and inhibiting abilities separately. We do not know of any studies to date that investigated the influence of response modality on updating abilities. We then describe the limitations that arise from the paradigms that were used in those studies. Finally, we introduce our paradigm designed to avoid those limitations, and to better disentangle cognitive from motor control.

### 1.1. The influence of response modality on shifting

In a neuroimaging study on the modality-specific nature of cognitive control, Brass and von Cramon (2007) investigated whether switching reaction times (RTs) and brain activity were sensitive to manual and foot response modalities. Participants were instructed to attend to a series of symbolic cues and make either a parity (odd or even) or magnitude (smaller or larger) judgement concerning the numbers that would follow each cue. Depending on the direction of the judgement, they were instructed to press either a left or right button with the corresponding hand or foot. They found a switch effect, where RTs were shorter for cues that repeated task instructions, and longest for cues that switched between task instructions. A modality effect was also observed, in which RTs were shorter on manual than on foot trials. However, there was no interaction between switch and modality, meaning that the magnitude of the switch effect in the RTs did not differ between response modalities, suggesting that “cognitive-control-related processes... do not depend on response modality” (p. 112). In line with these behavioral results, functional magnetic resonance imaging (fMRI) during these tasks did not show an interaction between switch and modality within motor or control-related brain regions. An RT study by Philipp and Koch (2011) also assessed task switching with manual and foot responses, in addition to vocal responses. This study replicated the findings from Brass and von Cramon, such that the switch cost did not differ as a function of response modality. Taken together, these results are in line with a supramodal view.

The two aforementioned studies assessed the influence of response modality using a cued task-switching paradigm, which made use of an arbitrary stimulus-response (S-R) mapping. Brass and von Cramon (2007) describe this as a mapping where the behavioral response is not automatically evoked by the stimulus but rather guided by abstract rules. The abstract rules that guide behavior are less automatic, and this suggests that additional cognitive control is needed. Conversely, a non-arbitrary S-R mapping is one where the stimulus more automatically elicits a behavioral response. For instance, a visuospatial stimulus (e.g., a left or right pointing arrow, or a target in a particular spatial position) would more directly elicit a directional or spatial behavioral response (e.g., a left- or right-related button press). It is possible that the lack of an interaction between cognitive and motor processes in the studies of Brass and von Cramon and Philipp and Koch (2011) might reflect the use of arbitrary S-R mapping, because arbitrary mappings call upon additional control strategies (such as rule retrieval). Non-arbitrary S-R mappings have been used in studies using particular Stroop-like tasks, which have been associated with the inhibiting component of cognitive control (e.g., Miyake et al., 2000). We discuss these studies next.

### 1.2. The influence of response modality on inhibiting

Baldo, Shimamura, and Prinzmetal (1998) examined the influence of response modality using arrow-word, arrow-arrow, and word-word versions of a Stroop-like task. For instance, in the arrow-word version, congruent and incongruent combinations of the words *left* or *right* and left- or right-pointing arrows (e.g., the word *right* combined with an incongruent left-pointing arrow) were presented. Participants responded to the word while ignoring the arrow (the word task) or to the arrow while ignoring the word (the arrow task). Responses were made

using either the right index finger (respond left) or right middle finger (respond right), or by saying “left” or “right”. The Stroop-like effect is the difference in RT between incongruent and congruent trials. Despite vocal RTs being generally longer than manual RTs, they found that the magnitude of the Stroop-like effect was larger for the word task when manual responses were required compared to when vocal responses were required. Conversely, for the arrow task, the Stroop-like effect was larger for vocal than manual responses. In a different study, Turken and Swick (1999) compared behavior between age-matched healthy controls and a patient with a lesion to part of the motor area of the right hemispheric anterior cingulate cortex (ACC) involved in manual control. The ACC is commonly associated with cognitive control (e.g., Barch et al., 2001; Miller & Cohen, 2001). Their results replicated those of Baldo et al. in that the Stroop-like effect for the patient and controls was larger for manual than for vocal responses in the word task, and larger for vocal than for manual responses in the arrow task. The patient showed a disproportionately larger Stroop-like effect in manual but not in vocal responding. These findings reported by Baldo et al., and further supported by Turken and Swick, suggest that inhibiting abilities are not independent of response modality, in line with the modality-specific view.

However, in the study from Baldo et al. (1998), the effect of response modality was further investigated in word-word and arrow-arrow versions of the Stroop-like task, which showed no influence of response modality. Thus, when participants were presented with congruent and incongruent combinations of two words (i.e., *left* or *right*), or combinations of two arrows (i.e., left- or right-pointing arrows), different response modalities did not influence the magnitude of the Stroop-like effect. This suggests that the modality effect was dependent on the stimulus rather than on the response dimensions. When the target and distractor dimensions of the stimuli are the same (i.e., two words or two arrows), the magnitude of the Stroop-like effect for the arrow and word tasks is not influenced by response modality.

To conclude, the literature provides support in favor and against both the supramodal and modality-specific views. The studies described in detail above used paradigms that evaluated the influence of response modality on shifting and inhibiting abilities separately, making use of arbitrary or non-arbitrary S-R mappings with discrete binary responses, in particular, hand, foot, or vocal responses. The current study uses non-arbitrary S-R mappings with discrete and non-discrete responses, and with stimulus dimensions that were exactly matched between response modalities. Moreover, we assessed effects of response modality not only for shifting and inhibition, but also for updating, in a within-subjects design. The evidence in favor or against the supramodal and modality-specific views was quantified using Bayesian statistical analyses.

### 1.3. Assessing updating, inhibiting, and shifting in a single paradigm

To examine the influence of response modality on updating, inhibiting, and shifting, we adopted a flanker-like paradigm (Eriksen & Eriksen, 1974) that was modeled after a spoken noun-phrase picture-naming task used in a study by Sikora, Roelofs, Hermans, and Knoors (2016). In this study, picture-description RTs were evaluated between more-demanding and less-demanding trials, by measuring effects of length (i.e., the difference between trials with long phrases and short phrases, assessing updating), distractor (i.e., the difference between trials with incongruent and congruent distractors, assessing inhibiting), and sequence (i.e., the difference between switch and repeat trials, assessing shifting). For example, participants said “the fork” to a black-and-white picture of a fork (the short phrase condition) and “the green fork” to a fork in green color (the long phrase condition). They heard spoken distractor words that were either congruent (“fork”) or incongruent (“spoon”) with the picture name. The required phrase type changed every second trial, so that picture description RTs were obtained on repeat trials (a short trial preceded by a short trial, or a long

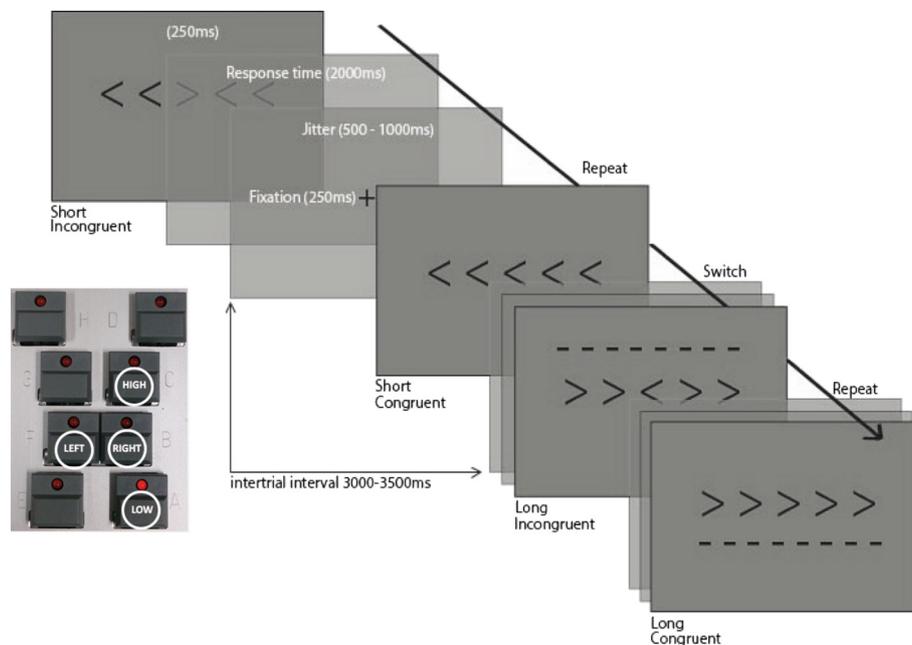


Fig. 1. Illustration of the task design and protocol. A trial consisted of a jitter (blank screen 500–1000 ms), a black fixation cross (250 ms), the stimulus (250 ms), and finally a blank response screen (2000 ms).

trial preceded by a long trial) and on switch trials (short-long and long-short). The length, distractor, and sequence effects correlated with performance on well-established cognitive tasks assessing updating, inhibiting, and shifting respectively (a complex-span task, stop-signal task, and a shape-color switching task; cf. Friedman et al., 2008; Miyake et al., 2000). By adopting this design in the present study, we assessed effects of length, distractor, and sequence in RTs when producing comparable vocal and manual responses to evaluate the influence of response modality on updating, inhibiting, and shifting abilities.

We expected that RTs on more control-demanding trials would be longer than on less demanding trials (i.e., effects of either length, distractor, or sequence), irrespective of response modality. However, the critical question was whether the magnitude of these control effects differs between vocal and manual responses, supporting the modality-specific view, or whether they do not differ, supporting the supramodal view. Under the supramodal view, we expect to find no interaction between any of the three cognitive control effects and response modality ( $H_0$ ). In contrast, under the modality-specific view, we do expect to find interactions between the modality and any of the three cognitive effects ( $H_1$ ). We adopted a Bayesian hypothesis testing approach to quantify the evidence against or in favor of each theoretical view.

## 2. Method

### 2.1. Participants

Forty native speakers of Dutch participated in the experiment (35 women and 5 men, mean age is 22.40 years, age range 18 to 30 years). The participants were recruited via the Radboud University SONA system. They received 5 Euros or study credit for their participation.

### 2.2. Materials

To examine the influence of response modality on updating, inhibiting, and shifting, we assessed length, distractor, and sequence effects when producing comparable vocal and manual responses. We used an arrow version of the Eriksen flanker task (Eriksen & Eriksen, 1974) with incongruent and congruent combinations of a left- or right-pointing arrow target (e.g., < or >) and two distractor arrows flanked

on each side (e.g., < < or > >). Thus, the target and distractor dimensions were matched (both arrows), so that an effect of response modality rather than mismatch in stimulus dimensions was assessed. The arrow stimuli were presented as simple stimuli (e.g. > > < > >) or as compound stimuli, in which they were combined with a dashed line positioned above or below the arrows (e.g., \_> > < > >). In response to the simple stimuli, participants indicated the direction of the target arrow by saying “left” or “right” (vocal response) or by pressing a left or right button (manual response). In response to the compound stimuli, participants indicated both the position of the line and the direction of the target arrow by saying, for example, “low, left” (vocal response) or by sequentially pressing two out of four spatially arranged buttons (always referring to first the position of the line and then the target arrow).

### 2.3. Procedure and design

Our task design ensured that a non-arbitrary mapping between the stimuli and the responses existed across both modalities. To measure each of the behavioral effects (i.e., length, distractor, and sequence), trials were juxtaposed between more and less cognitively demanding to assess control being exerted. For long trials, participants would process compound stimuli composed of a line and an arrow component, whereas short trials depicted simply the arrow component. Longer RTs and lower accuracy present on long than short trials would reflect the additional updating needed for the additional response element (i.e., the second word or button). Longer RTs and lower accuracy on incongruent than congruent trials would be indicative of interference posed by flankers and would demonstrate the degree of inhibiting needed to overcome interference. For switch trials, sequential trials would shift between long and short trials, or vice versa, which would result in longer RTs and lower accuracy due to a shift from one task set to another. Thus, responses were short or long (length: short vs. long, e.g., one word or button vs. two words or buttons), they were the same or different length between consecutive trials (sequence: repeat vs. switch, e.g., long-long vs. long-short), and they were produced in response to congruent or incongruent arrow stimuli (distractor), see Fig. 1.

All participants in the study gave written informed consent before

beginning the experimental session. After the experiment, participants were debriefed and paid for their participation. The duration of an experimental session was 35 min, and was completed in a soundproof recording booth, with participants seated facing a computer screen and positioned in front of a microphone connected to a voice key. Participants were given a Buttonbox to hold in their left hand for stability, and were instructed to use only their right thumb when pressing buttons accordingly. The four buttons on the Buttonbox were spatially arranged to be mappings primarily between left and right buttons and left and right target arrow stimuli and secondarily between top and bottom buttons and the high or low position of the line stimuli. Participants were further instructed to look at the screen and to respond as quickly and accurately as possible by speaking into the microphone or by pressing buttons with their right thumb. Trials began with a black fixation cross centered on a grey screen for a duration of 250 ms, see Fig. 1. Following the fixation, the stimuli were presented for 250 ms, followed by a 2000-ms period during which participants responded accordingly (i.e., a blank grey screen). Following the offset of the response period of 2000 ms, trials were jittered so that the interstimulus interval varied between 500 and 1000 ms, making the average intertrial interval between 3000 and 3500 ms. Trials were jittered so that we might utilize this similar design for future fMRI research.

Before beginning the experimental session, participants first completed four practice blocks of eight trials each to ensure compliance with the instructions: two practice blocks to practice vocal responses, and two blocks to practice manual (button) responses. There were a total of 384 experimental trials split into 32 blocks, with 12 experimental trials per block. At the start of every block, a black line-drawing of a mouth or hand would indicate whether to respond by speaking into the microphone or by pressing buttons for the upcoming block of trials. The blocks were split equally between each response modality (16 blocks), and blocks were randomized. The experiment was performed and reaction time data for both manual and vocal responses was retrieved using Presentation® software (Version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA). All materials were counterbalanced and pseudo-randomized in a custom script using MATLAB (2015, Mathworks Inc., Sherborn, MA, USA). Each block of trials was randomized with the constraint that the different sequences of lengths occurred equally often: long-long (repeat trials), long-short (i.e., switch trials), short-short (i.e., repeat trials), short-long (i.e., switch trials). Additionally, the randomization of trials within each block was further constrained to ensure that distractor information, incongruent and congruent flanking arrows, also occurred equally often. Each trial type of vocal, manual, long, short, incongruent, congruent, switch, and repeat occurred on a 192 experimental trials, six times within a block.

#### 2.4. Analysis

All 40 participants who partook in the study were included in the analysis. We analyzed both RTs and accuracy (percentage errors, PE). Raw RTs for vocal responses and raw RTs and PEs for manual responses were retrieved directly from Presentation® output. The speech onsets of all vocal trials were recorded using a voice key, and were marked for errors by the experimenter. Only responses that were completely accurate for both first and second response elements were included in the RT analysis, and those responses given inaccurately were used for the PE analysis. Trials that were excluded from the analyses included false starts occurring before 200 ms, responses that were not complete before the end of the trial (2000 ms), trials that followed an error, and self-corrections and disfluencies on vocal trials (approximately 7% of the total trials were excluded from the analyses). The mean RTs and PEs were calculated separately for both the manual and vocal response types and for each of the six condition types: long, short, incongruent, congruent, switch, and repeat. We ran Bayesian repeated-measures ANOVAs, with the following crossed within-subject factors: *modality* (vocal vs. manual trials), *length* (long vs. short trials), *distractor*

(incongruent vs. congruent trials), and *sequence* (switch vs. repeat trials). The Bayesian analyses were performed using JASP (Love et al., 2017).

In a Bayesian hypothesis testing approach, the Bayes factor ( $BF$ ) represents the ratio of evidence against or in favor of two competing hypotheses (Rouder, Morey, Verhagen, Swagman, & Wagenmakers, 2017; Wagenmakers et al., 2017). Thus, unlike using a  $p$ -value to quantify evidence only against the null hypothesis ( $H_0$ ), the Bayesian approach quantifies the evidence both against and in favor of the null hypothesis and the alternative hypothesis ( $H_1$ ). The subscript “10” in  $BF_{10}$  indicates that the Bayes factor quantifies the evidence that the data provide for  $H_1$  versus  $H_0$ , whereas the subscript “01” indicates the reverse (i.e.,  $BF_{10} = 1/BF_{01}$ ). Under a standard interpretation, a  $BF_{10}$  of 10–30 indicates “strong”, 30–100 “very strong”, and  $> 100$  “decisive” evidence for  $H_1$ . Conversely, a  $BF_{01}$  of 1/10–1/30 indicates “strong”, 1/30–1/100 “very strong”, and  $< 1/100$  “decisive” evidence for  $H_0$ .

We used Bayesian repeated-measures ANOVAs to test for the presence of main effects of modality, length, distractor, and sequence, and of interactions between modality and effects of length, distractor, and sequence for RT and PE. The analyses yield model comparisons, with quantified Bayes factors for all possible models in favor or against the null. Crucially, using this approach we tested whether the best models include or lack interaction terms between modality and the factors of length, distractor, or sequence. Evidence for the absence of an interaction term would provide support for the supramodal view on cognitive control. However, if the winning model does include such interactions, then this would be in favor of the modality-specific view.

### 3. Results

Table 1 presents the mean RT and PE for the trial types across response modalities. The table shows that RT was longer for vocal than manual responses, suggesting a general effect of response modality. Moreover, RTs varied with trial type in that they were longer for long-response than for short-response trials (length effect), longer for incongruent than for congruent trials (distractor effect), and longer for switch than repeat trials (sequence effect). Moreover, the length effect was larger for the manual than for the vocal responses (256 vs. 177 ms), whereas the distractor and sequence effects were more similar between the response modalities (respectively 44 vs. 57 ms and 29 vs. 43 ms), see Fig. 2. The PEs were generally in line with the RTs.

The statistical analyses of the RTs indicated decisive evidence for an effect of all factors: modality ( $BF_{10} = 1.9e+4$ ), length ( $BF_{10} = 6.41e+14$ ), distractor ( $BF_{10} = 1.5e+12$ ), and sequence ( $BF_{10} = 2.5e+9$ ). These findings indicate that RTs on cognitively more demanding trials were indeed slower than on less demanding trials, for updating, inhibiting, and shifting. The model comparison indicated that the RT data are more likely to occur under the alternative main effects model that included all four factors ( $BF_{10} = 3.4e+142$ ) with respect to the null than other models with fewer factors.

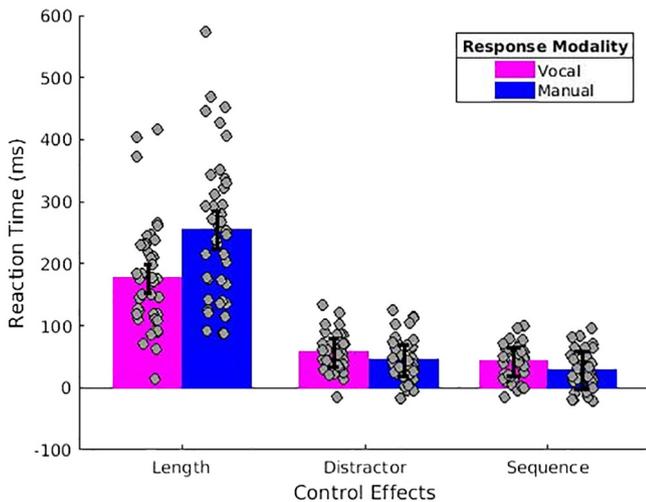
We used model comparison to further quantify the evidence in favor or against including interactions to explain the RT data. We first evaluated the influence of each interaction term of interest when added to the main effects model. These interactions of interest concerned the first-order interactions between modality and length, modality and distractor, and modality and sequence. The main effects model plus the modality  $\times$  length interaction yielded a  $BF_{10} = 3.3e+6$  relative to the main effects model without this interaction, indicating decisive evidence for an interaction between modality and length in the data. However, adding the modality  $\times$  distractor interaction to the main effects model reduced the  $BF_{10}$  by a factor of 0.19, and adding the modality  $\times$  sequence interaction reduced the  $BF_{10}$  by a factor of 0.20. Thus, there was no evidence for the presence of interactions between modality and distractor and between modality and sequence in the data.

Adding the length  $\times$  sequence interaction to the main effects model

**Table 1**

Mean reaction times (RT) and standard errors (SE) in milliseconds and percentage errors (PE) in the length, sequence, and distractor conditions for the vocal and manual response modalities.

Length	Sequence	Distractor	Modality								Total			
			Vocal				Manual							
			RT	SE	PE	SE	RT	SE	PE	SE	RT	SE	PE	SE
Long	Switch	Incongruent	897	20	6.5	1.0	828	25	5.2	1.4	863	23	5.9	1.2
		Congruent	856	19	6.4	1.3	826	27	5.2	0.8	841	24	5.8	1.1
		Total	877	20	6.5	1.2	827	26	5.2	1.2	852	23	5.8	1.2
	Repeat	Incongruent	879	21	6.2	0.9	844	30	3.3	0.9	861	26	4.8	0.9
		Congruent	832	21	4.0	0.9	813	28	5.2	1.0	822	25	4.6	1.0
		Total	855	21	5.1	0.9	828	29	4.2	1.0	842	25	4.7	0.9
	Total	Incongruent	888	21	6.4	0.9	836	28	4.3	1.2	862	24	5.3	1.1
		Congruent	844	20	5.2	1.1	819	28	5.2	0.9	832	24	5.2	1.0
		Total	866	20	5.8	1.0	828	28	4.7	1.1	847	24	5.3	1.0
Short	Switch	Incongruent	747	13	1.9	0.7	632	18	3.6	1.2	690	16	2.8	1.0
		Congruent	694	15	2.6	0.7	574	18	1.1	0.4	634	16	1.8	0.6
		Total	721	14	2.3	0.7	603	18	2.3	0.9	662	16	2.3	0.8
	Repeat	Incongruent	703	12	2.4	0.8	589	14	2.5	0.7	646	13	2.4	0.7
		Congruent	607	11	2.2	0.7	493	12	1.5	0.7	550	11	1.9	0.7
		Total	655	11	2.3	0.7	541	13	2.0	0.7	598	12	2.2	0.7
	Total	Incongruent	725	12	2.2	0.7	611	16	3.0	1.0	668	14	2.6	0.9
		Congruent	651	13	2.4	0.7	534	15	1.3	0.6	592	14	1.9	0.6
		Total	688	13	2.3	0.7	572	16	2.2	0.8	630	14	2.2	0.8
Total	Switch	Incongruent	822	17	4.2	0.8	730	22	4.4	1.3	776	20	4.3	1.1
		Congruent	775	17	4.5	1.1	700	23	3.1	0.7	738	20	3.8	0.9
		Total	799	17	4.4	1.0	715	22	3.8	1.0	757	20	4.1	1.0
	Repeat	Incongruent	791	17	4.3	0.8	716	23	2.9	0.8	754	20	3.6	0.8
		Congruent	719	17	3.1	0.8	653	22	3.4	0.9	686	19	3.2	0.8
		Total	755	17	3.7	0.8	685	23	3.1	0.8	720	20	3.4	0.8
	Total	Incongruent	806	17	4.3	0.8	723	23	3.7	1.1	765	20	4.0	0.9
		Congruent	747	17	3.8	0.9	677	22	3.2	0.8	712	20	3.5	0.9
		Total	777	17	4.0	0.9	700	22	3.5	0.9	738	20	3.7	0.9



**Fig. 2.** Magnitude of each control effect (length, distractor, sequence) as a function of response modality (vocal, manual), together with error bars indicating standard error and mean differences for each participant.

with the modality  $\times$  length interaction increased the  $BF_{10}$  by a factor of 481. In particular, the sequence effect was larger on short trials than on long trials (70 ms vs. 10 ms). Moreover, adding the length  $\times$  distractor interaction to this extended model further increased the  $BF_{10}$  by a factor of  $2.1 \times 10^4$ , as the distractor effect was also larger on short trials than on long trial (76 ms vs. 30 ms). Adding the distractor  $\times$  sequence interaction also further increased the  $BF_{10}$  by a factor of 54, as the distractor effect was larger on repeat trials than on switch trials (68 ms vs. 38 ms). No other 3- or 4-way interactions proved to explain the RT data. Therefore, the model that includes all main effects, 2-way interactions between length, sequence, distractor, and the interaction between

modality and length, best explains the RT data as compared to all lower and higher models.

Concerning PE, the model comparison revealed that the model that best explained accuracy simply included the factor of length,  $BF_{10} = 1.2 \times 10^{18}$ , with respect to the null. The inclusion of the other factors or interactions (of interest) did not provide any additional evidence in favor of the alternative model. There was no evidence for an influence of modality ( $BF_{10} = 0.38$ ), distractor ( $BF_{10} = 0.61$ ), or sequence ( $BF_{10} = 0.44$ ).

In summary, whereas PE showed an effect of length only, the model that best explains the RT data includes all main effects, an interaction between modality and length, and interactions between length and sequence, length and distractor, and distractor and sequence. Relevant for the adjudication between the supramodal and modality-specific views is that modality interacted with length, but not with distractor and sequence.

#### 4. Discussion

The aim of this study was to test between the supramodal and modality-specific views on the relation between cognitive and motor control. We examined the influence of response modality on updating, inhibiting, and shifting abilities using non-arbitrary S-R mappings and matching stimulus dimensions within a single paradigm. We observed all main effects, as indicated by RTs that were longer in more control-demanding conditions (i.e., long, incongruent, switch) than in less control-demanding conditions (i.e., short, congruent, repeat). Our RT results revealed an interaction between length and modality, an interaction that was absent for distractor and sequence effects. These results indicate that response modality influences cognitive control processes related to updating, but not to shifting and inhibiting. These findings have important implications for the debate regarding the relationship between cognitive and motor control processes, as we discuss below.

Response modality influenced the ability to update task representations in working memory. This suggests that updating is modality-specific, and “grounded” in motor control systems. In addition, we replicated previous findings revealing that response modality does not influence shifting (Brass & von Cramon, 2007; Philipp and Koch, 2011) or inhibiting (Baldo et al., 1998). This suggests that shifting and inhibiting represent supramodal processes that precede the selection of appropriate motor-related representations (Badre, 2008; Botvinick, 2008; Koechlin, Ody, & Kouneiher, 2003).

The difference in sensitivity to modality effects of updating on the one hand and shifting and inhibiting on the other, aligns well with their temporal manifestations as reported in another study by Sikora, Roelofs, and Hermans (2016). They measured event-related brain potentials (ERPs) during updating, shifting, and inhibiting in a noun-phrase production task and found that shifting was accompanied by an increase in both anterior and posterior N200. Conversely, updating was accompanied instead by decreases in a later ERP, the P300. Accordingly, the temporal manifestations of processes that relate to shifting (and inhibiting) occur earlier than those processes involved in updating, which is closer to the actual planning and production of the spoken phrases.

Some might argue that the interaction between modality and length effects reflects differences in S-R compatibility between the vocal and manual conditions (Eimer, 1995; Kunde, 2001). Thus, the visuospatial nature of the stimuli might be considered to be more compatible with the manual than the vocal modality of the responses. However, we believe that this account is unlikely, given the selectivity of the modality effect to updating, not extending to shifting and inhibiting. If S-R compatibility would have played a role, then effects should generalize to all three control conditions.

In addition to the interaction between modality and length, we found an interaction between distractor and length. On short trials, the response indicates the direction of the target arrow (e.g., “left” or left button press) and competition by the flankers must be resolved before the response can be given. However, on long trials, the second response element indicates the direction of the target arrow, and competition by the flankers may be resolved, at least partly, while planning (and producing) the first response element indicating the position of the line (e.g., low or high). This may explain why the distractor effect was larger on short trials than on long trials. The sequence effect was also larger on short trials than on long trials. The asymmetry suggests that disengagement takes longer for one task set than another due to differential enhancement or inhibition of the prior (Allport & Wylie, 1999; see also Sikora, Roelofs, Hermans, & Knoors, 2016). Specifically, the compound stimuli on long trials (e.g.,  $\_ > \_ > \_ < \_ > \_ > \_$ ) require a long response but also afford a short response (e.g., say “left” or a left button press), which needs to be inhibited. The simple stimuli on short trials (e.g.,  $> > < > >$ ) only afford a short response. When a switch from a long trial to a short trial occurs, the previous inhibition of the task set for short responses needs to be overcome, which delays responding. In contrast, when a switch from a short trial to a long trial occurs, there is no such previous inhibition that needs to be overcome, and responding is not delayed. This explains the asymmetrical switch costs obtained. Finally, the distractor effect was larger on repeat trials than on switch trials. As in the case of the distractor effect on short versus long trials, the additional time to shift task set may be exploited to resolve interference resulting in a difference in magnitude of the distractor effect for switch as compared to repeat trials.

It is important to note that there is an asymmetry in what one can infer from additive vs. interactive effects. An interaction suggests a close relation between cognitive control and response modality, but additivity is actually more neutral about the issue. We found additive effects of modality and sequence and of modality and distractor, suggesting that shifting and inhibiting are independent of motor control. However, it is also possible that shifting and inhibiting still interact with motor control, but in the same way for both modalities, making no

difference between them. This possibility may be assessed by running an fMRI study, as Brass and von Cramon (2007) did, and to test for an interaction between sequence and modality as well as distractor and modality, within motor and control-related brain regions (cf. Philipp, Weidner, Koch, & Fink, 2013). However, although we observed an interaction between length and modality, and interactions of length and sequence, and of length and distractor, there were no triple interactions of the latter with modality. We argued that the interaction of length and modality suggests that updating is modality-specific. If inhibiting and shifting also interact with motor control, but in the same way for both modalities (making no difference between them), then one might expect triple interactions of length, distractor, and modality, and of length, sequence, and modality. However, such triple interactions were not observed. This seems to slightly favor the view that inhibiting and shifting are modality-nonspecific rather than that they interact with motor control but in the same way for both modalities.

Furthermore, in our experiment, the short and long responses differed in the first element, namely ‘left or right’ vs. ‘high or low, left or right’. This raises the question of whether the main effect of length is due to the planning of one vs. two response elements (which we maintain) or due to a difference in motor requirements of ‘left or right’ vs. ‘high or low’ responses. The interaction between length and response modality would then suggest that the motor requirements are also different between modalities. The RTs are about the same for the long vocal and long manual responses, which would suggest that the motor requirements are about equal for ‘high or low’ vocal and manual responses. In contrast, the RTs are much shorter for the short manual responses than the short vocal responses, which would suggest that the motor requirements differ between ‘left or right’ manual and vocal responses. However, under this alternative motor requirements account, it remains unclear why the motor demands differ between ‘left or right’ vs. ‘high or low’ responses, and why this depends on modality for ‘left or right’ but not for ‘high or low’. In contrast, independent evidence suggests that it takes longer to plan a two-element than a one-element response (e.g., Rosenbaum, 2010). Moreover, independent evidence suggests that RTs for ‘left or right’ responses to arrows are shorter in the manual than in the vocal modality (Baldo et al., 1998). Thus, existing evidence favors an updating account over a motor requirements account. Still, a follow-up study where the first element of long and short responses would remain unchanged could definitely settle the issue.

One may also argue that the sequence effect is not due to shifting but is the result of repetition priming. However, although repetition priming may account for the main sequence effect, it has difficulty explaining the observed interaction with length. If repetition of a response type (e.g., short or long) causes the sequence effect, this should hold for both short and long responses. Of course, it is possible that short responses may benefit less from repetition priming than long responses, because short responses are already made quickly (i.e., there is less room for improvement). However, this was clearly not the case. Although we observed an interaction between sequence and length, the sequence effect was observed for the short responses (64 ms) but not for the long ones (7 ms). This is exactly the opposite of what one would expect under a repetition priming account. In contrast, the asymmetrical switch cost that we observed is readily explained by assuming differential difficulty in disengaging from the previous task set (i.e., differential task-set inertia).

Lastly, we used flanker stimuli to engage the inhibiting component of cognitive control, whereas Sikora, Roelofs, and Hermans (2016) and Sikora, Roelofs, Hermans, and Knoors (2016) used auditory-visual Stroop-like stimuli. One may further argue that flanker and Stroop-like stimuli tap into different types of inhibition, namely interference control and prepotent response inhibition, respectively. However, in a large-scale latent variable analysis, Friedman and Miyake (2004) observed that although these types of inhibition may be separated, the best model had them as a single latent variable. Still, future research may examine whether Stroop-like distraction and response modality

yield additive effects, as we observed for flanker distraction, or whether an interaction is observed.

## 5. Conclusion

The aim of this study was to adjudicate between the supramodal and modality-specific views on the relation between cognitive and motor control. To this end, we examined the influence of response modality on three core subcomponents of cognitive control, namely updating, inhibiting, and shifting. We replicated findings from previous studies, showing that shifting and inhibiting are not influenced by response modality, while, in contrast, updating does vary as a function of

response modality. These results suggest that inhibiting and shifting are supramodal cognitive control functions that precede motor control, whereas updating is a cognitive control function that exhibits greater proximity to motor control functions.

## Acknowledgments

The authors wish to thank two reviewers for helpful comments. This study was supported by the Gravitation Grant 024.001.006 of the Language in Interaction Consortium from the Netherlands Organization for Scientific Research (NWO).

## Appendix A

Table A1  
Results of the repeated-measures ANOVA.

Main effects and interactions	F(1,39)	$\eta_p^2$
Mode	33.77	0.464***
Length	220.90	0.850***
Distractor	157.36	0.801***
Sequence	94.25	0.708***
Length × Distractor	42.65	0.522***
Length × Sequence	35.73	0.478***
Sequence × Distractor	11.64	0.229**
Modality × Length	22.52	0.366***
Modality × Distractor	4.24	0.098
<sup>a</sup> Modality × Sequence	5.34	0.120*
<sup>a</sup> Modality × Length × Distractor	7.04	0.152*

<sup>a</sup> Not surviving Bayesian hypothesis testing.

\*  $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .

## References

- Allport, A., & Wylie, G. (1999). Task-switching: Positive and negative priming of task-set. In G. W. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, space, and action: Studies in cognitive neuroscience* (pp. 273–296). Oxford: Oxford University Press.
- Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, *12*, 193–200.
- Baldo, J. V., Shimamura, A. P., & Prinzmetal, W. (1998). Mapping symbols to response modalities: Interference effects on Stroop-like tasks. *Perception & Psychophysics*, *60*, 427–437.
- Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex*, *11*, 837–848.
- Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, *12*, 201–208.
- Brass, M., & von Cramon, Y. (2007). How motor-related is cognitive control? In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition: Attention and performance XXII* (pp. 101–119). Oxford, UK: Oxford University Press.
- Eimer, M. (1995). S-R compatibility and automatic response activation: Evidence from psychophysiological studies. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 837–854.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143–149.
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*, 101–135.
- Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*, 201–225.
- Haggard, P., Rossetti, Y., & Kawato, M. (Eds.). (2007). *Sensorimotor foundations of higher cognition: Attention and performance XXII*. Oxford, UK: Oxford University Press.
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, *302*, 1181–1185.
- Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 387–394.
- Love, J., Selker, R., Verhagen, J., Marsman, M., Gronau, Q. F., Jamil, T., et al. (2017). *JASP (Version 0.8.1.2) [Computer software]*.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.
- 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.
- Pecher, D., & Zwaan, R. A. (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. Cambridge University Press.
- Philipp, A. M., & Koch, I. (2011). The role of response modalities in cognitive task representations. *Advances in Cognitive Psychology*, *7*, 31–38.
- Philipp, A. M., Weidner, R., Koch, I., & Fink, G. (2013). Differential roles of inferior frontal and inferior parietal cortex in task switching: Evidence from stimulus-categorization switching and response-modality switching. *Human Brain Mapping*, *34*, 1919–1920.
- Rosenbaum, D. A. (2010). *Human motor control* (2nd Edition). Elsevier: Amsterdam.
- Rouder, J. N., Morey, R. D., Verhagen, A. J., Swagman, A. R., & Wagenmakers, E.-J. (2017). Bayesian analysis of factorial designs. *Psychological Methods*, *22*, 304–321.
- Sikora, K., Roelofs, A., & Hermans, D. (2016). Electrophysiology of executive control in spoken noun-phrase production: Dynamics of updating, inhibiting, and shifting. *Neuropsychologia*, *84*, 44–53.
- Sikora, K., Roelofs, A., Hermans, D., & Knoors, H. (2016). Executive control in spoken noun-phrase production: Contributions of updating, inhibiting, and shifting. *Quarterly Journal of Experimental Psychology*, *69*, 1719–1740.
- Turken, U., & Swick, D. (1999). Response selection in the human anterior cingulate cortex. *Nature Neuroscience*, *2*, 920–924.
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, A. J., Love, J., et al. (2017). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*, 35–57.