



ELSEVIER

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych



Talking while looking: On the encapsulation of output system representations



Lynn Huestegge^{a,b,*}, Aleksandra Pieczykolan^{a,b}, Iring Koch^b

^a Institute of Psychology, University of Würzburg, Würzburg, Germany

^b Institute of Psychology, RWTH Aachen University, Aachen, Germany

ARTICLE INFO

Article history:

Accepted 10 June 2014

Keywords:

Modularity
Encapsulation
Crosstalk
Dual-task interference
Saccadic eye movements
Vocal responses

ABSTRACT

The idea that the human mind can be divided into distinct (but interacting) functional modules is an important presupposition in many theories of cognition. While previous research on modularity predominantly studied input domains (e.g., vision) or central processes, the present study focused on cognitive representations of output domains. Specifically, we asked to what extent output domain representations are encapsulated (i.e., immune to influence from other domains, representing a key feature of modularity) by studying determinants of interference between simultaneous action demands (oculomotor and vocal responses). To examine the degree of encapsulation, we compared single- vs. dual-response performance triggered by single stimuli. Experiment 1 addressed the role of stimulus modality under dimensionally overlapping response requirements (stimuli and responses were spatial and compatible throughout). In Experiment 2, we manipulated the presence of dimensional overlap across responses. Substantial performance costs associated with dual-response (vs. single-response) demands were observed across response modalities, conditions, and experiments. Dimensional overlap combined with shared spatial codes across responses enabled response-code priming (i.e., beneficial crosstalk between output domains). Overall, the results are at odds with the idea of strong encapsulation of output system representations and show how processing content determines the extent of interdependency between output domains in cognition.

© 2014 Elsevier Inc. All rights reserved.

* Corresponding author at: Institute of Psychology, University of Würzburg, Röntgenring 11, 97070 Würzburg, Germany. Fax: +49 931 3182616.

E-mail address: lynn.huestegge@uni-wuerzburg.de (L. Huestegge).

1. Introduction

Cognitive psychology has always aimed at specifying contents, processes, and components of the mind. An important part of this endeavor has been the development of theories on information processing and its structural bases. A central characteristic of a broad range of theories about the mind is the modularity assumption, that is, the notion that at least some of the mental phenomena can be explained in terms of the interplay of distinct cognitive modules. However, only little consensus has been reached on the question of what exactly qualifies as a module, or to what extent certain “parts” of the mind can be adequately described as being modular (Barrett & Kurzban, 2006).

1.1. Modularity in cognition

While the assumption of cognitive modules (or mental faculties) can easily be traced back to at least ancient Greek philosophy, the debate about modularity in cognition has been substantially revitalized by Fodor (1983). According to his conceptualization, cognitive modules represent functionally specialized cognitive systems that primarily occur at peripheral processing stages. In addition to functional specialization and peripheral localization, Fodor also proposed further properties of cognitive modules, for example, domain specificity (e.g., the visual system on the input side of processing, see Marr, 1982), and brain localization (i.e., that modules are realized in dedicated parts of the brain, e.g., Friston & Price, 2011). In a subsequent discussion, researchers suggested several candidates for a key feature of modularity, for example, domain specificity (Coltheart, 1999), or functional specialization (Barrett & Kurzban, 2006). However, many other researchers (including Fodor) agreed that probably the most important property of a module is informational encapsulation, which refers to the idea that a module is cognitively impenetrable (i.e., inaccessibility to the influence of other modules or higher cognitive processes) and does not refer to other psychological systems in order to operate (Fodor, 1983; Pylyshyn, 1999).

Fodor’s original work has been criticized on a number of fronts (e.g., Buller, 2005). For example, some researchers claimed that modularity may also occur at central processing stages (e.g., Magen & Cohen, 2010; Sternberg, 2011), a view that has also been termed “massive modularity” in the context of a functional (evolutionary) view of cognition (see Carruthers, 2005; Frankenhuys & Ploeger, 2007; Pinker, 2005; Sperber, 2005; Tooby & Cosmides, 1992). Furthermore, it has been argued that many empirical observations are incompatible with the assumption of (both peripheral and central) modularity (e.g., J. J. Prinz, 2006). For example, brain plasticity appears to speak against a fixed localization of modules in the brain, and the many instances of information crosstalk between processing domains (e.g., the McGurk effect, see MacDonald & McGurk, 1978, showing an interaction between vision and audition on the input side of processing) typically count as evidence against strong encapsulation (see also Jiang & Egner, 2013, for more recent evidence). However, it should also be noted that Fodor (1983, p. 37) himself conceded that the notion of modularity “ought to admit of degrees”, and should not be judged in terms of an all-or-nothing phenomenon. Thus, cognitive system interdependency could be considered as a research-guiding dimension with the (heuristically valuable) poles of strong modularity/encapsulation on one side and strong crosstalk on the other side. In this way, the crucial research questions no longer refer to the existence (or non-existence) of modularity, but rather to the factors and mechanisms determining the degree of modularity in specific domains. In the present study, we specifically focus on the issue of encapsulation with respect to cognitive representations of output systems.

1.2. Encapsulation and output system modularity

Interestingly, most empirical research on cognitive encapsulation mainly focused on the input side of processing (e.g., vision, see Barrett & Kurzban, 2006, for a review), whereas substantially less research effort has been put into studying cognitive representations of output (action) domains. However, researchers typically used the term “peripheral systems” to comprise both input and output

systems, and to contrast them with “central” systems (see [Barrett & Kurzban, 2006](#)), showing that at least theoretically the issue of output system modularity was always an integral part of the discussion. While for input systems the notion of encapsulation typically comprises inaccessibility to the influence of both other (input) modules and central cognitive operations (see [Fodor, 1983](#), for an extensive discussion), this definition needs to be refined for output system representations, where in a trivial sense actions are usually the effect of some variant of cognitive processing (except for the special case of reflexes). Thus, we here provisionally refer to an operational definition of encapsulation of output system representations in terms of immunity to concurrent action processing demands in other output systems. More specifically, we propose that the representations of two output systems should be considered (relatively) encapsulated if corresponding actions can be triggered simultaneously without significant cross-response interference. Based on this ad hoc definition, the silver bullet to the study of encapsulation of output system representations appears to be a paradigm which explicitly tests for potential adverse effects of a secondary response in one output domain on response control in another output domain, involving a comparison of performance under single-response vs. dual-response demands. Up to now, secondary (cognitive) demands have only been utilized to study another potential attribute of modularity, namely the automaticity of computations (see [DeSteno, Bartlett, Braverman, & Salovey, 2002](#); but see [Barrett, Frederick, Haselton, & Kurzban, 2006](#), for a critical assessment of the assumptions underlying this approach), but not to examine encapsulation on the output side of processing.

Given that the empirical study of output system encapsulation is a comparatively novel approach to research in the field of cognitive modularity, some further clarifications and methodological decisions are important. In general, the issue of output system encapsulation can be studied from a peripheral processing (motor control) point of view or from a cognitive (response selection/specification) point of view. The former perspective would be concerned with potential interactions between peripheral effector systems on a late motor level (e.g., corresponding to interactions on final command levels of motor cortex). For example, [D’Avella and Pai \(2010\)](#) conducted research in the field of motor adaptation in muscle control. In contrast, the latter perspective rather relates to cognitive representations of output systems (i.e., on a late central level) and mechanisms associated with implementing an already generated abstract response code (e.g., a spatial code associated with a “left” response) to a code specifying the specific output module (e.g., “vocal”) to ensure full specification of a response prior to its ultimate initiation (e.g., [Huestegge & Koch, 2010a](#)). Note that the exact mechanisms associated with specifying output modalities have often been ignored in research on response selection/specification, even though many dual-task studies involved different output modalities and thus the need to disambiguate responses in terms of their associated effector systems (see, e.g., [Pashler, 1994](#), for a review of typical dual-task studies).

From the peripheral processing point of view, a certain degree of output system encapsulation appears indisputable if we assume that some neural populations in primary motor cortex are strictly linked to specific effectors. At this peripheral level, it also appears conceivable that effector systems compete for other basic neurophysiological resource related to brain metabolism, such as blood oxygen or glucose, which could result in some form of cross-modal response interference even despite the presence of strong effector encapsulation. However, these issues of peripheral output system encapsulation will not be at stake in the present study, which instead rather focuses on cognitive representations of output systems and the underlying determinants of interference.

Interference effects between two responses are traditionally studied in the field of multitasking research. However, research designs utilized in typical multitasking studies appear to be unsuited to address the issue of encapsulation of output system representations. For example, task switching studies do not involve a comparison of single-response vs. dual-response performance and are usually carried out within a single output domain (see [Kiesel et al., 2010](#), for a review). Furthermore, the lack of single-response control conditions also holds for most studies utilizing the psychological refractory period (PRP) paradigm, which is probably the most common method employed in basic multitasking research and involves the systematic manipulation of the temporal interval between the onset of two stimuli that are mapped to two distinct responses (see [Pashler, 1994](#)). Another drawback of PRP studies in particular and many other dual-task studies in general to address the issue of encapsulation of output system representations is the employment of two distinct stimuli. For example, any costs

associated with dual-task conditions (vs. single-task conditions) cannot solely be traced back to interference between representations of response domains, since such costs can also be attributed to the processing of two (vs. one) stimuli, and to the involvement of two (vs. one) response selection processes (based on the two individual stimuli).

Based on these considerations, we reasoned that the issue of output system modularity could best be addressed with a research design involving only one stimulus for both single-response and dual-response conditions. We will refer to such a design as *single-onset cross-modal response paradigm*.

1.3. Previous research on dual-response interference

Interestingly, some previous dual-task studies appear to fulfill the methodological requirements outlined above, although they were not originally intended to address the issue of encapsulation of output system representations. For example, [Holender \(1980\)](#) and [Fagot and Pashler \(1992\)](#) had participants respond to single stimuli with either manual responses, vocal responses, or both. Interestingly, they reported significant dual-response costs for vocal responses, whereas no significant costs were found for manual responses. This pattern of results could be interpreted in terms of encapsulation of manual response control. However, another more recent single-onset study combined manual and oculomotor responses ([Huestegge & Koch, 2009](#)) and reported substantial dual-response cost for manual responses, a finding that clearly contradicts the idea that manual response control is encapsulated. Finally, [Huestegge and Koch \(2013\)](#) presented data on pairwise combinations of manual, vocal, and oculomotor responses within a single-onset paradigm and reported evidence for dual-response costs for each of these modalities.

Taken together, these previous studies appear to have yielded rather mixed results regarding the issue of encapsulation of output system representations. However, this observation also indicates that our working definition of encapsulation of output system representations (see above) appears to be informative in that both empirical outcomes (i.e., the presence/absence of dual-response costs) are principally conceivable. It may well be that differences in study design, for example, with respect to stimulus modalities or response content, may account for the empirical heterogeneity. In sum, these previous results strongly suggest that a more systematic and explicit examination of encapsulation with respect to output system representations is necessary.

It should be noted that all of these previous single-onset studies involved dimensional overlap (see [Kornblum, Hasbroucq, & Osman, 1990](#)) between stimuli and responses (S–R) or across the two responses (R–R). For example, in [Holender's \(1980\)](#) study participants responded to the presentation of a letter with naming and/or typing of that letter (conceptual response overlap on the letter identity dimension). The study by [Fagot and Pashler \(1992\)](#) involved presenting colored letters and naming their color (overlap on the color dimension), and the experiments of [Huestegge and Koch \(2009, 2013\)](#) involved left/right stimuli and corresponding left/right responses (overlap on the spatial dimension). Thus, stimulus and/or response processing always shared a common processing dimension, so that drawing on corresponding limited processing resources might principally be responsible for the occurrence of dual-response costs.

At first sight, one might argue that these studies typically involved S–R and/or R–R compatible responses (e.g., the execution of two “left” responses in the same trial), and compatible conditions should not yield any strong processing conflict compared to studies involving crosstalk based on incompatible codes (e.g., involving spatially incompatible responses, see [Hommel, 1998](#); [Lien & Proctor, 2002](#); [Logan & Schulkind, 2000](#)). However, [Koch \(2009\)](#) observed that dimensional overlap across response sets in dual tasks (e.g., the mere fact that both response sets are spatial) can increase dual-task costs even when the specific responses are spatially compatible, compared to a condition with unrelated responses (i.e., without dimensional overlap). This observation suggests that between-task crosstalk based on dimensional overlap across response sets actually contributes to (adverse) dual-task interference and that even response-compatible actions do not represent a condition that generally facilitates dual-task performance. Within multiple resource theory (e.g., [Wickens, 2008](#)), such findings could be explained by assuming that the two response selection processes draw on a common limited resource pool that is devoted to a specific (e.g., spatial) dimension. Thus, the previous studies on dual-response interference referred to above (which did not take the potentially

important role of dimensional overlap into account) are not suited to finally address the issue of encapsulation of output system representations.

1.4. Rationale of the present study

In the present study, we asked to what extent cognitive representations of output domains are encapsulated by focusing on simultaneously executed oculomotor and vocal actions. How can the study of oculomotor and vocal actions be especially informative regarding the issues of encapsulation and modularity? Previous evidence suggests that several properties of modules appear to apply.

First, both systems are clearly domain specific in the sense that they are responsible for the control of a distinct set of motor actions (ocular movements vs. movements of the larynx). Second, both domains appear to be reasonably distinct with respect to the localization of corresponding control areas in the brain. While cortical control over laryngeal motor responses is mainly instantiated via the primary cortex in the posterior frontal lobe (e.g., [Simonyan & Horwitz, 2011](#)), oculomotor responses are controlled through different networks involving the frontal eye fields and superior colliculi (e.g., [Horn & Leigh, 2011](#)). Third, there is also some tentative evidence for encapsulation of oculomotor control. Specifically, a PRP study by [Pashler, Carrier, and Hoffman \(1993\)](#) suggested that oculomotor responses might be able to bypass central processing bottlenecks, and other previous studies reported no delay of oculomotor responses when additional manual response demands were present (e.g., [Bekkering, Adam, Kingma, Huson, & Whiting, 1994](#); [Lünenburger, Kutz, & Hoffmann, 2000](#)). Fourth, the combination of vocal and oculomotor responses allows us to systematically manipulate dimensional overlap both between stimuli and responses and across responses, which is not possible when two inherently spatial response modalities are combined (e.g., manual and oculomotor responses).

Taken together, the discussion of the previous evidence suggests that a single-onset study utilizing the combination of vocal and oculomotor responses may be especially informative with respect to the issue of encapsulation of output system representations. To examine this issue, we conducted two experiments.

1.5. Overview of Experiments: The role of stimulus modality and dimensional overlap

In Experiment 1, we maximized dimensional overlap by utilizing spatial (left/right) location stimuli and spatial responses in both response modalities (saccades to the left/right [i.e., physical spatial location] and vocal responses “left”/“right” [i.e., semantic spatial meaning]). Moreover, specific spatial code overlap was always present between stimuli and responses (S–R compatibility) as well as across responses (R–R compatibility).

This experiment enabled us to test whether saccades and vocal responses are subject to dual-response interference under most optimal conditions for conflict-free performance (i.e., in a setting that is both S–R and R–R compatible), *but within the limits of dimensionally overlapping response requirements*. Note that any potential conflict on perceptual and central processing stages is avoided by using only one stimulus for both responses (eliminating inter-stimulus conflict on the perceptual stage), and by using two spatially compatible S–R mappings (enabling the implementation of only one common selection process for both responses, see [Fagot & Pashler, 1992](#)). Thus, any residual interference should be attributable to late central processes, where a selected spatial response code (left/right) needs to be linked to a specific response modality code (oculomotor/vocal) prior to response execution.

Importantly, we systematically varied the stimulus modality (auditory vs. visual stimuli) in Experiment 1 to assess the dependency of dual-response costs on input modality. [Fodor \(1983\)](#) reasoned that modularity of cognitive systems may only arise given adequate input modalities (i.e., specialization of modules regarding specific trigger signals). In line with this theoretical claim, previous research in the context of multiple response control suggested that stimulus and response modalities and their specific pairings can substantially affect processing efficiency. For example, vocal responses are more efficiently triggered by auditory stimuli (as opposed to visual stimuli), while manual responses are

more efficiently triggered by visual stimuli (as opposed to auditory stimuli) (input–output modality compatibility [IOMC] effect, e.g., [Hazeltine, Ruthruff, & Remington, 2006](#); [Hazeltine & Wifall, 2011](#); [Stelzel & Schubert, 2011](#); [Stephan & Koch, 2010, 2011](#)). When transferred to the oculomotor domain, these observations suggest that the oculomotor system could be specifically tuned to process visual stimuli, since the same organ (i.e., the eye) is involved in both visual perception and oculomotor control (see [Stephan, Hendler, Koch, & Huestegge, 2013](#)). This view is backed up by literature suggesting a tight neurophysiological link between maps coding locations of visual input and oculomotor targets at the level of the frontal eye fields and the colliculi superiores (e.g., [Hutton, 2008](#); [Munoz, Armstrong, & Coe, 2007](#); [Sweeney, Luna, Keedy, McDowell, & Clementz, 2007](#)). Furthermore, even early visual processing areas (V1) directly project toward oculomotor control areas (e.g., [Isa & Yoshida, 2009](#)). Based on this literature, it appears reasonable to assume that the oculomotor control domain may show characteristics of a module (e.g., encapsulation) only when it receives input through an appropriate (e.g., visual) channel. Note that this hypothesis is further corroborated by the fact that previous studies on oculomotor control that utilized visual input for saccades yielded tentative evidence for encapsulation ([Bekkering et al., 1994](#); [Pashler et al., 1993](#)), whereas studies using auditory stimuli did not ([Huestegge & Koch, 2009, 2013](#)).

The role of dimensional overlap will be specifically addressed in Experiment 2. To this end, we removed dimensional overlap between stimuli and responses (S–R compatibility) by using non-spatial stimuli (the words “Hund”/“Maus”, meaning “dog”/“mouse”) instead of left/right stimuli. This allowed us to introduce a meaningful manipulation of dimensional overlap across responses: While in one condition both responses were spatial (R–R code overlap; similar to Experiment 1), we additionally implemented another condition in which we avoided dimensional overlap across responses by having participants respond with a (spatial) saccade and a non-spatial vocal response (responding “gelb”/“grün”, meaning “yellow”/“green”). If dual-response costs were still present in Experiment 1, Experiment 2 allows us to decide whether such residual dual-response costs could be extinguished through an elimination of dimensional overlap across response requirements (i.e., elimination of the potential for drawing on the same limited spatial processing resource). Again, the absence of dual-response costs in Experiment 2 (especially in conditions without dimensional overlap) would be evidence for encapsulation, whereas the presence of dual-response costs (as well as any evidence for information crosstalk, e.g., priming effects across overlapping response codes) would rather speak against strong encapsulation of output system representations.

2. Experiment 1

2.1. Introduction

Based on the consideration that encapsulation of output system representations may be strongly determined by the modality of the trigger signal, Experiment 1 examined the effects of stimulus modality on dual-response costs for combined saccades and vocal responses. Specifically, we derived the following hypotheses based on the theoretical considerations regarding IOMC effects: If visual input is used to trigger combined oculomotor and vocal responses, one might expect empirical markers for encapsulation specifically for the oculomotor control domain (i.e., the absence of dual-response costs for oculomotor responses and the presence of dual-response costs for vocal responses). Conversely, if auditory input is used to trigger both oculomotor and vocal responses, one might expect the reversed pattern (i.e., absence of dual-response costs for vocal responses and presence of dual-response costs for oculomotor responses).

However, if encapsulation is strong enough to occur irrespective of the specific modality of the trigger signal, one might expect a general absence of dual-response costs across both response modalities. This would then be evidence for particularly strong encapsulation of output system representations. Finally, if encapsulation of output system representations is rather weak (i.e., if there is strong interdependency among response domains irrespective of specific trigger signals), one would expect to find robust dual-response costs for both response modalities regardless of the stimulus modality. Note, however, that any presence of dual-response costs in Experiment 1 could still be due to dimensional

overlap (e.g., the fact that stimulus and response processing draw on common spatial resources), an issue that will be further addressed in Experiment 2.

In Experiment 1, auditory imperative stimuli were presented either to the left or right ear, whereas imperative visual stimuli were arrows pointing to the left or right above central fixation. Note that in line with previous experiments (Stephan, Koch, Hendler, & Huestegge, 2013), we refrained from using peripheral visual targets which would elicit highly automatic orientation responses that would not be comparable to the saccades elicited by the auditory stimuli. We asked participants to perform a spatially corresponding saccade (toward one of two lateralized targets on the screen), a vocal response, or both. Dual-response costs (as an inverse index of encapsulation, see above) were defined as the difference between single-response and dual-response performance for each individual response modality.

2.2. Materials and methods

2.2.1. Participants

Twenty-four university students with normal or corrected-to-normal vision took part in this experiment, eighteen female and six male. Mean age was 22 years ($SD = 2.5$, range: 19–31).

2.2.2. Apparatus and stimuli

Participants were seated 67 cm in front of a 21" cathode ray monitor (temporal resolution: 100 Hz, spatial resolution: 1024×768 pixels). Saccade latencies were registered using a head-mounted Eye-link II infrared reflection system (SR Research, Canada). An eye camera measured the position of the pupil of the right eye with a temporal resolution of 500 Hz and a spatial resolution $<0.022^\circ$. A chin rest was used to minimize head movements.

The experiment was programmed using Experiment Builder (SR Research, Canada) and utilized its built-in voice key functionality. A green fixation cross (size: $1/3^\circ$) in the middle of the screen as well as two green rectangular squares ($1/3^\circ$ each) at 8° to the left and right of the fixation cross remained present throughout. Participants were wearing headphones for the presentation of the imperative auditory stimuli. The auditory stimuli consisted of a 1000 Hz sine wave presented for 80 ms to either the left or right ear with an intensity that was easily audible. The visual stimuli consisted of left vs. right arrows (size: $1/3^\circ$) presented for 80 ms right above the central fixation cross. Vocal responses were recorded throughout the experiment by utilizing a voice key.

2.2.3. Procedure

Each trial began with the presentation of the left vs. right imperative auditory or visual stimulus. Subjects were instructed to respond as fast and accurately as possible either by moving their gaze to the spatially compatible square on the screen (saccade response in single blocks), by pronouncing the spatially compatible direction (vocal response in single blocks), or both (dual-response blocks).

In the two conditions that require saccades (saccade response in single and dual-response blocks), subjects were instructed to return to the central fixation cross after response. Each subject completed nine blocks of 30 trials with one stimulus type (e.g., visual), followed by the same amount of blocks with the remaining stimulus type (e.g., auditory, with stimulus modality counterbalanced across participants). Within each block, left vs. right stimuli were presented in a randomized sequence with an inter-stimulus interval of 3000 ms. Prior to each experimental block, subjects underwent a calibration routine in which the spacebar of the keyboard was used. Prior to the experiment, subjects performed 30 practice trials (10 trials of each response condition) that were not further analyzed.

2.2.4. Design

The variables response modality (saccade vs. vocal response), stimulus modality (visual vs. auditory), and response load (single- vs. dual-response condition) were manipulated intraindividually and blockwise. For example, one subject performed three identical sequences of three blocks in the order "saccade response (single response)", "dual response", "vocal response (single response)". A Latin square design ensured that each of the three conditions occurred equally often at the three possible positions within a sequence across participants. As dependent variables, we measured saccade RTs, vocal RTs, and errors.

2.3. Results

Because of blinks or measurement error, we discarded 4.1% of trials in single saccade response blocks and 0.5% of trials in dual-response conditions. In the single vocal-response blocks, 7.3% of trials with erroneously executed saccades were excluded.

Table 1 shows the mean RTs of saccades and vocal responses in single- and dual-response conditions. An ANOVA with response load, stimulus modality, and response modality as independent variables yielded a significant main effect for response load (single vs. dual), $F(1,23) = 45.41$, $p < .001$, $\eta_p^2 = .66$, indicating longer RTs in dual-response vs. single-response conditions (509 ms vs. 481 ms). The main effect of response modality (saccade vs. vocal) was significant, too, $F(1,23) = 670.28$, $p < .001$, $\eta_p^2 = .97$. RTs for saccades were generally shorter than RTs for vocal responses (262 ms vs. 728 ms). There was a significant interaction of response modality and response load, $F(1,23) = 14.41$, $p = .001$, $\eta_p^2 = .39$, indicating larger dual-response costs for vocal responses than for oculomotor responses. Separate paired *t*-tests (one-tailed) revealed significant dual-response costs for vocal responses (auditory stimuli: 49 ms; visual stimuli: 44 ms) and for saccades (auditory stimuli: 7 ms; visual stimuli: 10 ms), all $ps < .05$ (see Fig. 1). Note that due to the substantial RT differences between response modalities, Fig. 1 additionally displays proportional dual-response costs (middle panel).

There was a main effect of stimulus modality on RTs, $F(1,23) = 61.15$, $p < .001$, $\eta_p^2 = .73$, indicating longer RTs for visual stimuli (523 ms) than for auditory stimuli (467 ms). The interaction of stimulus modality and response modality was significant, too, $F(1,23) = 33.95$, $p < .001$, $\eta_p^2 = .60$, indicating that the overall RT difference between saccades and vocal responses was more pronounced for auditory stimuli than for visual stimuli. There was no interaction of stimulus modality and response load and no significant three-way interaction, both $F_s < 1$, showing that the pattern of dual-response costs was unaffected by stimulus modality.

Vocal response errors occurred rarely ($M = 0.2\%$ in single-response conditions vs. $M = 1\%$ in dual-response conditions), $t(23) = 1.21$, $p = .239$. Saccade error rates in dual-response conditions ($M = 1.6\%$) were significantly higher than in single-response conditions ($M = 0.5\%$), $t(23) = 3.33$, $p = .003$, thus not contradicting RT patterns. Saccades were initiated earlier than vocal responses throughout in dual-response trials.

Finally, we additionally tested whether the data provide evidence for another type of interaction between the two response domains (and thus, against encapsulation) in dual-response conditions. Specifically, we asked whether the two responses within a trial were generally grouped with a fixed inter-response interval (IRI). This could occur due to a strategy to withhold the first response (saccade) until the second response is prepared. If this was the case, one would expect that the first response is always executed time-locked to the second response, a finding that would count as empirical evidence for (cognitively controlled) cross-response interaction. Across the RT distribution of the slower (vocal) response, vocal RTs were thus ranked for each subject and divided into four bins (i.e., fastest to slowest trials). A one-way ANOVA with vocal RT bin as an independent variable and IRI as a dependent

Table 1

Response times (ms) in Experiment 1 and 2 as a function of response modality (saccade vs. vocal), response load (single vs. dual), and response relatedness (related vs. unrelated). SEs are presented in parentheses.

Experiment	S–R relation	R–R relation	Stimulus modality	Response modality	Single-response RTs (ms)	Dual-response RTs (ms)
1	compatible	compatible	visual	saccade	299 (6.9)	309 (6.9)
				vocal	720 (21.1)	764 (22.1)
	compatible	compatible	auditory	saccade	217 (6.6)	224 (7.3)
				vocal	689 (21.1)	738 (23.6)
2	arbitrary	compatible	auditory	saccade	300 (10.3)	317 (12.6)
				vocal	701 (20.7)	724 (27.5)
	arbitrary	arbitrary	auditory	saccade	300 (10.3)	320 (13.2)
				vocal	732 (23.4)	783 (28.8)

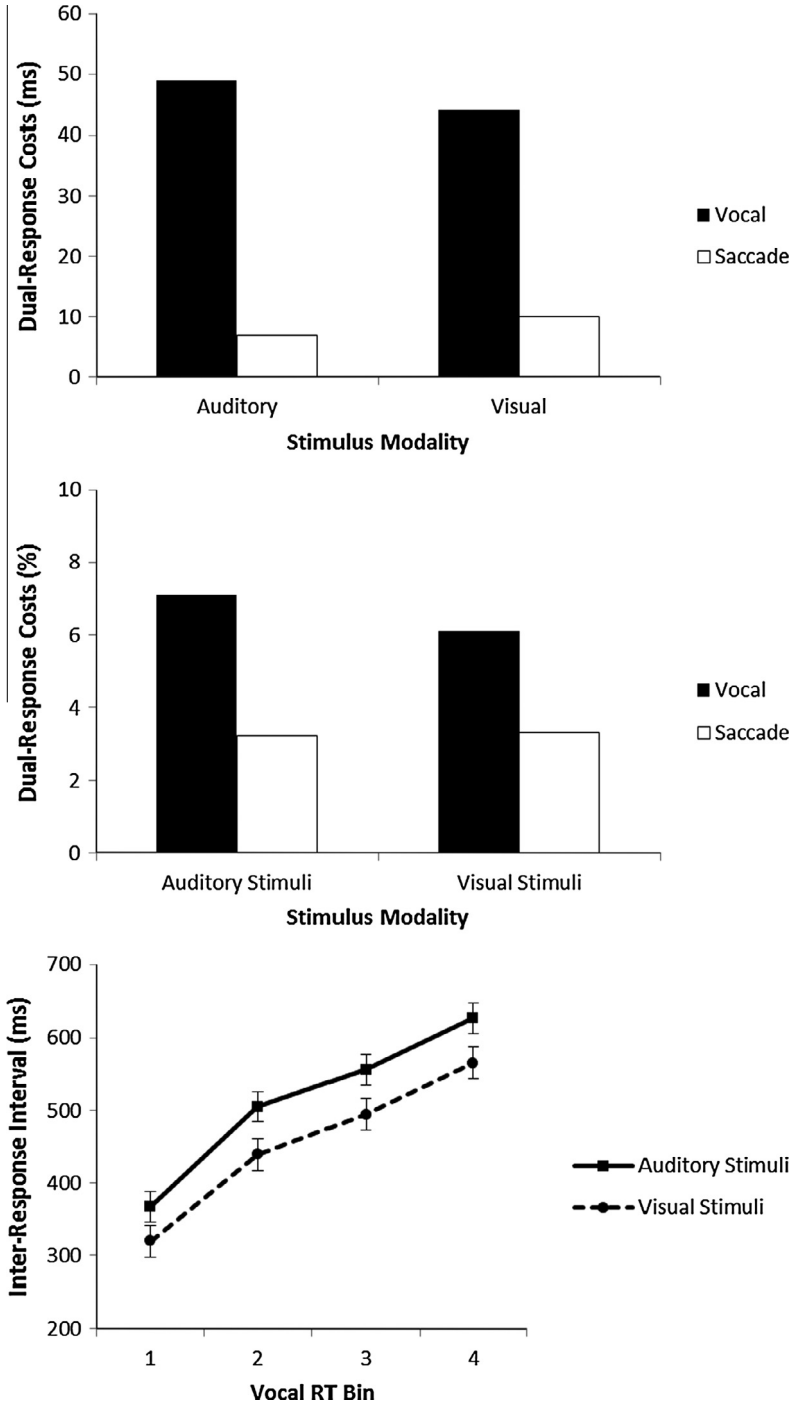


Fig. 1. Mean absolute (upper panel) and proportional (middle panel) dual-response costs and the distribution of inter-response intervals (lower panel) in Experiment 1.

variable resulted in a significant linear trend for both visual stimuli, $F(1,23) = 217.13, p < .001, \eta_p^2 = .90$, and auditory stimuli, $F(1,23) = 231.54, p < .001, \eta_p^2 = .91$, indicating that IRIs increased with increasing vocal latencies (see Fig. 1). Thus, responses were not generally grouped with a fixed IRI, a finding that can be interpreted against a strategic (cognitively mediated) interaction across response modalities.

2.4. Discussion

Experiment 1 revealed significant dual-response costs in RTs for both saccades and vocal responses, regardless of stimulus modality. Errors occurred only rarely, and the observed pattern was not in line with a speed-accuracy trade-off as an alternative explanation of the data. Thus, the present results indicate a strong interdependency among response domains irrespective of specific trigger signals, a conclusion that is at odds with a strong account of encapsulation of output system representations. Note that the observed cross-modal interference effects were present regardless of the fact that responses were always spatially S–R and R–R compatible, representing a condition which – within the limits of an overlapping spatial response dimension on a more general level – appears to be quite optimal for interference-free performance. Within the classic framework of stage processing, these results point to a late central source of interference. Specifically, any potential conflict on perceptual and central processing stages was avoided by using only one stimulus for both responses (eliminating inter-stimulus conflict on the perceptual stage) and by using two spatially compatible S–R mappings (enabling the implementation of only one common selection process for both responses, see Fagot & Pashler, 1992). Thus, the robust amount of interference should be mainly attributable to late central processes, where a selected spatial response code (left/right) is linked to a specific response modality code (oculomotor/vocal). Possibly, this specific type of interference could (at least partially) be due to the need to hold the goal to execute two responses (including information about respective response modalities) in working memory, perhaps in verbal format (e.g., Miyake, Emerson, Padilla, & Ahn, 2004). Regardless of the specific mechanisms, the results generally suggest that even under fairly optimal conditions for dual-response execution robust performance costs still emerge, a finding that is at odds with a strong encapsulation account.

Interestingly, dual-response costs for the individual response modalities were not modulated by the input modality (visual vs. auditory). Based on previous reports of IOMC effects in the context of other dual-task paradigms (Hazeltine & Wifall, 2011; Hazeltine et al., 2006; Stelzel & Schubert, 2011; Stephan & Koch, 2010, 2011), one might have expected that a specific stimulus type (auditory) could be appropriate for one response type (vocal responses) but not for the other response type (saccades), and the reversed pattern may hold for the other (visual) stimulus type. However, this reasoning would imply that the auditory (vs. visual) stimulus condition should yield substantially reduced dual-response costs for vocal responses, which we did not observe. Similarly, we did not observe any evidence for encapsulation when the oculomotor output system is fed with (probably more appropriate) visual instead of auditory input. Instead, our results are more in line with previous work in the field of task switching suggesting that in terms of dual-response coordination, saccades respond equally well to both visual and auditory stimulation (Stephan et al., 2013).

The overall slower RTs for visual (vs. auditory) stimuli may be due to additional time associated with decoding the meaning of the symbolic visual cue, whereas the auditory stimulus was defined on a physical level (tone on left vs. right ear). Importantly, however, these potential differences in stimulus processing did not substantially affect the observed pattern of dual-response costs.

It is important to note that previous research in the field of dual-task control has repeatedly shown that it is principally possible to establish experimental conditions for observing saccades without any sign of central processing conflict or dual-response costs, namely when peripheral visual stimuli are used as a trigger signal (e.g., Bekkering et al., 1994; Lünenburger et al., 2000; Pashler et al., 1993). However, our present results clearly show that these findings do not generalize to the visual modality per se, but are rather restricted to the special case of orientation responses triggered by salient peripheral stimuli. Probably, any need to transform spatial coordinates (from stimulus characteristics to required response demands) may suffice to produce cross-modal costs across response domains. Thus, it is important to explicitly address the issue of dimensional (here: spatial) overlap (see Experiment 2). Probably, one could also conclude that the influence of input modality (in terms of receptor systems

involved) may be less relevant for determining the amount of interference than the specific presentation mode of the stimulus (e.g., peripheral vs. central visual stimuli).

Finally, the IRI analysis was conducted to test whether response grouping occurred in dual-response conditions. At first sight, such a strategy could be considered likely since both responses were triggered by a common response selection process based on a single stimulus (Fagot & Pashler, 1992). Thus, one might have expected a relatively constant IRI that should mainly reflect the temporal difference in response execution processes across both effector systems. However, while the IRI data certainly do not rule out the possibility that grouping may have occurred in some of the dual-response trials, the substantial variability of the IRI suggests that additional processing appears to occur after the common response selection stage, but prior to response execution. This additional processing may well reflect the process of binding the selected spatial response code to specific response modalities (see above). Finally, the systematic variability of the IRI also renders it unlikely that in the majority of trials the first response was withheld until the preparation of the second response was completed.

3. Experiment 2

3.1. Introduction

While Experiment 1 showed that dual-response costs occurred under optimal conditions for dual-response processing in the presence of (spatial) dimensional overlap across both responses and across stimuli and responses, Experiment 2 was designed to explicitly examine the role of dimensional overlap across responses (R–R overlap) in the absence of dimensional overlap between stimuli and responses. To this end, we had participants simultaneously respond to non-spatial auditory stimuli (i.e., the words “Hund”/“Maus”, meaning “dog”/“mouse”) with a (spatial) saccade (the same response as in Experiment 1) and either a spatial (“left”/“right”) or a non-spatial (“yellow”/“green”) vocal response. While using the same stimuli as in Experiment 1 would have made the present experiment more comparable at first sight, we decided against this option because then both responses would differ with respect to their dimensional overlap with the relevant stimulus dimension. Hence, we decided to implement an arbitrary S–R mapping (i.e., no dimensional overlap of stimulus set and response set, see Kornblum et al., 1990), and the crucial manipulation refers to the presence vs. absence of dimensional (here: spatial) overlap across responses. Given that Experiment 1 yielded evidence against a substantial influence of stimulus modality, Experiment 2 involved only one stimulus modality (auditory).

Based on the theoretical rationale outlined in the introduction, we put the following predictions to the test: A strong encapsulation account would predict lower (or absent) dual-response costs in the absence (vs. presence) of dimensional overlap across responses (and across stimuli and responses), since the removal of dimensional overlap should minimize the potential of drawing on shared (spatial) processing resources across response domains. In contrast, a strong cross-modal interaction account would probably predict even greater dual-response costs in the absence (vs. presence) of dimensional overlap, since response-based interaction across response modalities in Experiment 1 may have resulted in information crosstalk across spatially compatible responses, eventually resulting in priming benefits for the second of the two responses.

3.2. Materials and methods

3.2.1. Participants

Eighteen new participants (university students) participated in this experiment (mean age = 23.55 years, $SD = 2.28$), ten female and eight male. They had normal or corrected-to-normal vision and received course credits or money for participation.

3.2.2. Apparatus and stimuli

The apparatus and overall setup was the same as in Experiment 1. Unlike in Experiment 1, we used two spoken words as imperative auditory stimuli, (“Hund”/“Maus”, meaning dog/mouse) with a

spoken duration of 300 ms for both words. We used these stimuli because we reasoned that they are not (or at least not strongly) associated with the spatial concepts of left/right (unlike, e.g., tones of different pitch). Given the lack of an effect of input modality on dual-response costs in Experiment 1, we only utilized uni-modal (i.e., auditory) stimuli in Experiment 2.

3.2.3. Procedure

Prior to the first trial of each block the display screen (including the fixation cross and the two saccade targets) was presented for 1 s. Each trial started with the presentation of one of the two imperative auditory stimuli. There were five different blocks of trials. Participants responded either with a) a saccade toward one of the two lateralized saccade targets (*single-response saccade condition*, e.g. “Hund”: left saccade, “Maus”: right saccade), b) a vocal response by saying either “grün” (green) or “gelb” (yellow) (*single-response R–R non-overlap vocal condition*, e.g. “Hund”: “grün”, “Maus”: “gelb”), c) a vocal response by saying “links” (left) or “rechts” (right) (*single-response R–R overlap vocal condition*, e.g. “Hund”: “links”, “Maus”: “rechts”), d) a saccade and an unrelated vocal response (*non-overlap dual-response condition*, e.g. “Hund”: left saccade and “grün”, “Maus”: right saccade and “gelb”), and e) a saccade and a related vocal response (*overlap dual-response condition*, e.g. “Hund”: left saccade and “links”, “Maus”: right saccade and “rechts”). The overlap dual-response condition only included spatially compatible responses. The inter-stimulus interval amounted to 2500 ms. Participants were asked to respond as fast and accurately as possible. Vocal response errors were coded manually. Note that we carefully ensured that vocal responses in overlap and non-overlap conditions did not differ with respect to their overall difficulty, which is a precondition for meaningful comparisons between dual-response costs across conditions. First, S–R mappings were equally arbitrary for both spatial and non-spatial vocal responses (see above). Second, all words (“links”/“rechts”, “grün”/“gelb”) were closely matched in logarithmized lemma frequency (all within the narrow range of 1.5–1.8) according to CELEX database (celex.mpi.nl). Third, while it is not possible to use single-response RT data as an index of difficulty here (since this would compare voice key latencies for different responses with different associated pronunciation characteristics), comparable difficulty should result in similar error rates. To anticipate the results, single-response error rates for vocal responses in the related vs. unrelated responses conditions were virtually identical (1.3% vs. 1.2%, $t < 1$).

Each participant completed fifteen blocks of 30 trials each with calibration routines at the beginning of each block. A fixed sequence of the five different blocks was repeated three times for each participant. The order of the five blocks within each sequence was randomized. Within each block, the order of imperative auditory stimuli was randomized, and the assignment of response direction to the imperative auditory stimuli was counterbalanced. The experiment lasted for about 45 min.

3.2.4. Design

The three independent within-subject variables were response modality (saccade vs. vocal), response load (single- vs. dual-response condition), and R–R overlap (overlap vs. non-overlap). The dependent variables were RTs (for correct trials only) and errors.

4. Results

Because of blinks or measurement error, we discarded 0.28% of trials in the single saccade conditions and 1.22% of trials in the dual-response conditions. In the single vocal-response conditions, 1.15% of trials with erroneously executed saccades were excluded.

Table 1 depicts mean RTs as a function of the experimental conditions. The three-way ANOVA revealed a significant main effect of response modality on RTs, indicating longer RTs for vocal responses (735 ms) as compared to saccades (309 ms), $F(1, 17) = 310.83$, $p < .001$, $\eta_p^2 = .95$. Furthermore, there was a significant main effect of response load, indicating shorter RTs for single response conditions (508 ms) as compared to dual-response conditions (536 ms), $F(1, 17) = 17.94$, $p = .001$, $\eta_p^2 = .51$. There was no significant interaction between response modality and response load, $F < 1$. Overall, the general result pattern reported so far matches the pattern from Experiment 1. There

was also a significant main effect of R–R overlap, indicating longer RTs for non-overlapping responses (534 ms) as compared to overlapping responses (511 ms), $F(1, 17) = 26.34$, $p < .001$, $\eta_p^2 = .61$.

Importantly, there was a significant interaction between response condition and R–R overlap with greater dual-response costs for non-overlapping responses (35 ms) as compared to overlapping responses (20 ms), $F(1, 17) = 4.57$, $p = .047$, $\eta_p^2 = .21$. Moreover, there was a significant interaction between response modality and R–R overlap, $F(1, 17) = 27.02$, $p < .001$, $\eta_p^2 = .61$, indicating that the overall difference between saccades and vocal responses was greater for the non-overlapping responses condition (448 ms) than for the overlapping responses condition (405 ms).

Finally, there was also a significant three-way interaction, $F(1, 17) = 5.62$, $p = .030$, $\eta_p^2 = .25$, indicating that the difference in dual-response costs for non-overlapping vs. overlapping responses was specifically pronounced for vocal responses (23 ms for overlapping vs. 51 ms for non-overlapping response conditions), and did not differ in the same way for saccades (17 ms for related and 20 ms for unrelated conditions, see Fig. 2, which also displays proportional dual-response costs).

The corresponding three-way ANOVA for accuracy only revealed a significant main effect of modality, $F(1, 17) = 11.09$, $p = .004$, $\eta_p^2 = .40$, indicating a higher error rate for vocal responses (1.2%) as compared to saccade responses (0.1%). None of the remaining main effects or interactions were statistically significant, all $ps > .10$. In dual-response conditions, saccades were nearly always executed first, and eliminating the few reversals of this response sequence (<0.5% of all trials) did not change the pattern of results.

IRIs were analyzed in the same way as in Experiment 1 (see Fig. 2). As in the previous experiment, we found significant linear trends (related responses: $F(1, 17) = 101.75$, $p < .001$, $\eta_p^2 = .86$; unrelated responses: $F(1, 17) = 123.99$, $p < .001$, $\eta_p^2 = .88$).

4.1. Discussion

The data from conditions with dimensional overlap across responses (which did not involve spatial S–R dimensional overlap) mainly replicated the main findings from Experiment 1 (which included both R–R and S–R overlap), namely significant dual-response costs for saccades as well as for vocal responses. The three-way interaction indicates that the difference in dual-response costs between saccades and vocal responses is more pronounced in the non-overlapping responses condition compared to the overlapping responses condition. Thus, at least in the non-overlapping responses condition vocal responses exhibited greater dual-response costs than saccades. Although Experiments 1 and 2 cannot be compared directly due to the different stimuli and hence different levels of S–R compatibility, the overall similarity of the result pattern suggests that dual-response costs are not substantially affected by the presence or absence of spatial overlap between stimulus and response codes (S–R compatibility). This can be interpreted as evidence for rather general dual-response execution costs associated with co-ordinating two responses and thus against strong encapsulation of output systems.

Most importantly, however, the pattern of the data showed that the execution of two responses without dimensional overlap significantly *increased* dual-response costs for vocal responses. If general (S–R or R–R) code overlap would be the single source of interference in executing two responses across output domains, one would expect the opposite data pattern, namely a decrease of dual-response costs in the absence of dimensional overlap (see Koch, 2009, for corresponding data within a paradigm involving two separate stimuli for two simultaneous responses). Instead, the results indicate that for spatially compatible responses (with code overlap), the selection of the (slower) vocal response appeared to benefit from response code priming through the pre-activation of the same spatial code for the first (saccade) response. Since in non-overlapping responses this priming effect cannot take place, this may have slowed down the activation of the response code in the slower (vocal) response. These priming effects demonstrate that information crosstalk occurs across response domains, a finding that is again compatible with the assumption of strong interdependency across cognitive representations of output systems.

Finally, the IRI analyses in the dual-response conditions again did not support the idea that the first response was generally withheld in order to produce a coordinated compound response after completion of the programming of the second response. Thus, there is no evidence for a strong temporal coordination strategy in dual-response execution.

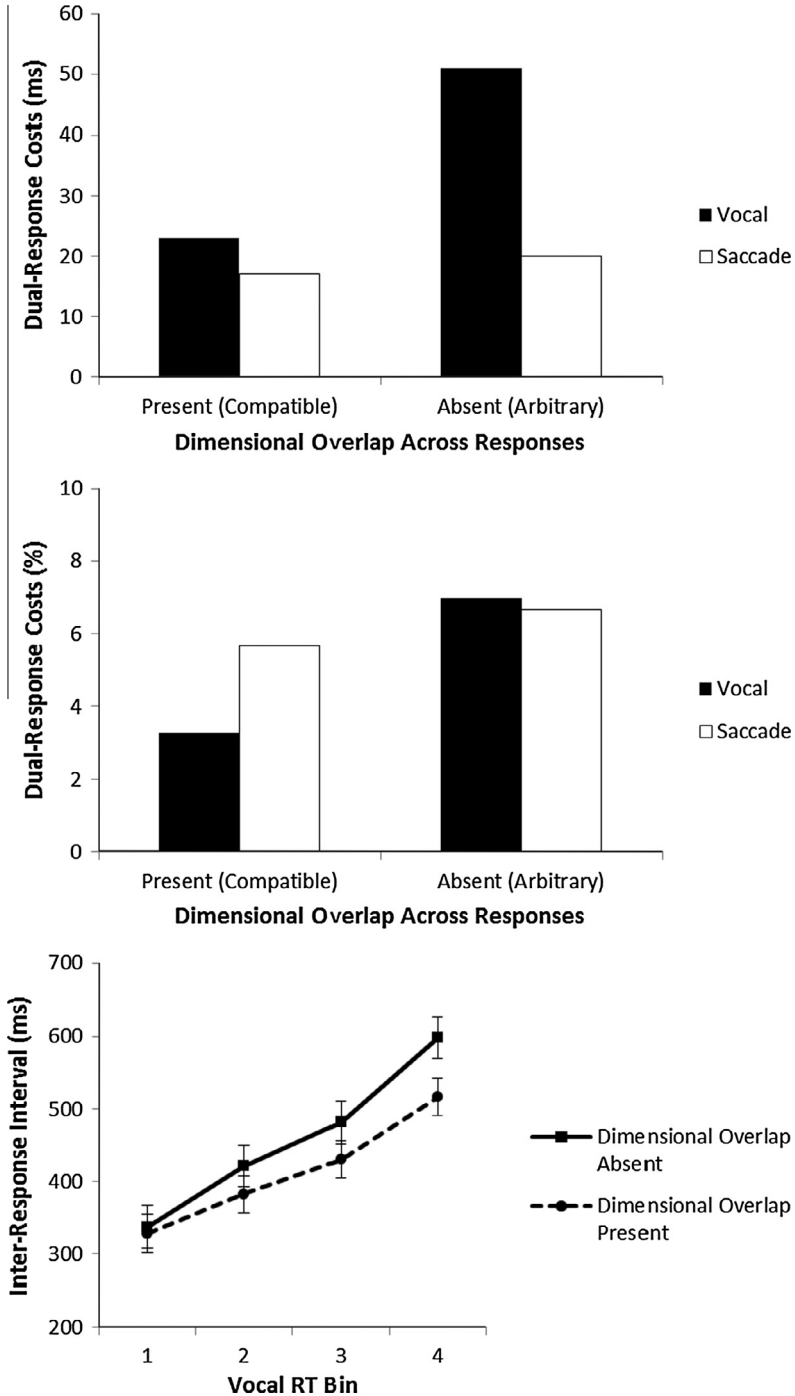


Fig. 2. Mean absolute (upper panel) and proportional (middle panel) dual-response costs and the distribution of inter-response intervals (lower panel) in Experiment 2.

5. General discussion

The assumption of distinct functional modules is an important presupposition in many theories in the field of cognition (Barrett & Kurzban, 2006; Fodor, 1983). Previous research on modularity predominantly studied input domains (e.g., vision; see Marr, 1982) and/or central processes (massive modularity; see Carruthers, 2005; Frankenhuys & Ploeger, 2007; Pinker, 2005; Sperber, 2005; Tooby & Cosmides, 1992), whereas output domains (i.e., systems devoted to motor control of specific effector systems) were largely neglected. As a consequence, there is only little research regarding the general interplay of output modalities, including manual action, vocal action, and oculomotor action when these actions are not inherently coordinated through the presence of a common action target.

The present study focused on a potential key feature of cognitive modules, namely their encapsulation, that is, their immunity to the influence of other domains (Pylyshyn, 1999). While for input systems the notion of encapsulation typically comprises inaccessibility to the influence of both other (input) modules and central cognitive operations (see Fodor, 1983, for an extensive discussion), this interpretation needs to be refined for output system representations, where in a trivial sense actions are typically the effect of some variant of cognitive processing (except for the special case of reflexes). Thus, we here defined encapsulation of output system representations in terms of immunity of action processing in one modality to other concurrent action processing demands in other output systems (i.e., interaction on a late central level). Based on theoretical considerations and previous empirical work, our main goals were to determine the extent of encapsulation of output system representations and to specify relevant factors determining the overall amount of output system interdependency, which can be regarded as a heuristically valuable continuum with the poles of strong encapsulation on the one side and strong interaction/crosstalk on the other side. We reasoned that the question of encapsulation of output system representations can best be addressed through the study of interference effects (and their potential determinants) in simultaneous oculomotor and vocal responses, and by employing an experimental paradigm involving only one stimulus that either triggers one or two responses (single-onset paradigm; Fagot & Pashler, 1992; Holender, 1980). Since the single-onset paradigm is known to avoid interference on perceptual and central response selection stages of information processing (Fagot & Pashler, 1992), it appears to be particularly suited to address late central mechanisms of cross-response control, and thus the issue of encapsulation of output system representations. However, previous single-onset studies were not designed to address the issue of encapsulation of output system representations, and thus did not control for potentially important factors such as input modality and dimensional overlap (Fagot & Pashler, 1992; Holender, 1980; Huestegge & Koch, 2009; Huestegge & Koch, 2013).

5.1. Evidence against strong encapsulation: Dual-response costs

Fodor (1983) proposed that modules should exhibit domain specificity in that they may only operate on a specific kind of input. In Experiment 1, we explicitly studied the role of stimulus modality under dimensionally overlapping response requirements by testing hypotheses derived from previous research and theory on input–output modality compatibility (e.g., Hazeltine & Wifall, 2011; Hazeltine et al. 2006; Stelzel & Schubert, 2011; Stephan & Koch, 2010; Stephan & Koch, 2011). Since stimuli and responses were spatially compatible throughout, one could assume that this represents an optimal condition for interference-free performance (within the limits of dimensionally overlapping response requirements). In contrast, Experiment 2 addressed the issue of dimensional overlap across responses. A key result of the present study is that substantial dual-response costs emerged for both response modalities across the whole variety of conditions and experiments. In Experiment 1, dual-response costs were observed in the presence of both S–R and R–R compatibility and for both stimulus modalities, suggesting that none of the response modalities involved displays strongly encapsulated performance. Interestingly, we also observed mutual interference between responses when there was no dimensional overlap between stimuli and responses and even when spatial dimensional overlap across responses was additionally removed (Experiment 2). If general (S–R and/or R–R) dimensional overlap was the central source of interference (e.g., Koch, 2009), one would have expected a decrease

of dual-response costs in the absence of dimensional overlap. Note that without Experiment 1, one could argue that the dual-response costs in Experiment 2 may simply be due to comparatively difficult response selection processes based on arbitrary S–R mappings. However, Experiment 1 rules out this possibility by showing that even under most beneficial conditions (spatially compatible S–R and R–R mappings) substantial dual-response costs emerged. Taken together, our consistent finding of dual-response costs throughout response modalities, conditions, and experiments is clearly at odds with the idea of encapsulated output systems.

These results appear especially interesting since for a long time, it has been assumed that saccades are special in the sense that they might not be affected by a central processing bottleneck (Pashler et al., 1993) or any kind of dual-response interference (e.g., Bekkering et al., 1994; see Huestegge, 2011, for a review). However, we would not like to conclude that it is impossible to find circumstances under which dual-response costs (e.g., for saccades) can be eliminated. In fact, previous research as well as research from our own lab suggests that special cases like a) visual orientation responses of the eyes toward suddenly appearing, salient stimuli in the periphery (Bekkering et al., 1994; Pashler et al., 1993) or b) eye movements in the context of reaching and grasping (e.g., Lünenburger et al., 2000) may occur without costs under additional manual action demands. However, we feel that a truly encapsulated system should exhibit a more general immunity to interference, at least within the limits of a certain trigger dimension (e.g., visual stimuli; see Fodor, 1983; see also theoretical explanations of the IOMC effect, e.g., Stephan & Koch, 2011). For example, one would not call motor control of leg movements encapsulated just because there may be instances of leg responses that can be elicited without any interference (e.g., the special case of patellar reflexes). Taken together, our present results thus show that there is clearly no generalized immunity to interference for the output systems under study. Therefore, our results represent evidence against strong accounts of encapsulation of output system representations.

5.2. Evidence against strong encapsulation: Information crosstalk

In Experiment 2, the data pattern showed that the execution of two unrelated responses (without dimensional overlap) significantly increased dual-response costs for the (slower) vocal responses. This finding indicates that in the case of compatible responses, the selection of the (slower) vocal response could benefit from response code priming through a pre-activation of the same spatial code for the first (saccade) response. Since in unrelated responses this priming effect cannot take place, this may have slowed down the activation of the response code in the slower (vocal) response in dual-response conditions. Thus, the present benefit for shared response codes (see also Koch & Prinz, 2002) represents an instance of beneficial (i.e., facilitating) crosstalk. Probably, the fact that responses were always logically coupled in our study (i.e., in dimensional overlap conditions, a left saccade was always followed by a “left” vocal response) may have further increased the potential of finding beneficial crosstalk, since experimental designs involving logically independent responses revealed negative effects of dimensional overlap on performance (i.e., performance in trials without dimensional overlap was still superior than that in trials involving compatible responses; Koch, 2009). Most importantly, however, these instances of information crosstalk between responses in different modalities are clearly not compatible with the idea of strongly encapsulated output systems.

5.3. Mechanisms of multiple-response control

In the following, we will further speculate about the specific mechanisms behind the observed interference effects in dual-response control by referring to previous empirical findings and theoretical frameworks in the related field of dual-task control. Most notably, Fagot and Pashler (1992) convincingly demonstrated that responding to (the same attribute of) a single stimulus with two responses (in different modalities) involves only one single central response selection process, as opposed two distinct, serial response selection processes (i.e., due to a central response selection bottleneck, see Anderson, Taatgen, & Byrne, 2005; Pashler, 1994) that occur whenever two stimuli (or two different aspects of one stimulus) need to be transformed into two responses.

When we transfer these theoretical conclusions by [Fagot and Pashler \(1992\)](#) to our present design, we can effectively rule out two potential loci of interference: First, there is no potential for any conflict on the perceptual processing stage, since only a single stimulus was implemented to trigger both responses (eliminating inter-stimulus conflict on the perceptual stage). Second, there was also no potential for conflict on a central stage that has traditionally been associated with response selection, since only a single, common response selection process was involved ([Fagot & Pashler, 1992](#)). Thus, the robust amount of interference should be mainly attributable to later (post-selection) processes. Probably, the interference observed here is due to mechanisms associated with implementing an (already generated) abstract response code (e.g., a spatial code associated with a “left” response) to a specific output module code (e.g., “vocal”) to ensure full specification of a response prior to its ultimate initiation (e.g., [Huestegge & Koch, 2010a](#)). Note that this specification of output modalities has traditionally been ignored in research on response selection/specification, even though corresponding studies often utilized research designs in which responses needed to be disambiguated in terms of their associated effector systems (see, e.g., [Pashler, 1994](#), for a review of traditional dual-task studies). Probably, the linkage between an abstract response code (as a result of the central S–R translation process) and a code defining the corresponding effector modality is held in working memory in terms of a goal to execute the two responses (including information about respective response modalities), perhaps in verbal format (see [Miyake et al., 2004](#)). Regardless of the specific mechanisms, the results generally suggest that even under fairly optimal conditions for dual-response execution robust performance costs still emerge, a finding that is at odds with a strong encapsulation account.

There is further evidence in our data that suggests additional mechanisms (e.g., in terms of the aforementioned linkage of spatial response codes to response modality codes) after the central response selection stage. Specifically, the notion of a single response selection process in our paradigm *without any additional processing afterwards* would suggest that the onset of response execution in both modalities is solely determined by the individual temporal characteristics of the response modalities, so that one would expect a constant (or near-constant) IRI in dual-response conditions across the distribution of vocal response latencies, which we did not observe. This renders it unlikely that the result of the computations related to response selection were directly transferred to a single execution signal for both modalities. Instead, these data further underline that another (late central) process, most probably the individual specification of response modalities, must be involved.

Another class of dual-task models that might be considered to frame the present data (and which is in fact opposed to the idea of motor system encapsulation) is based on the idea of capacity sharing for simultaneous response control (e.g., [Meyer & Kieras, 1997](#); [Navon & Miller, 2002](#); [Tomblu & Jolicoeur, 2003](#)), at least when we assume that the coordination of two responses across response domains per se always implies a certain amount of central processing. Especially the assumption of parallel processing in these models nicely fits our finding of a mutual influence between both response modalities. While the notion of capacity sharing is usually relatively unspecific with respect to the exact underlying mechanisms, one more specific explanation of the general occurrence of dual-response costs within this framework is that interference may occur due to the mere fact that two responses or motor programs (instead of one) need to be selected and prepared (e.g. [Logan & Gordon, 2001](#)). However, this would not readily explain why some of the previous single-onset studies reported evidence for response execution without interference from other concurrent response demands (e.g., [Holender, 1980](#)) without further assumptions. Taken together, we thus think that our interpretation (i.e., that the dual-response costs in our paradigm mainly reflect processes associated with a post-response selection modality-specification stage at a late central processing level) appears to be a plausible alternative.

It should be noted that the dual-task models mentioned above are generally limited with respect to explaining other core results of the present study. More specifically, these frameworks are inherently “content-blind”, that is, information processing for simultaneous responses is modelled without reference to the specific attributes of the responses involved (i.e., whether a response requires a left vs. right response, or whether there is a content-specific interrelation between attributes of both tasks based on a shared spatial dimension). In contrast to these models, other frameworks of multitasking explicitly refer to such content-based interference, for example, crosstalk models ([Navon & Miller, 1987](#)), or multiple resource models ([Wickens, 2008](#)).

However, while crosstalk models alone cannot account for interference in situations without any dimensional overlap between responses (which we observed in Experiment 2), multiple resource models cannot readily account for performance differences between two dual-task situations that involve the same input-, central- and output-related processing modalities. Additionally, the models discussed so far have in common that they were mainly designed to explain *adverse* effects of executing multiple responses by assuming processing bottlenecks, crosstalk, resource conflicts, or executive processes with the aim to coordinate multiple response demands. Thus, these frameworks cannot really capture our present facilitation effects based on cross-response priming (i.e., beneficial crosstalk) without additional assumptions. A notable exception is a hybrid model proposed by [Lehle and Hübner \(2009\)](#), which extends previous (content-blind) capacity-sharing models (e.g., [Tombu & Jolicoeur, 2003](#)) by additionally taking (content-based) crosstalk and priming effects into account.

As outlined above (see section on information crosstalk), a more specific explanation of beneficial crosstalk effects could refer to the assumption that the activation of a response code in the second of two responses (e.g., the spatial code “left” for the vocal response) benefits from a pre-activation of the same code from the first response (e.g., the spatial code “left” from the leftward saccade). Recently, we proposed a framework of cross-modal action control based on such codes (i.e., spatial codes and modality codes which need to be bound together in accordance with task instructions) that principally allows us to cover such facilitation effects (see [Huestegge, 2011](#); [Huestegge & Koch, 2010a](#)). Specifically, the activation of one spatial code (representing, for example, “left”) may be used for both the execution of the first response (saccade) and the second response (vocal) within a trial, without the need to activate another spatial code for spatially specifying the second response.

5.4. Implications for related fields of research

While our present study focused on response-based interference and modularity issues, our results might also bear some implications for more general aspects of the interplay between spatial attention and (inherently spatial) oculomotor control. Specifically, spatial dimensional overlap between saccade execution and additional perceptual tasks (i.e., tasks without immediate response requirements; see [Huestegge & Koch, 2010b](#); [Kowler, Anderson, Doshier, & Blaser, 1995](#)) may have been a major cause determining corresponding interference effects, suggesting that visual attention and oculomotor control might draw on common limited resources. For example, [Tibber, Grant, and Morgan \(2009\)](#) explicitly compared non-spatial vs. spatial perceptual tasks in the context of saccade execution and observed stronger interference for the latter condition, suggesting that spatial processing overlap may increase the level of cross-task conflict. At first sight, this appears to be at variance with our observation that interference is independent of content-based (spatial) processing overlap. Note, however, that the setup in [Tibber et al. \(2009\)](#) involved simultaneous processing of two spatially separated stimuli (saccade target and perceptual target), which might have specifically emphasized spatial processing demands. In contrast, our design only involved the processing of one stimulus for both responses. Thus, we minimized processing demands associated with visuospatial attention to specifically focus on the interplay between output domains.

The finding of mutual interference between saccades and vocal responses could also be of interest to many applied research fields wherever corresponding action demands are present in more complex cognitive tasks. For example, fixation durations in oral reading are known to be longer than for silent reading (e.g., [Rayner, 2009](#)), and talking on the cell phone is known to impair visual orientation during driving (e.g., [Atchley & Dressel, 2004](#); [Atchley, Dressel, Jones, Burson, & Marshall, 2011](#); [Strayer, Drews, & Johnston, 2003](#)). Probably, the interference effects shown in our present study (at least partly) contribute to such effects in more complex environments.

5.5. Conclusion

Modularity has been a central assumption in many theories of the mind within the area of cognitive sciences. However, after a revitalization of the concept of modularity by [Fodor \(1983\)](#) and more recently by proponents of massive modularity ([Carruthers, 2005](#); [Frankenhuis & Ploeger, 2007](#); [Pinker, 2005](#); [Sperber, 2005](#); [Tooby & Cosmides, 1992](#)), many researchers raised serious doubts (on either

theoretical or empirical grounds) regarding the extent to which our mind can be adequately described by assuming strong modularity (e.g., see Barrett & Kurzban, 2006; Buller, 2005; Prinz, 2006). However, the notion of modularity (and its counterpart, the notion of system interdependency) can still be considered a fruitful research-guiding heuristic to assess the fundamental determinants of the extent to which cognitive systems interact. The present research should be considered as a potential starting point to study the issue of encapsulation on the output side of cognitive processing, and in this way represents another step toward an understanding of crucial factors affecting the interplay among output systems in cognition.

Acknowledgments

The present research was funded by the Deutsche Forschungsgemeinschaft (HU 1847/3-1). We thank Magali Kreutzfeldt and Linn-Elena Zulka for the collection of the data, and those who kindly volunteered to participate in the study.

References

- Anderson, J. R., Taatgen, N. A., & Byrne, M. D. (2005). Learning to achieve perfect timesharing: Architectural implications of Hazeltine, Teague, and Ivry (2002). *Journal of Experimental Psychology: Human Perception and Performance*, 33, 749–761.
- Atchley, P., & Dressel, J. (2004). Conversation limits the functional field of view. *Human Factors*, 46, 664–673.
- Atchley, P., Dressel, J., Jones, T. C., Burson, R., & Marshall, D. (2011). Talking and driving: Applications of crossmodal action reveal a special role for spatial language. *Psychological Research Psychologische Forschung*, 75, 525–534.
- Barrett, H. C., Frederick, D. A., Haselton, M. G., & Kurzban, R. (2006). Can manipulations of cognitive load be used to test evolutionary hypotheses? *Journal of Personality and Social Psychology*, 91, 513–518.
- Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, 113, 628–647.
- Bekkering, H., Adam, J. J., Kingma, H., Huson, A., & Whiting, H. T. A. (1994). Reaction time latencies of eye and hand movements in single- and dual-task conditions. *Experimental Brain Research*, 97, 471–476.
- Buller, D. J. (2005). *Adapting minds: Evolutionary psychology and the persistent quest for human nature*. Cambridge, MA: MIT Press.
- Carruthers, P. (2005). The case for massively modular models of mind. In R. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 205–225). Oxford, England: Blackwell.
- Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences*, 3, 115–120.
- D'Avella, A., & Pai, D. K. (2010). Modularity for sensorimotor control: Evidence and a new prediction. *Journal of Motor Behavior*, 42, 361–369.
- DeSteno, D., Bartlett, M., Braverman, J., & Salovey, P. (2002). Sex differences in jealousy: Evolutionary mechanism or artifact of measurement? *Journal of Personality and Social Psychology*, 83, 1103–1116.
- 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.
- Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Frankenhuis, W. E., & Ploeger, A. (2007). Evolutionary psychology versus Fodor: Arguments for and against the massive modularity hypothesis. *Philosophical Psychology*, 20, 687–710.
- Friston, K. J., & Price, C. J. (2011). Modules and brain mapping. *Cognitive Neuropsychology*, 28, 241–250.
- Hazeltine, E., Ruthruff, E., & Remington, R. W. (2006). The role of input and output modality pairings in dual-task performance: Evidence for content-dependent central interference. *Cognitive Psychology*, 52, 291–345.
- Hazeltine, E., & Wifall, T. (2011). Searching working memory for the source of dual-task costs. *Psychological Research Psychologische Forschung*, 75, 466–475.
- Holender, D. (1980). Interference between a vocal and a manual response to the same stimulus. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 421–431). Amsterdam: North-Holland.
- Hommel, B. (1998). Automatic stimulus-response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1368–1384.
- Horn, A. K., & Leigh, R. J. (2011). The anatomy and physiology of the ocular motor system. *Handbook of Clinical Neurology*, 102, 21–69.
- Huestegge, L. (2011). The role of saccades in multitasking. Towards an output-related view on eye movements. *Psychological Research Psychologische Forschung*, 75, 452–465.
- Huestegge, L., & Koch, I. (2009). Crosstalk between simultaneously executed saccades and manual responses. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 352–362.
- Huestegge, L., & Koch, I. (2010a). Crossmodal action selection: Evidence from dual-task compatibility. *Memory and Cognition*, 38, 493–501.
- Huestegge, L., & Koch, I. (2010b). Fixation disengagement enhances peripheral perceptual processing: Evidence for a perceptual gap effect. *Experimental Brain Research*, 201, 631–640.
- Huestegge, L., & Koch, I. (2013). Constraints in task-set control: Modality dominance patterns among effector systems. *Journal of Experimental Psychology: General*, 142, 633–637.
- Hutton, S. B. (2008). Cognitive control of saccadic eye movements. *Brain and Cognition*, 68, 327–340.
- Isa, T., & Yoshida, M. (2009). Saccade control after V1 lesion revisited. *Current Opinion in Neurobiology*, 19, 608–614.
- Jiang, J., & Egner, T. (2013). Using neural pattern classifiers to quantify the modularity of conflict-control mechanisms in the human brain. *Cerebral Cortex*. <http://dx.doi.org/10.1093/cercor/bht029>.

- Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A., et al (2010). Control and interference in task switching – A review. *Psychological Bulletin*, 136, 849–874.
- Koch, I. (2009). The role of crosstalk in dual-task performance: Evidence from manipulating response-set overlap. *Psychological Research Psychologische Forschung*, 73, 417–424.
- Koch, I., & Prinz, W. (2002). Process interference and code overlap in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 192–201.
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis of stimulus–response compatibility – A model and taxonomy. *Psychological Review*, 97, 253–270.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35, 1897–1916.
- Lehle, C., & Hübner, R. (2009). Strategic capacity sharing between two tasks: Evidence from tasks with the same and different task sets. *Psychological Research Psychologische Forschung*, 73, 707–726.
- Lien, M.-C., & Proctor, R. W. (2002). Stimulus–response compatibility and psychological refractory period effects: Implications for response selection. *Psychonomic Bulletin and Review*, 9, 212–238.
- Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, 108, 393–434.
- Logan, G. D., & Schulkind, M. D. (2000). Parallel memory retrieval in dual-task situations: I. Semantic memory. *Journal of Experimental Psychology: Human Perception & Performance*, 26, 1072–1090.
- Lünenburger, L., Kutz, D. F., & Hoffmann, K. P. (2000). Influence of arm movements on saccades in humans. *European Journal of Neuroscience*, 12, 4107–4116.
- MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception processes. *Perception and Psychophysics*, 24, 253–257.
- Magen, H., & Cohen, A. (2010). Modularity beyond perception: Evidence from the PRP paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 395–414.
- Marr, D. (1982). *Vision*. New York: H. Freeman.
- Meyer, D. E., & Kieras, D. W. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, 104, 3–65.
- Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression. *Acta Psychologica*, 115, 123–142.
- Munoz, D. P., Armstrong, I., & Coe, B. (2007). Using eye movements to probe development and dysfunction. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 100–124). Oxford, UK: Elsevier.
- Navon, D., & Miller, J. (1987). Role of outcome conflict in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 435–448.
- Navon, D., & Miller, J. (2002). Queuing or sharing. A critical evaluation of the single-bottleneck notion. *Cognitive Psychology*, 44, 193–251.
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, 116, 220–244.
- Pashler, H., Carrier, M., & Hoffman, J. (1993). Saccadic eye movements and dual-task interference. *Quarterly Journal of Experimental Psychology*, 46A, 51–82.
- Pinker, S. (2005). So how does the mind work? *Mind and Language*, 20, 1–24.
- Prinz, J. J. (2006). Is the mind really modular? In R. J. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 22–36). Malden, MA: Blackwell.
- Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, 22, 341–423.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62A, 1457–1506.
- Simonyan, K., & Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *Neuroscientist*, 17, 197–208.
- Sperber, D. (2005). Modularity and relevance: How can a massively modular mind be flexible and context-sensitive? In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Structure and content* (pp. 53–68). Oxford, England: Oxford University Press.
- Stephan, D. N., & Koch, I. (2010). Central crosstalk in task switching – Evidence from manipulating input-output modality compatibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1075–1081.
- Stephan, D. N., & Koch, I. (2011). The role of input–output modality compatibility in task switching. *Psychological Research Psychologische Forschung*, 75, 491–498.
- Stephan, D. N., Koch, I., Hendler, J., & Huestegge, L. (2013). Task switching, modality compatibility, and the supra-modal function of eye movements. *Experimental Psychology*, 60, 90–99.
- Sternberg, S. (2011). Modular processes in mind and brain. *Cognitive Neuropsychology*, 28, 156–208.
- Strayer, D., Drews, F., & Johnson, W. (2003). Cell-phone induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, 9, 23–32.
- Stelzel, C., & Schubert, T. (2011). Interference effects of stimulus–response modality pairings in dual tasks and their robustness. *Psychological Research Psychologische Forschung*, 75, 476–490.
- Sweeney, J. A., Luna, B., Keedy, S. K., McDowell, J. E., & Clementz, B. A. (2007). fMRI studies of eye movement control: Investigating the interaction of cognitive and sensorimotor brain systems. *NeuroImage*, 36, 54–60.
- Tibber, M. S., Grant, S., & Morgan, M. J. (2009). Oculomotor responses and visuospatial perceptual judgments compete for common limited resources. *Journal of Vision*, 9(12): 21, 1–13.
- Tombu, M., & Jolicoeur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 3–18.
- Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 19–136). Oxford: Oxford University Press.
- Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, 50, 449–455.