Attentional Flexibility is Imbalanced: Asymmetric Cost for Switches between External and Internal Attention

Sam Verschooren[1], Baptist Liefooghe[1], Marcel Brass[2], Gilles Pourtois[1]

[1]Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium;
[2]Department of Experimental Psychology, Ghent University, Ghent, Belgium

Word count: 11.288

Corresponding author
Sam Verschooren
Cognitive & Affective Psychophysiology Laboratory
Department of Experimental Clinical and Health Psychology
Ghent University
Ghent
Belgium
Tel: +32 9 264 86 56
Email address: sam.verschooren@ugent.be

Abstract

Whereas the effects of attention switches occurring within perception or memory are relatively well understood, much less is known about switches of attention between them. We discuss the methodological limitations of initial research on this topic, which was never integrated with the broader cognitive literature. On the basis of this discussion, we present here a new paradigm, in which participants performed a simple probe-to-target matching task where targets were either perceived on screen or retrieved from memory. Across successive trials, repetitions or alternations (in both directions) between these two conditions were created, and eventually compared to each other. In line with our prediction, derived from the assumption of a top-down control mechanism, we found a cost for switching between external and internal attention in Experiment 1. Furthermore, this switch cost was asymmetric, being substantially larger when switching from (external) perception to (internal) memory than the other way around. In Experiments 2-4, we ruled out an imbalance in practice, learning, and preparation as confounds for this asymmetry. We propose that switches of attention between internal and external information are underpinned by a supervisory attention control mechanism, and that this asymmetry can be explained in terms of priming, associative interference or memory retrieval.

Significance statements

1.   This study sheds light on the remarkable cognitive ability that allows to switch attention rapidly between external information available in the environment and internal representations held in memory.

2.   Results show that it is more difficult to switch to internally-stored information stored in memory when beforehand one processes external competing information than the other way around. We discuss this cost asymmetry in the light of different theoretical frameworks compatible with it.

Introduction

Humans maintain a seemingly rich mental representation of their external environment, as is amply demonstrated in research on perception (Simons & Chabris, 1999) and memory (Schacter, Addis, & Buckner, 2007). Whereas perception depends on the immediate presence of sensory input (Mesulam, 1998), mental imagery (see Nanay, 2015), which underlies memory maintenance, has been defined as perceiving "in the absence of the appropriate immediate sensory input" (Kosslyn, Behrmann, & Jeannerod, 1995). Given that cognitive resources are limited (Lennie, 2003; Marois & Ivanoff, 2005), we need to constantly balance them between perceptual and memory processing. Or put differently, resources need to be flexibility shifted from internal stimulus representations in memory (internal attention) to external stimuli in the environment (external attention) or vice versa (Chun, Golomb, & Turk-Browne, 2011). Yet, despite such flexibility being central in many everyday situations, the cognitive architecture underlying resource allocation between internal and external attention remains currently poorly understood (Burgess, Dumontheil, & Gilbert, 2007; Myers, Stokes, & Nobre, 2017).

Weber, Burt, and Noll (Weber, Burt, & Noll, 1986) were the first to empirically investigate this process and used a list-completion procedure (e.g., Jersild, 1927). On each trial, participants were presented with two displays, each containing a series of letters or asterisks (e.g., '***', 'RNH', '*NH'). During the encoding phase, participants were required to memorize the letters on the first display in their correct order for future use. Afterwards, the second display was presented, which also contained strings of letters and asterisks. Participants were asked to 'complete' this second display. Whereas letters of the second display simply needed to be read aloud, asterisks needed to be replaced by the corresponding letter presented in the first (memorized) display. Weber et al. (1986) observed shorter list-completion times (i.e., the time between the onset of the display and the completion of the whole list of items) when the second display only required either reading

(e.g., "P Q G"; external attention list) or recall (e.g., "***"; internal attention list) compared to displays which required the coordination of both modalities (e.g., "*Q*"; mixed list). This performance cost increased as a function of list length (e.g., sequences of 3 vs. 6 items). Based on these findings, Weber et al. (1986) proposed that, in this task, internal and external attention are underlain by distinct processing systems, which roughly coincided with Baddeley's (1986) division of working memory (WM) between the articulatory loop and the visuo-spatial sketchpad, respectively. More precisely, in their proposal, internal information is kept accessible in the articulatory loop and external information is processed through the visuo-spatial sketchpad. When switching between both modalities, central resources (i.e., the central executive) are needed to selectively access information from one of these subsystems. Because this switch is resource demanding, it comes with a performance cost. This hypothesis was supported by a trade-off in resources between the processes involved in switching between modalities (i.e., the central executive) and the processes involved in the maintenance of the items (e.g., articulatory loop and visuo-spatial sketchpad).

The hypothesis that switching between external and internal attention comes with a performance cost was, however, challenged in two ways. First, Dark (1990) criticized the use of an average completion time over both the external and internal lists, as it was shorter for the internal than for the external one (Dark, 1990; Experiment 1), and comparing it to the mixed list led to an overestimated switch cost. Dark (1990; Experiment 2) used lists with five switches (e.g. * P * Q * G) or one switch (* * * P Q G) and calculated the switch cost by taking the difference in completion time between them, divided by four (i.e. the number of remaining switches after subtracting). Performing this analysis separately for lists starting with memory (switch towards perception; external switch cost) and perception (switch towards memory; internal switch cost) items revealed a larger internal than external switch cost. In Experiment 3, she investigated whether this larger

internal switch cost was actually associated with accessing the internal modality (i.e. a switch cost) or rather with selecting the relevant item within this modality (i.e. a memory retrieval cost). In this experiment, she did use Weber et al. (1986)'s approach of comparing unmixed and mixed lists, but had participants report only one pair of items (i.e. a single external or internal switch in the mixed lists versus a modality repetition in unmixed lists). This pair was selected via a cue appearing either 750ms before (pre cue) for some participants or simultaneously with (simultaneous cue) the second display for others. There was no significant switch cost in the pre-cued group (and even a benefit for the external switch), but in the simultaneously cued group she found a larger internal than external switch cost. As the only difference between the pre- and simultaneous cue condition was whether the relevant item could be selected in advance, Dark (1990) suggested that the performance cost observed by Weber et al. (1986) was not induced by a switching mechanism between internal and external attention, but by the mere retrieval of items from memory when an asterisk was presented.

Subsequently, Carlson, Wenger, and Sullivan (1993) registered responses to each item separately during list completion. With this approach, they could investigate changes in response time as function of the serial position of an item within a list. Doing so, they observed a scalloped pattern of reaction times, suggesting that in mixed lists internal and external information was coordinated by first combining both sources into a single representation. This representation was then reproduced overtly by chunking responses to items in pairs of two, with small latencies within these pairs but large ones between them. According to Carlson et al. (1993), performance costs observed in the list-completion procedure when comparing single-modality with mixed lists were best explained in terms of higher-order control processes that organize predictable sequences in chunks of information, which can be easily processed and produced. This conclusion invalidated

the findings by Dark (1990) as well, as it could not be excluded that they merely reflected artifacts inherent to the procedure.

As suggested by Carlson et al. (1993), performance in the list-completion procedure is thus difficult to interpret because, on the one hand, list-completion time comparisons only offer a raw and indirect measure of performance and, on the other hand, the predictable nature of a list invokes an additional number of processing steps (e.g., planning, chunking,…). Such drawbacks have also been documented in the task-switching literature (e.g. Miyake, Emerson, Padilla, & Ahn, 2004), which advocates the use of cued task switching (Meiran, 1996) as a better means to investigate processes underlying cognitive flexibility. In cued task switching (Meiran, 1996; Sudevan & Taylor, 1987), bi-dimensional stimuli are usually used on which two choice-reaction tasks can be performed. For instance, coloured shapes on which a colour or a shape judgment can be applied. On each trial, a task cue (e.g. the word COLOUR or SHAPE) is presented, which indicates the task to perform on that trial. Importantly, the sequence in which the tasks are presented is unpredictable for the participants and they can only rely on the task cue in order to know which task to perform. With this kind of experimental design, two types of transitions between a trial n and a trial n−1 can be discerned: task repetitions (e.g. trial n: shape task; trial n−1: shape task) and task switches (e.g. trial n; shape task; trial n−1: colour task). A cornerstone finding is that participants are slower and more error-prone on task switches compared to repeat trials (see Kiesel et al., 2010; Monsell, 2003; Vandierendonck, Liefooghe, & Verbruggen, 2010 for reviews). The switch cost supposedly reflects the operation of processes related to the coordination of different tasks. It is commonly assumed that the switch cost has different sources: (a) processes related to the reconfiguration or retrieval of task-sets (e.g. Mayr & Kliegl, 2000, 2003; Rogers & Monsell, 1995; Rubinstein, Meyer, & Evans, 2001); (b) facilitating and/or inhibitory priming of the task-set, which was active on the previous trial (e.g. Allport, Styles, & Hsieh, 1994; Sohn & Anderson, 2001); and (c) automatic retrieval of

a task-set on the basis of stimulus–task associations formed through practice (Waszak, Hommel, & Allport, 2003; Wylie & Allport, 2000).

Cued task switching thus provides a fine-grained tool to investigate the processes underlying attention flexibility. Not only is the task order completely unpredictable, but repetition and switch trials can be derived from the same mixed blocks. This offers an additional advantage compared to the list-completion procedure, in which mixed lists are compared to single-modality lists. Whereas mixed lists require both to switch between modalities as well as to keep both modalities accessible, single-modality lists require to access only one modality, and no switching. A difference at the behavioral level between both types of lists is thus potentially contaminated by other factors than attention flexibility alone. In view of these considerations, the present study aimed to reevaluate the hypothesis of Weber et al. (1986) stating that switching between internal and external attention is associated with a switch cost. To this end, we introduced a new procedure for attention switching (see Fig. 1), which was inspired by cued task switching (Meiran, 1996). We used a single category of stimuli for both (internal) memory and (external) perception, and randomly presented participants on each trial with a cue that informed them to use either visual perception or (working) memory to perform a matching task that was held constant for these two modalities. More precisely, on external trials, an arrow pointed towards four figures on the screen and, after an interval, one of them was selected to be matched to a probe appearing at the end of the trial (see Fig. 1C, first and second trial). On an internal trial, the same task was used, but the arrow pointed towards four question marks (see Fig. 1C, third trial). Participants then retrieved the figure for the selected location from four figures that were memorized before the start of the block (Fig. 1A), and matched this figure to the probe. With this procedure, four types of transitions between trial n-1 and trial n can be discerned: perceptual following perceptual ("External-Repeat"), perceptual following memory ("External-Switch"), memory following memory ("Internal-

Repeat"), and memory following perceptual ("Internal-Switch"). Note that these transitions are defined by different modalities (e.g., internal vs. external), but always comprise the same matching task. Accordingly, during these transitions, pure and direction-specific switches of attention can be examined, as opposed to switches of task for example. Furthermore, these transitions are unpredictable and part of the same mixed list requiring a frequent access to both modalities.

Based on the aforementioned procedure, we conducted four experiments. In Experiment 1 we tested whether switching between internal and external attention was associated with a cost. Although the studies presented here above did not reach a consensus on the presence of this switch cost (Carlson, Wenger, and Sullivan, 1993; Dark, 1990; Weber, Burt, & Noll, 1986), we did expect to find one however. This prediction was derived from work by Burgess and colleagues (Burgess et al., 2007; Gilbert, Frith, & Burgess, 2005), who have provided evidence for a top-down mechanism that regulates these switches. The recruitment of such a mechanism on switch trials likely taxes attentional resources and leads to a cost (see "General Discussion" for an extensive discussion of this account; see also Verschooren, Schindler, De Raedt, & Pourtois, 2019). In line with this prediction, the results of Experiment 1 indicated the presence of a switch cost. Moreover, this cost was reliably larger cost on internal than on external trials, hence revealing an unpredicted asymmetry in attention flexibility. In Experiment 2, we tested whether asymmetric switch cost we observed in Experiment 1 was induced by uncontrolled differences in the amount of practice participants encountered in both modalities (see Yeung & Monsell, 2003b, 2003a). Equating the amount of practice, however, did not change the results. In Experiment 3, we controlled for the potential contribution of location-based learning to performance in the external modality. A similar asymmetric switch cost was, however, obtained. Finally, in Experiment 4, we explored the effect of preparation by reducing the interval between the cue indicating the relevant modality and the onset of the prime stimulus (i.e., cue-target interval or CTI). This manipulation did not alter the

pattern of results obtained in the previous experiments and suggesting an asymmetry in the ability to switch between internal and external attention.
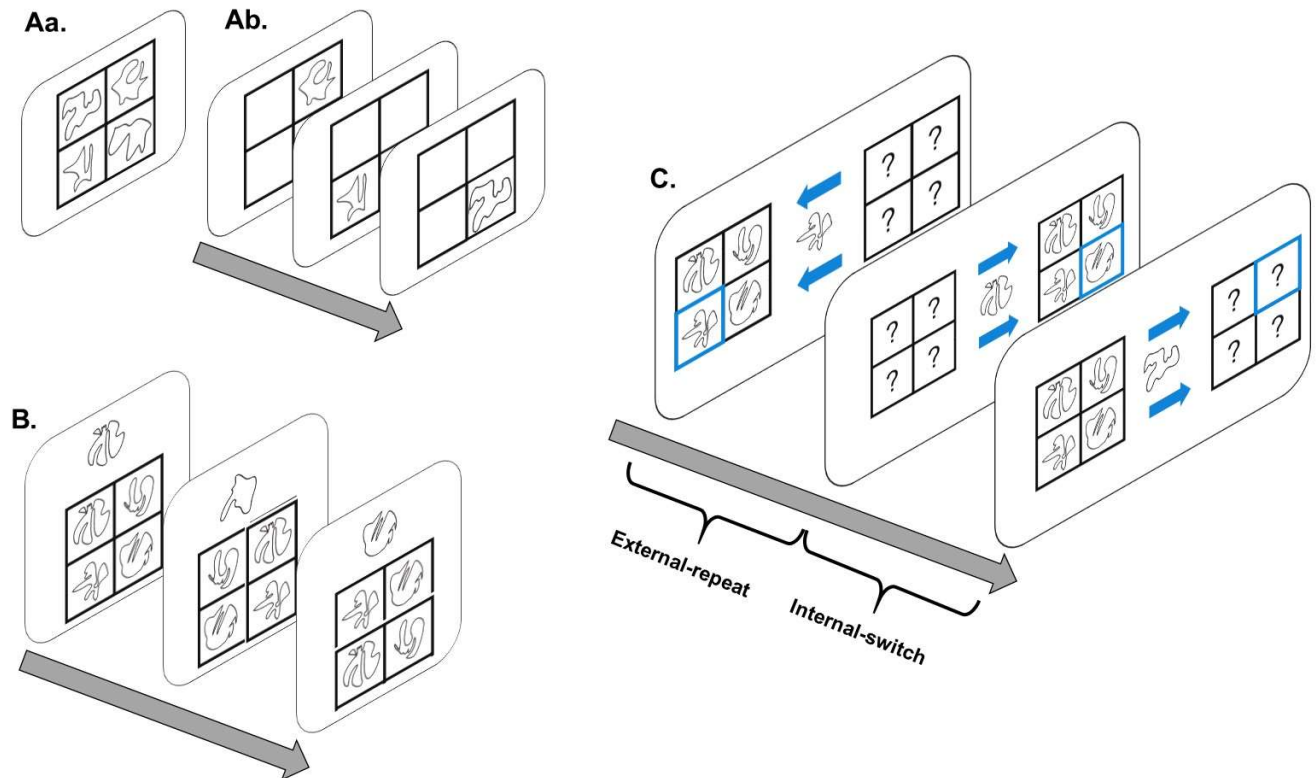
Figure 1. Experimental paradigm. Aa. Before each block, four figures were presented, which were committed to memory. **Ab.** During the internal training phase, participants decided whether the figure was in the correct location, until they reached the accuracy criterion (.8 correct) **B.** During the external training phase (Experiments 2-4), participants decided whether the figure presented above was part of the sample below, until the accuracy criterion was reached. **C.** On each trial, a central probe was presented with arrows pointing towards the external set (first two trials depicted) or internal set (last trial depicted). When the target was selected (blue), participants decided whether it matched the central probe. The conditions of interest were the four (equally frequent) trial-to-trial transitions: External-Repeat, External-Switch, Internal-Repeat, and Internal-Switch.

Experiment 1

Methods

Participants

We recruited 47 participants (18 to 30 years old, 38 women) through Sona Systems, an online recruiting tool at Ghent University. This number was based on an a priori power analysis for a 2 x 2 repeated measures analysis of variance (RM ANOVA) to detect a small effect size (partial η^2 of 0.04) with 80% power, which translated into a minimum sample size of 36 participants.

Materials

The experiment was programmed in the PsychoPy2 library in Python (version 1.85.6; Peirce, 2008). We used non-verbalizable figures from a set of stimuli provided by Endo and colleagues (see Fig. 1A-C; Endo, Saiki, Nakao, & Saito, 2003). These black figures had a size of approximately 70 by 70 pixels and were presented at a visual angle of 3°50' (on a white background). Twelve sets of four stimuli were sampled, of which six were used for one half of the participants and six for the other half. These six sets were used in pairs of two (one set for the External and one for the Internal trials) in the practice block, block one to four, and block five to eight. Moreover, for each participant, we randomized which set was used for External and which for Internal trials. Finally, as all stimuli were unique, the sets never overlapped.

Design

The experiment used a 2x2 within-subjects design. Modality and repetition were the two main independent variables of interest and each trial was coded following their interaction, i.e. External-Repeat, External-Switch, Internal-Repeat, or Internal-Switch. Whereas for Repeat trials the modality on the current trial was the same as on the previous one, they were different for Switch trials. For External-switch, participants switched from memory to visual perception, whereas for

Internal-Switch, they switched in the opposite direction (i.e. from visual perception to memory). The main dependent variables were the reaction times and accuracy scores on each trial.

The experiment consisted of one practice block and eight experimental blocks. Each block consisted of 34 trials. We counterbalanced trial order within each block to have an equal number of trials for each combination of modality (external or internal), switch type (repeat or switch) and probe-target compatibility (match or mismatch), restricted by a maximum of four switches/repeats in a sequence. In addition, immediate repetitions of target position or probe were excluded. On each trial, the (left or right) location for the square filled with question mark and the one filled with figures was either maintained or reversed. We included this factor to avoid that modality switches consistently coincided with spatial attention switches. Over the course of the experiment, for each of the trial combinations (modality x switch x match), there was an equal amount of side switches (location of the squares maintained or reversed) and relevant locations (target in left/right upper/lower compartment). This trial list was subdivided in 11 blocks and the block order was randomized for each participant, but the trial sequence within these blocks was not changed. Different sets of stimuli were used for the practice block, the first four, and the last four experimental blocks.

Procedure

Each trial started with a 300 ms fixation cross, after which a central probe appeared with blue arrows pointing either to the right or left side (Figure 1). A square filled with four question marks was presented on one side and one filled with four figures on the other. For Internal trials, the arrows pointed towards the question marks and for External trials towards the figures. After another 500 ms, the location of the targeted figured lit up in blue. After another 600 ms, the stimuli

disappeared from screen until response (1600 ms response limit). The location of the external and internal stimuli never changed (but see Experiments 3 and 4).

At the beginning of each block, participants first memorized the spatial location of four different stimuli, by means of a separate and short training phase (Figure 1). On each trial, only one figure appeared in either the correct compartment or the opposite (along the diagonal axis) one. Participants had to perform a two-alternative forced choice (2-AFC) task. They had to indicate, by pressing the key "d" or "k" of the keyboard, if the stimulus was presented at the correct or incorrect location, respectively (10s response limit). Only when participants reached the accuracy criterion on this task (80 percent correct), did the actual experimental task start. This ensured that all participants had an equally reliable memory representation for the Internal trials. The training phase was presented at the start of each block, but was more extensive when new stimuli were introduced (i.e. for the practice block and for the first and fifth experimental block) than when the same stimuli were used as in the previous block. In the former case, participants were required to have an accuracy of at least 80 percent after at least 20 correct trials and in the latter case after at least eight correct trials. That is, the experiment continued only when the conjunction of these two conditions, proportion and amount of trials correct, was met.

During the main experimental session, a similar 2-AFC task was used. More specifically, participants decided whether the central probe matched the figure highlighted (blue frame, see Figure 1). Whereas this required a visual matching on External trials, on Internal trials (represented by the question mark), they needed to compare the probe to the figure they had memorized at this specific location beforehand, hence to perform the same matching task but based on an internal representation of the display. Participants responded with the keys "d" and "k" of the keyboard for match and mismatch, respectively.

Data analysis

Data preprocessing, visualization and analysis were carried out in R (R core team, 2018; version 1.1.456; data and analysis scripts: https://osf.io/62u7b/), using the tidyverse (Wickham, 2017), ggplot2 (Wickham, 2016), lme4 (Bates, Mächler, Bolker, & Walker, 2015) and emmeans (Lenth, 2018) packages. As a preprocessing plan for the data, we removed the training trials and the first two trials of each block. Incorrect trials, or correct trials preceded by incorrect responses were removed. In addition, participants who performed around chance level (< .55) in at least one of the four conditions were removed (average accuracy for included participants was around .9). Accuracy data and reaction times (RT) for correct responses were the main dependent variable, which we analyzed using a generalized linear mixed model (GLMM) approach. GLMMs incorporate the advantages of linear mixed models (LMMs), which do not average over sets of data points but instead treat each one as a single observation belonging to a specified level (here: participant). Hence, they retain richer information than other common approaches such as repeated measures analysis of variance (RM ANOVA) and do not violate the implicit assumption that an individual's performance fluctuates over time (Lo & Andrews, 2015; Speelman & McGann, 2013). Simultaneously, the GLMMs transcend the LMMs in an important way, by relaxing their assumptions concerning the normality of the dependent variable and the residuals (Lo & Andrews, 2015). This element is important, as accuracy and RT data are almost never normally distributed. Accuracy data are discrete (i.e. either 0 or 1) rather than continuous at the trial level and are restricted to the 0 to 1 range when looking at the proportion of correct responses. LMMs often violate this restriction (i.e. a confidence interval (CI) larger than 1) when accuracy is high (Dixon, 2008). Dixon (2008) demonstrated that for analyzing accuracy data, it is therefore desirable to use a logistic regression with binomial distribution in a GLMM. In this model, the log odds of the

proportion of correct trials, which is not restricted to a 0-1 interval, is used to linearly model the effect of the experimental factors.

Similarly, for the RT data, the assumption of normality is rarely met (Lo & Andrews, 2015; Speelman & McGann, 2013). Even though this can be circumvented by transforming the data to fit the mathematical requirements of LMMs, the consensus view in the statistics literature is against this, as the mean of the transformed data is not the same as the transformation of the mean (and violates the assumption behind mental chronometry, see Lo & Andrews, 2015; Speelman & McGann, 2013)). Here again, a powerful solution is offered by GLMMs, which allow researchers to both use the most appropriate dependent variable (i.e. the reaction times themselves instead of their natural log function) and meet the mathematical assumptions of the model (i.e. normality). More precisely, using a link function, we are actually allowed to transform the regression coefficients of the model back to the response scale, which is part of the reason why GLMMs are deemed useful and powerful. We selected the most appropriate distribution and link function by looking at the Akaike and Bayesian information criterions (AIC/BIC), which can be used as a description of how closely a model fits the data while punishing it for its complexity (see Burnham & Anderson, 2004). Based on this analysis (see Supplementary Materials), we selected a GLMM with an inverse Gaussian distribution and natural log link function for all experiments.

As a next step, we tested the hypothesis that a cost was present for the switches compared to repetitions, and whether it could be different for External and Internal trial types. To do so, we carried out a model comparison for both accuracy and RT data, in which we incrementally compared the fit of a model with a random varying intercept only, or with fixed main effects for Switch and for Modality, and/or their interaction effect. These models all had a random intercept for each participant. The p-values for these comparisons were obtained by a likelihood ratio test of the model with the effect under scrutiny being compared against the model without it (e.g. to

investigate whether a different cost was associated with each modality, the model with the interaction effect was compared to the one with the main effects only, and a significant p-value for this comparison therefore signified the presence of an interaction effect).

For the best fitting model, we provided the estimates for its coefficients, its 95% CI, and the associated p-value, which were calculated by the asymptotic Wald test (for which no degrees of freedom are provided). We transformed the estimates for the planned contrasts (External-Switch > External-Repeat and Internal-Switch > Internal-Repeat for the interaction effect) back from the log-scale to the response scale, for which the 95% CI and p-values were obtained by the asymptotic Wald tests as well, and which were corrected for multiple comparisons with the Tukey method. Assumptions were visually inspected with the residual plots, which did not reveal any obvious deviations from homoscedasticity or normality for the residuals of the GLMM (but these deviations were present for the LMM, see Supplementary Materials).

Finally, as it has been demonstrated previously that experimental factors can differentially affect the skewed tail of the RT distribution, we carried out a shift function analysis (Rousselet, Pernet, & Wilcox, 2017; see Supplementary Materials).

Results
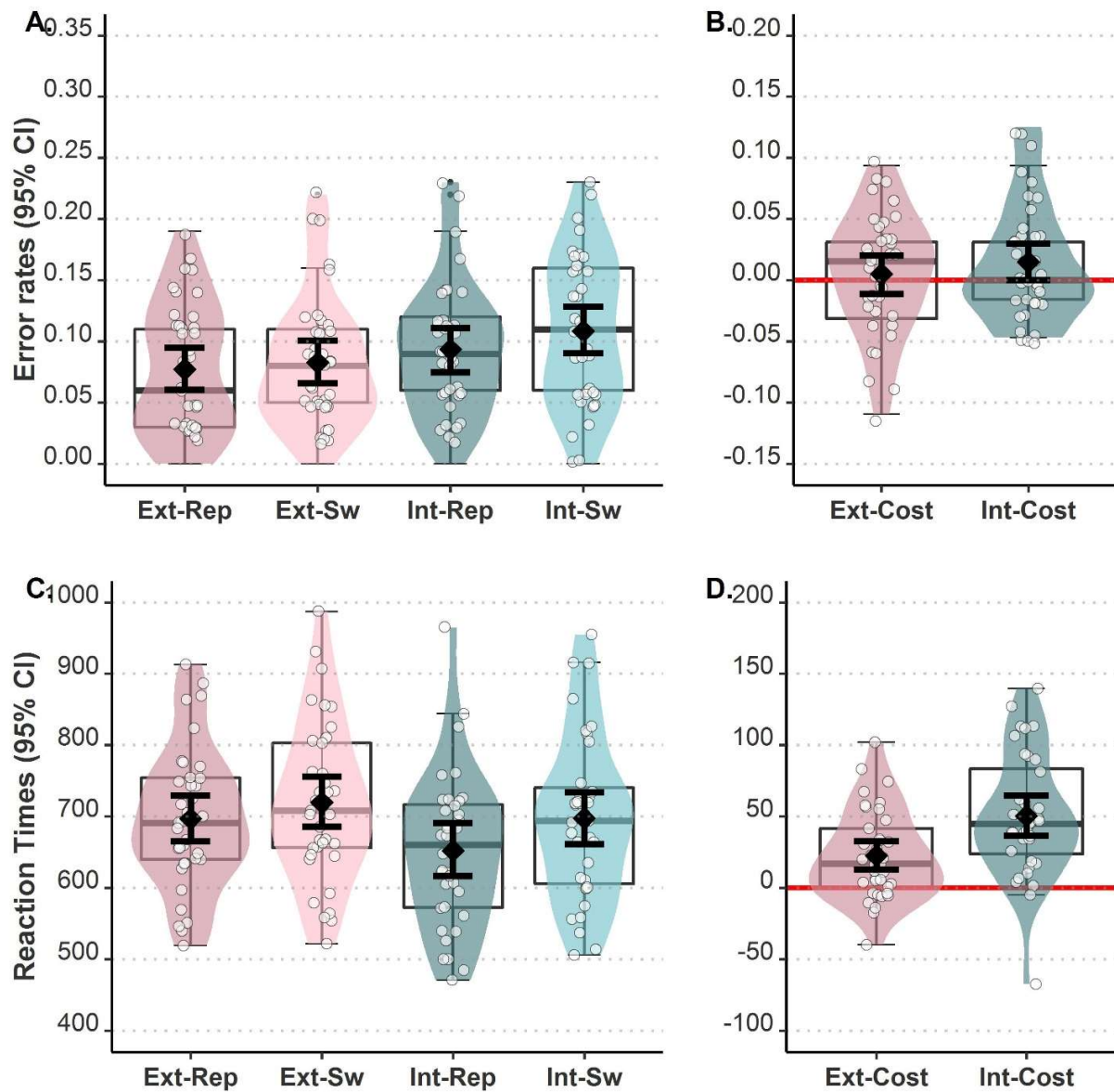
Figure 2. Results Experiment 1. In black: Mean sample value and 95% confidence interval (CI) in black. In white: Individual participants' mean. **A)** Error rates for External-Repeat, External-Switch, Internal-Repeat and Internal-Switch conditions. **B)** Error rate difference scores for external (External Cost: External-Switch > External-Repeat) and internal (Internal Cost: Internal-Switch > Internal-Repeat) switches. **C)** Reaction times for the conditions. **D)** Reaction times for the difference scores.

Ten participants[1] were excluded from the analysis due to low accuracy (see "Data Analysis" section). For the 37 included participants, accuracy was overall very high (see Fig. 2A). The model comparison for the accuracy data revealed that a main effect for Modality best explained the data ($\chi^2_3 = 10.99$, p < 0.001; see table 2). The model's coefficient for Modality, transformed to an odds ratio, was estimated at 1.31 (CI = [1.14-1.51], Z = -3.73, p < 0.001), meaning that it was 1.31 times more likely to answer correctly on External than on Internal trials. Planned contrast further estimated the error rates at 0.07 (CI = [0.06-0.09]) for External and at 0.09 (CI = [0.08-0.11]) for Internal trials. The model comparison for the RT data (see fig. 2C) showed that the model including the interaction effect fit the data best, better than the model with the main effects only ($\chi^2_6 = 20.66$, p < 0.001; see table 2). In this model, Switch Type ($\beta = 0.03$, 95% CI = [0.01;0.04], T = 3.60, p < 0.001), Modality ($\beta = -0.07$, 95% CI = [-0.08;-0.05], T = -8.51, p < 0.001), and their interaction ($\beta = 0.05$, 95% CI = [0.03;0.07], T = 4.55, p < 0.001) all predicted the participants' RTs. Estimates of the planned contrasts for this interaction effect (see Fig. 2D), transformed back to the original response scale in milliseconds, revealed a steeper slope from Repeat to Switch in the Internal modality (see Fig. 2D; $\beta = 54.6$, 95% CI = [41.0;67.9], Z = 10.05, p < 0.001), compared to the External one (see Fig. 2D; $\beta = 20.4$, 95% CI = [5.9;34.6], Z = 3.60, p = 0.002).

Table 1

Model comparison for accuracy

	Df	AIC	BIC	χ^2	P-value
Intercept only	2	5652.71	5667.02	-	-
Switch Type	3	5651.83	5673.30	2.88	0.09
Modality	3	5640.84	5662.31	10.99	< 0.001
Both	4	5639.95	5668.57	2.89	0.09

[1] This relatively large number of excluded participants was likely due to a misinterpretation of the instructions. Accordingly, after adjusting the instructions in Experiments 2-4, the number of participants removed due to a poor accuracy substantially dropped.

Table 1

Model comparison for accuracy

	Df	AIC	BIC	χ^2	P-value
Interaction	5	5641.52	5677.30	0.43	0.51

Model comparison for reaction times

	Df	AIC	BIC	χ^2	P-value
Intercept only	3	102812.8	102833.7	-	-
Switch Type	4	102727.4	102755.3	87.46	< 0.001
Modality	4	102765.0	102792.9	0.00	1.00
Both	5	102674.5	102709.3	92.51	< 0.001
Interaction	6	102655.8	102697.6	20.66	< 0.001

Discussion

In Experiment 1 we investigated whether switching between internal and external attention could lead to a cost in performance. Accuracy performance was overall very high, but participants were somewhat worse on internal than external trials (approximately 2%). Two main results emerged from the analysis of the RTs. First, attention switching was associated with a substantial processing cost, compared to the condition where no switch occurred across successive trials. This result converges with previous findings based on a different paradigm (Verschooren, Schindler, De Raedt, & Pourtois, 2019), as well as older results available in the literature (Carlson et al., 1993; Dark, 1990; Weber et al., 1986). Second, the switch cost was larger when switching from external towards internal attention (approximately 55 ms), than the reverse (approximately 20 ms).

The results of Experiment 1 suggest that switching from external to internal attention is associated with a larger cost than switching from internal to external attention. However, before endorsing this conclusion, an alternative account needs to be considered. Research on task switching has demonstrated that an imbalance in practice between two tasks can induce an asymmetric switch cost. More precisely, Yeung and Monsell found a larger switch cost for more recently practiced tasks (Yeung & Monsell, 2003b), as well as for tasks that received a substantial

amount of pre-experimental practice (Yeung & Monsell, 2003a). In Experiment 1, participants received training before the start of the experiment on the internal task (the memory training). Thus, our findings could potentially be explained by both the amount and recency of the internal practice, rather than an asymmetry in attention flexibility between internal and external attention.

To address this concern, we ran a second experiment that was identical to Experiment 1, with the notable exception that we also added a separate training phase for external trials to familiarize the participants with the figures used in the two-alternative forced choice task.[2] Hence, participants were familiarized with both modalities separately before they actually started the test phase where both modalities were combined.[3] To account for the potential effects of practice recency, we counterbalanced the order of the external and internal practice across participants.

---

[2] Even though we attempted to equate the external and internal training phase as much as possible, some differences necessarily remained. More specifically, whereas the goal of the internal training was to encode the stimulus-location associations, we actually needed to avoid this component in the external training. That is, encoding these associations in the external training phase would confound the key difference between these conditions in the experimental task. Accordingly, the external training corresponded more to a familiarization for the participants with the stimuli rather than a real training of external attention.

[3] A detailed analysis of the results of Experiment 1 in which we modeled the effect of time on performance (see Supplementary Materials), suggested that changing the stimulus set used for the working-memory load after four blocks (i.e. in the middle of the test phase) influenced performance on external trials more than on internal trials (see supplementary Fig. 2). Presumably, this was due to an imbalance in training as well, given that participants were familiarized with the new internal stimuli before the start of the experimental task, but not with the new external stimuli. Since this unexpected difference could artificially enlarge the difference between external and internal trials (and thus inflate the asymmetric switch cost seen between them in Experiment 1), we used, for each participant thoughout all experimental blocks, a unique set of external stimuli and a unique set of internal stimuli. The central question of Experiment 2 was whether we could replicate the asymmetric switch cost observed in Experiment 1, when practice with each of the two modalities was carefully controlled and balanced; and the same set of four figures was used throughout.

Experiment 2

Methods

Participants

We recruited 32 participants (18 to 30 years old, 25 women) through Sona Systems. This number was based on a power analysis run to find the interaction effect from Experiment 1 with 90% power, which revealed that 30 participants had to be included to achieve this.

Procedure

The task and stimuli were mostly identical to the ones used in Experiment 1. However, we added an external training phase (see Figure 1B) before the start of the experiment. In each trial, a square appeared with the four figures from the external set. The location of these figures was randomized on each trial. Above the square, either a figure from this set or a different one was presented. Participants performed a 2-AFC task, responding with the "d"-key if this figure was part of the set and the "k"-key if it was not. The same learning criterion as the one used for the internal one was used, i.e. 20 and 8 correct trials (for first encounter and repetition, respectively) and > .8 accuracy. The order of the external and internal training phase was counterbalanced across participants. In addition, we used only one set of four figures for internal trials (see footnote 2) and another set of four for the external trials throughout the entire experiment, except for the ones used in the practice block. Finally, we added some minor changes to optimize further the paradigm. More specifically, we used ten blocks instead of eight. We improved the randomization by avoiding probe repetitions across successive trials. Last, when the probe did not match the figure at the cued-location, it matched with an equal probability any of the three other figures (i.e. non-cued locations).

Results
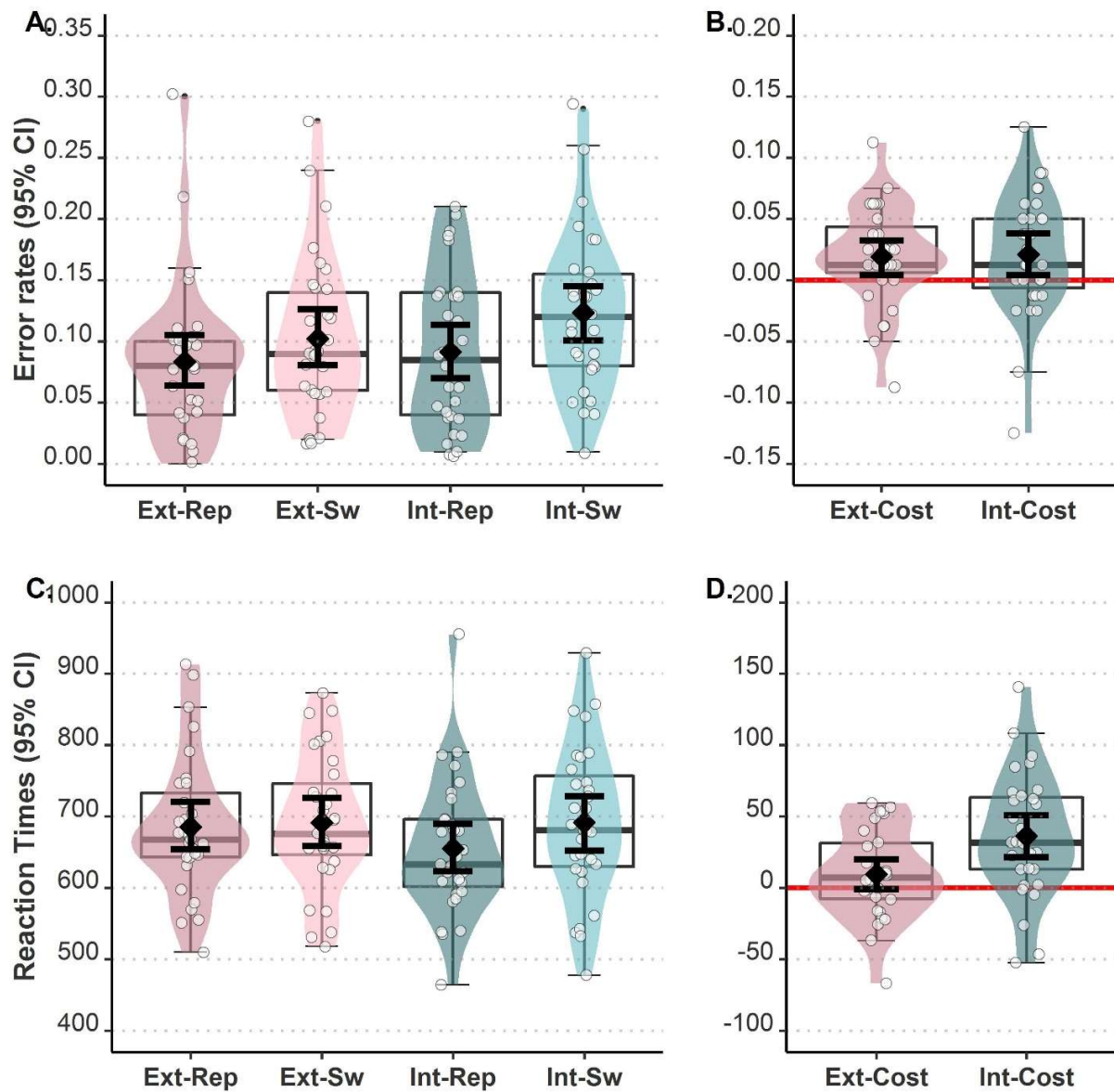
Figure 3. Results Experiment 2. In black: Mean sample value and 95% confidence interval (CI) in black. In white: Individual participants' mean. **A)** Error rates for External-Repeat, External-Switch, Internal-Repeat and Internal-Switch conditions. **B)** Error rate difference scores for external (External Cost: External-Switch > External-Repeat) and internal (Internal Cost: Internal-Switch > Internal-Repeat) switches. **C)** Reaction times for the conditions. **D)** Reaction times for the difference scores.

One participant was excluded from the analysis (see section "Data Analysis"). Accuracy was very high, overall (see Fig. 3A). The model comparison showed that the model with main effects for Modality and Switch Type fit the data best (χ^2_4 = 11.39, p < 0.001; see table 3). In this model, the coefficients, transformed to the odds ratio, were estimated at 1.24 (CI = [1.09;1.41], Z = -3.18, p = 0.001) and 1.26 (CI = [1.09;1.41], Z = -3.38, p < 0.001) for Modality and Switch Type, respectively. Planned contrasts for Modality estimated the error rate for External trials at 0.07 (CI = [0.06;0.10]) and at 0.10 (CI = [0.08;0.12]) for Internal ones. For Switch Type, planned contrasts estimated error rate at 0.08 (CI = [0.06;0.10]) for Repeat and 0.10 (CI = [0.08;0.12]) for Switch trials. The model comparison run on the RT data (see Figure 3B) showed that the model including the interaction effect fit the data best (χ^2_6 = 11.69, p < 0.001 when compared to the model with main effects only; see table 3). RTs were predicted by Modality (β = -0.05, 95% CI = [-0.07; -0.03], T = -5.80, p < 0.001) and importantly, the interaction with Switch Type (β = 0.04, 95% CI = [0.02;0.07], T = 3.42, p < 0.001), but not by Switch Type alone (β = 0.01, 95% CI = [-0.01;0.03], T = 1.20, p = 0.23). Planned contrasts, transformed to the response scale, showed that whereas the slope from Repeat to Switch for External trials was not significant (see Fig. 3D; β = 7.2, 95% CI = [-8.3;22.2], Z = 1.20, p = 0.629), for the Internal trials it was highly significant (see Fig. 3D; β = 34.9, 95% CI = [20.3; 49.1], Z = 6.064, p < 0.001).

Table 2

Model comparison for accuracy

	Df	AIC	BIC	χ^2	P-value
Intercept only	2	6309.30	6323.71	-	-
Switch Type	3	6299.93	6321.53	11.38	0.001
Modality	3	6301.25	6322.86	0.00	1.00
Both	4	6291.86	6320.67	11.39	0.001
Interaction	5	6293.84	6329.85	0.02	0.89

Table 2

Model comparison for accuracy

	Df	AIC	BIC	χ^2	P-value
Model comparison for reaction times					
Intercept only	3	106568.5	106589.5	-	-
Switch Type	4	106545.5	106573.5	24.99	< 0.001
Modality	4	106549.3	106577.2	0.00	1.00
Both	5	106524.8	106559.8	26.42	< 0.001
Interaction	6	106515.1	106557.1	11.69	0.001

Discussion

Results of Experiment 2 replicated those of Experiment 1. Error rates were low overall, but with increased for internal and switch trials (approximately by 2%). An asymmetric switch cost for the reaction times was again observed: the switch cost was larger for external-internal transitions (approx. 35 ms) compared to internal-external transitions (approx. 7 ms). Importantly, participants received a roughly equivalent amount of practice for both modalities (cfr. Yeung & Monsell, 2003b, 2003a).

Although results of Experiment 2 clearly suggest that mostly switching away from the external and towards the internal modality was associated with a substantial performance cost, the current design does not ensure that the external and internal condition were completely distinct. More precisely, a potential concern of Experiment 2 (and Experiment 1 as well) is that location-based learning took place during the course of the experiment for the external stimuli, facilitating their processing, and complicating in turn a direct comparison with the internal stimuli. This is possible because the four external stimuli were presented at the same location throughout the experiment. As a result, participants may have memorized the stimulus-location associations, and – over the course of the experiment – used an internal representation to perform the task in the external modality. Accordingly, performance on the external modality may have been

contaminated by the use of internal representations. In response to this concern, we randomized the stimulus-location mappings of the external modality on a trial-by-trial basis in Experiment 3. There, we tested the hypothesis of an asymmetric switch cost between internal and external attention, when the putative implicit learning of external stimuli was cancelled.

Experiment 3

**Methods**

Participants

32 participants (18 to 30 years old, 27 women) were recruited through Sona Systems, based on the same power analysis used for Experiment 2.

Procedure

The task and stimuli were mostly identical to the ones used in Experiment 2. Compared to this previous experiment, the most important adjustment was that across successive trials, the four figures used for external trials were presented at different locations each time, thereby blocking any implicit learning of the associations between location and identity. In addition, as the number of trials required in the external and internal training phase seemed to point towards ceiling effects, we decreased the amount of correct trials required to start the main task (i.e. from 20 to 14). That is, when participants had at least 14 trials correct in total in conjunction with an overall accuracy of at least 80 percent, the training phase ended and the experimental phase began.

Results
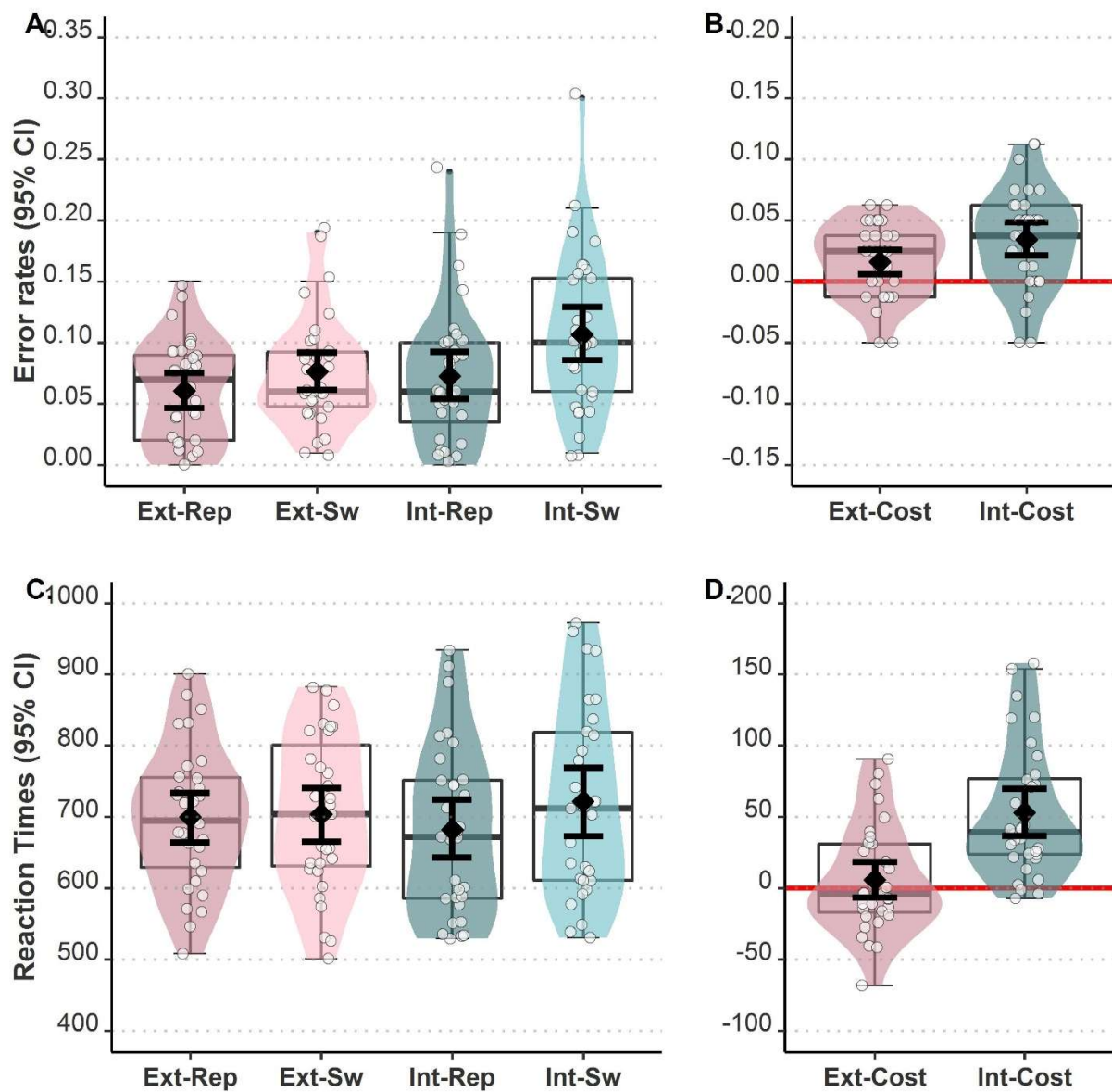
Figure 4. Results Experiment 3. In black: Mean sample value and 95% confidence interval (CI) in black. In white: Individual participants' mean. **A)** Error rates for External-Repeat, External-Switch, Internal-Repeat and Internal-Switch conditions. **B)** Error rate difference scores for external (External Cost: External-Switch > External-Repeat) and internal (Internal Cost: Internal-Switch > Internal-Repeat) switches. **C)** Reaction times for the conditions. **D)** Reaction times for the difference scores.

No participants were excluded from the analysis (see "Data Analysis" section). Accuracy was very high, overall (see Fig. 4A). When comparing the models for the accuracy data, the one with the main effects for Modality and Switch Type fit them best (χ^2_4= 23.03, p < 0.001; see table 4), with odds ratios estimated at 1.36 (CI = [1.18;1.58], Z = 4.134, p < 0.001) and 1.43 (CI = [1.24;1.66], Z = 4.80, p < 0.001) for Modality and Switch Type, respectively. Planned contrasts for Modality estimated the error rate for External trials at 0.06 (CI = [0.05;0.07]) and for Internal ones at 0.08 (CI = [0.06;.0.10]). For Switch Type, the planned contrasts estimated the error rate for Repeat trials at 0.06 (CI = [0.04-0.07]) and at 0.08 (CI = [0.06;0.10]) for Switch trials. The model comparison for the RT data (see Fig. 3B) revealed that the model including the interaction effect fit the data best (χ^2_6 = 28.096, p < 0.001 when compared to the model with main effects only; see table 4). RTs were predicted by the effect of Modality (β = -0.03, 95% CI = [-0.05; 0.02], T = -4.27, p < 0.001) and the interaction between the two main factors (see Fig. 4D; β = 0.06, 95% CI = [0.04;0.09], T = 5.30, p < 0.001), but not by the effect of Switch Type alone (β = 0.01, 95% CI = [-0.01;0.02], T = 0.56, p = 0.575). The planned contrasts for the interaction, transformed back to the response scale, showed a practically flat slope for the External trials (see Fig. 4D; β = 3.4, 95% CI = [-12.2;18.6], Z = 0.561, p = 0.94) but a very steep one for the Internal ones (see Fig. 4D; β = 48.6, 95% CI = [33.4;63.5], Z = 8.01, p < 0.001).

Table 3

Model comparison for accuracy

	Df	AIC	BIC	χ^2	P-value
Intercept only	2	5495.27	5509.74	-	-
Switch Type	3	5474.29	5495.99	22.98	< 0.001
Modality	3	5480.27	5501.97	0.00	1.00

Table 3

Model comparison for accuracy

	Df	AIC	BIC	χ^2	P-value
Both	4	5459.25	5488.18	23.03	< 0.001
Interaction	5	5459.86	5496.03	1.39	0.24
Model comparison for reaction times					
Intercept only	3	116517.8	116539.0	-	-
Switch Type	4	116483.8	116512.1	35.96	< 0.001
Modality	4	116519.4	116547.7	0.00	1.00
Both	5	116485.2	116520.6	36.24	< 0.001
Interaction	6	116459.1	116501.6	28.10	< 0.001

Discussion

In Experiment 3, we ruled out the possible confound of location-based learning for external stimuli by randomizing the stimulus-location mappings on a trial-by-trial basis. As in Experiment 2, we also ensured that practice was equated for both modalities. Notwithstanding this manipulation, results of Experiment 3 clearly showed a similar asymmetric switch cost as in Experiment 1 and 2: approximately 49 ms and 3 ms for the internal-external and external-internal transitions, respectively. For the accuracy data, we observed again small decreases for internal and switch trials (approximately 2%). These results show that participants did not exploit the opportunity to internalize the external stimuli in Experiment 1 and 2, but rather outsourced them to the external environment. In other words, in those experiments, as well as in Experiment 3, the switch cost asymmetry was due to intrinsic differences between external and internal attending.

The results of Experiments 1-3 thus suggest that switching away from internal and towards external attention did not come with a substantial cost. In a last experiment, we aimed to test whether this conclusion could hold when the difficulty of the matching task was increased by lowering the time given to the participants to scan and encode the stimulus-location mappings in

advance. Up to this point (Experiments 1-3), a 500ms cue-target interval (CTI) was used, which may be thought relatively long in view of the simple matching task we used. It is reasonable to assume that, for the external modality, this interval was sufficient for participants to scan and encode the four figures. In this context, participants may have been optimally prepared when the probe was presented on screen, even if the previous trial was internal. Optimal preparation is known to reduce (e.g. Meiran, Chorev, & Sapir, 2000) or even eliminate (Verbruggen, Liefooghe, Vandierendonck, & Demanet, 2007) the switch cost. In Experiment 4, we therefore used a much shorter CTI (i.e. 150 ms), thereby strongly reducing preparation. The question was whether in these conditions, the consistent pattern of results observed in Experiments 1-3 could be replicated.
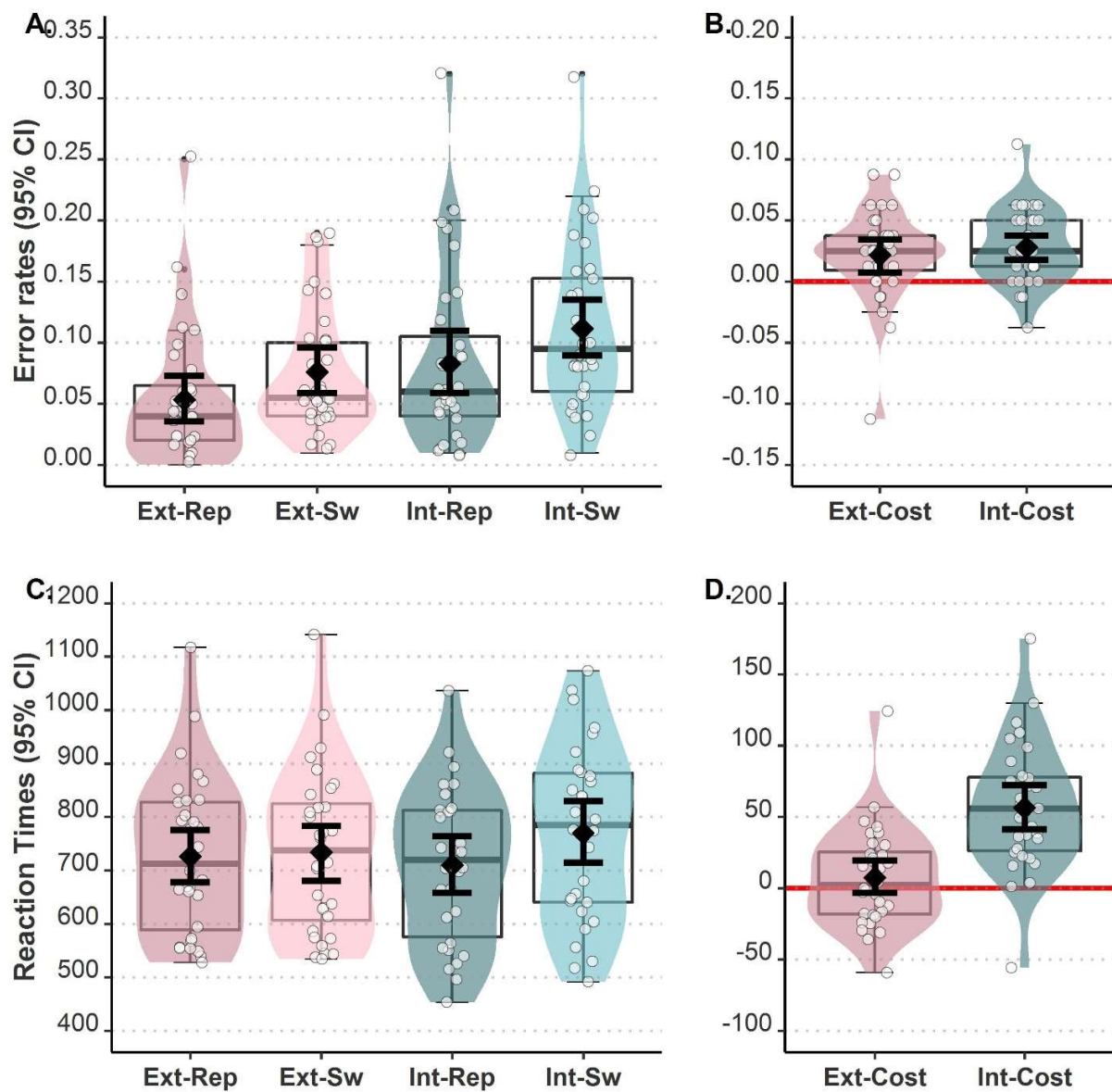
Experiment 4

Methods

Participants

37 participants (18 to 30 years old, 29 women) were recruited through Sona Systems. We initially recruited 32 participants, based on the same power analysis used for Experiment 2, but included an additional five, as more participants reached the exclusion criterion (see "Data Analysis" section).

Procedure

In Experiment 4, we adjusted the duration of the fixation cross (300 ms to 150 ms) and CTI (500 ms to 150 ms). As a result, the stimuli remained on screen longer (600 ms to 850 ms) and the response deadline (from the moment the target was presented) was increased from 1600 to 2000ms. The total trial duration and other task parameters were identical to Experiment 3.

Results

Figure 5. Results Experiment 4. In black: Mean sample value and 95% confidence interval (CI) in black. In white: Individual participants' mean. **A)** Error rates for External-Repeat, External-Switch, Internal-Repeat and Internal-Switch conditions. **B)** Error rate difference scores for external (External Cost: External-Switch > External-Repeat) and internal (Internal Cost: Internal-Switch > Internal-Repeat) switches. **C)** Reaction times for the conditions. **D)** Reaction times for the difference scores.

Five participants were excluded from the analysis (see "Data Analysis" section). Accuracy for the 32 included participants was very high, overall (see Fig. 5A). A model comparison for accuracy data showed that the model with the main effect for Modality and Switch Type fit the data best ($\chi^2_4 = 22.09$, $p < 0.001$; see table 5), with odds ratios for Modality estimated at 1.58 (95% CI = [1.36;1.83], Z = 6.087, $p < 0.001$) and for Switch Type at 1.42 (95% CI = [1.23;1.64], Z = 4.701, $p < 0.001$). Planned contrasts for Modality estimated the error rate for External trials at 0.05 (95% CI = [0.04;0.07]) and for Internal trials at 0.08 (95% CI = [0.07;0.10]). For Switch Type, the planned contrasts estimated the error rate for Repeat trials at 0.06 (95% CI = [0.05;0.08]) and at 0.08 [0.06;0.10] for Switch trials. The best fitting model to the RT data (see Fig. 5B) was the one including the interaction term, better than the one with the main effects only ($\chi^2_6 = 28.37$, $p < 0.001$; see table 5). In this model, the data were not predicted by the main effects of Switch Type ($\beta = 0.001$, 95% CI = [-0.02;.02], T = 0.13, $p = 0.90$) or Modality alone ($\beta = 0.01$, 95% CI = [-0.01;0.03], T = 1.01, $p = 0.31$), but solely by their interaction effect (see Fig. 6D; $\beta = 0.07$, 95% CI = [0.04;0.09], T = 5.33, $p < 0.001$). Planned contrasts for the interaction effect showed a practically flat slope for the External trials (see Fig. 5D; $\beta = 6.6$, 95% CI = [-10.3;23.1], Z = 1.01, $p = 0.74$) but a very steep one for the Internal ones (see Fig. 5D; $\beta = 58.4$, 95% CI = [41.0;75.3], Z = 8.42, $p < 0.001$).

Table 4

Model comparison for accuracy

	Df	AIC	BIC	χ^2	P-value
Intercept only	2	5524.31	5538.76	-	-
Switch Type	3	5504.32	5526.00	21.99	< 0.001
Modality	3	5489.09	5510.78	15.22	< 0.001
Both	4	5469.00	5497.91	22.09	< 0.001
Interaction	5	5470.95	5507.09	0.05	0.83

Table 4

Model comparison for accuracy

	Df	AIC	BIC	χ^2	P-value
Model comparison for reaction times					
Intercept only	3	116735.6	116756.8	-	-
Switch Type	4	116691.5	116719.8	46.04	< 0.001
Modality	4	116706.4	116734.6	0.00	1.00
Both	5	116665.1	116700.4	43.32	< 0.001
Interaction	6	116638.7	116681.1	28.37	< 0.001

Discussion

Results of Experiment 4 replicated the findings of Experiments 1-3: switching away from internal towards external attention only induced a minimal performance cost (7 ms), whereas switching away from external towards internal attention induced a substantial performance cost (58 ms). In other words, reducing preparation for the external switches did not alter the pattern of results. At the accuracy level, there was a decrease for internal and switch trials (3% and 2%, respectively).

Table 5

Exp.	Mod.	RT (SD)	ER (SD)	Switch	RT (SD)	ER (SD)	RT Cost	ER Cost
1	Ext	706 (193)	.08 (.27)	Rep	696 (185)	.08 (.27)	21	.00
				Sw	717 (200)	.08 (.27)		
	Int	684 (246)	0.10 (.30)	Rep	655 (230)	.09 (.29)	59	.02
				Sw	714 (259)	.11 (.31)		
2	Ext	690 (200)	.09 (.29)	Rep	687 (203)	.08 (.28)	7	.02
				Sw	694 (198)	.10 (.30)		
	Int	671 (254)	.11 (.32)	Rep	651 (243)	.10 (.30)	40	.02
				Sw	691 (263)	.12 (.33)		
3	Ext	716 (219)	.07 (.25)	Rep	713 (213)	.06 (.24)	6	.02
				Sw	719 (226)	.08 (.27)		
	Int	718 (284)	.09 (.29)	Rep	690 (269)	.07 (.26)	57	.04
				Sw	747 (296)	.11 (.31)		
4	Ext	730 (251)	.07 (.25)	Rep	728 (731)	.05 (.23)	3	.03
				Sw	731 (252)	.08 (.26)		
	Int	767 (340)	.10 (.30)	Rep	735 (323)	.08 (.28)	63	.03
				Sw	798 (353)	.11 (.31)		

Summary statistics for Exp. 1-4. For each experiment, the mean reaction times (RT) and error rates (ER) and their standard deviations (SD) are provided, for Modality and for its interaction with Switch separately. The two final colums display the RT and ER cost for the External and Internal conditions.

General Discussion

In this series of experiments, we investigated whether switches of attention between (external) perception and (internal) memory led to a cost, and whether it was comparable for the two directions. To improve some methodological shortcomings of previous studies available in the literature on this topic, we designed a novel procedure in which participants performed a simple probe-to-target matching task with externally presented (external trials) or internally represented (internal trials) targets. On each trial, a cue indicated the relevant modality. In line with our prediction, across four experiments (see **Table 5**), we consistently found evidence for a switch cost. This cost was asymmetric, however, with a small cost for switches from internal to external attention and a substantial switch cost was observed for switches from external to internal attention. Additional control for stimulus familiarity (Experiment 2), learning of stimulus-location mappings (Experiment 3) or extensive preparation (Experiment 4) did not alter this pattern of results. Taken together, we can thus conclude that, in these experiments, switching is more efficient when moving from internal to external attention than the other way around. Here after, we first present in more detail a (neuro)cognitive mechanism that can account for the presence of a cost when switching between external and internal attending (Burgess et al., 2007; Gilbert, Frith, & Burgess, 2005; see also Verschooren et al., 2019). However, even though this mechanism could in principle explain why we found it, it needs to be complemented with an account for the unexpected asymmetry we observed. To address this, we further discuss different theoretical frameworks available in the literature that can accommodate this switch cost, as well as tentatively explain its asymmetry, including priming, associative interference, and retrieval.

Cost for switches between external and internal information

In all four experiments, participants were slower on switch than repeat trials. This main effect was, however, qualified by the interaction with modality – a substantially larger cost for switches towards internal trials. Given the small magnitude of the external switch cost in the experiments, we carried out a more sensitive analysis by combining the data of all four experiments (see Supplementary Materials). This analysis did provide evidence for a cost in both modalities, albeit reduced for external switch cost (9 ms; 95% CI = [2;16]) compared to internal switch cost (49 ms; 95% CI = [41;56]).

A general mechanism for accessing different subroutines in external and internal processing was proposed from a neurocognitive perspective (Burgess et al., 2007; Gilbert et al., 2005). Burgess and colleagues (2007) noted that behavior and thought can be initiated through both (external) stimulus-oriented and (internal) stimulus-independent processing, which are in constant competition (for recent reviews on resource sharing between external and internal attention, see Chun & Johnson, 2011; Kiyonaga & Egner, 2013; Verschooren et al., 2019). To investigate how this competition is channeled at the neural level, these authors (Gilbert, Frith, & Burgess, 2005) used a set of different tasks that alternated between an external and internal phase, in which participants responded on the basis of visually presented versus self-generated information, respectively (e.g. visually versus mentally tracking a clock hand or interleaved letters of the alphabet). They found that participants were slower on the internal trials and on trials following a switch. In that study, participants also showed a somewhat larger external than internal switch cost, unlike what we have found here. Their design was, however, not optimized for detecting a (behavioral) switch cost and there were large differences in costs between the different tasks. Based on fMRI data collected with this paradigm, they postulated the existence of a top-down gating mechanism: the supervisory attentional gateway (SAG; Burgess et al., 2007), amplifying one

attention modality over the other. Even though Burgess and colleagues (2007) did not explicitly refer to a switch cost associated with this mechanism, it can explain indirectly the existence of such a cost (see Verschooren et al., 2019 and under review for such an application). On switch trials, attention control has to be exerted, arguably by the SAG, because information available in the competing internal or external modality was recently attended. The recruitment of this top-down SAG taxes attentional resources and hence likely explains the cost on these switch trials (Verschooren et al., 2019). However, even though Burgess and colleagues (2007) provided a fine-grained conceptualization of the putative attention control mechanism responsible for the switch cost itself, their account eventually needs to be complemented with an explanation for its asymmetry.

Switching towards internal information is less efficient: priming, associative interference, or retrieval?

An important contribution of our study is to show that the switch cost is asymmetric, being larger for switches towards internal than external trials. In the task-switching literature, such asymmetries have already been observed in the past when switching between a dominant and non-dominant task. Initial observations come from task-switching studies using Stroop stimuli (Stroop, 1935), such as a color word printed in a particular color (e.g., the word "red" printed in blue). Two different tasks can be applied to the same Stroop stimuli: either naming the color word (word naming) or naming the color of the color word (color naming). Word naming (i.e., reading) is a highly practiced, nearly automatic task, whereas color naming is not automatic (Stroop, 1935). Allport et al. (1994) observed that word naming is associated with a larger switch cost than color naming. Two main explanatory frameworks for the (asymmetric) task- switch cost have been proposed in the literature (Vandierendonck et al., 2010): trial-to-trial task priming (Gilbert &

Shallice, 2002; Meuter & Allport, 1999; Yeung & Monsell, 2003a) and associative interference (Bryck & Mayr, 2008; Mayr, Kuhns, & Hubbard, 2014).

Priming

The task priming hypothesis proposes that task sets are carried over from one trial to the next, which interferes with efficient task execution on switch trials. This hypothesis accounts for switch costs asymmetries by assuming that dominant and non-dominant task sets are associated with different attentional control settings when both are relevant within the same context. More specifically, Allport and colleagues initially argued that because word naming is more automatic, it must be inhibited to a large extent on color naming trials to reduce interference (Allport et al. 1994; see also Meuter & Allport, 1999). This high amount of inhibition of the dominant task on non-dominant trials persists in the subsequent trial. If this subsequent trial is a switch toward word naming, a lot of inhibition needs to be overcome and a large switch cost occurs as a result. In contrast, because color naming is not automatic, it does not need to be strongly inhibited on word naming trials. Less inhibition thus persists in the subsequent trial. If this subsequent trial is a switch toward word naming, there is in turn less interference created to overcome, and a smaller switch cost eventually occurs.

Yeung & Monsell (2003a) similarly argued that task priming is responsible for the switch cost asymmetry, but provided a mathematical model in which priming is positive (activation) rather than negative (inhibition). The authors assumed that task priming, i.e. transient increases in activation for recently performed tasks, is especially strong following a trial of the non-dominant task (see also Gilbert & Shallice, 2002) and that (costly) control is minimized where possible. More precisely, on non-dominant trials, control is recruited by necessity to activate the relevant and suppress the irrelevant task. Given this task's low default activation, however, task priming effects

are especially strong. On trials of the dominant task, which have a high default level of activation, no strong task priming effects are observed. These opposite effects can account for the switch cost asymmetry. That is, the dominant task's high activation does not change for repetition or switch trials of the non-dominant task. Given this stable amount of competition, the difference between repetition and switch trials of the non-dominant task lies mostly in the small benefit on repetition trials due to the priming of the non-dominant task. As a result, the switch cost (or repetition benefit) for the non-dominant task is small. On switch trials of the dominant task, on the other hand, there is strong competition with the non-dominant task because of the large priming effect, which is absent on repetition trials. This results in a large switch cost for the dominant task. The advantage of this model over negative priming is that it does not require the assumption of an additional inhibition process – though the authors do acknowledge that this can play a role as well.

These (negative and positive) task priming accounts propose that competition increases specifically following a trial of a weak task and that this increase can explain the cost asymmetry. Translated to our new results, it would then seem that internal attention is dominant and has a stronger default activation than external attention: when participants had to switch from external to internal trials, the additional activation (or inhibition of internal attending) required for externally attending on trial n-1 (because less default activation) hindered the processing of internal information on trial n. Hence, this account assumes that the asymmetry originates from switches between a dominant and non-dominant task set. However, this assumption might be too strong in our case, as there was no evidence that internal attention dominated external attention – and this would actually be somewhat counterintuitive as one would probably a priori predict external attention to dominate internal attention. In addition, even in the context of dominant and non-dominant tasks, these switch cost asymmetries have not always been observed (Monsell, Yeung,

& Azuma, 2000; Rubinstein et al., 2001) and have even been shown to be reversible (Yeung & Monsell, 2003a). Consequently, it is probably worthwhile to consider alternative interpretations.

Associative interference

The associative interference hypothesis for cost asymmetries can be applied more broadly than the task priming hypothesis, as it proposes that switch trials are just a specific instance of WM updating (Mayr et al., 2014). This hypothesis argues, firstly, that switch costs are due to memory interference exerted by earlier trials on the current one where a switch of attention is required (i.e. activation of memory traces from both previously encountered tasks, see Logan, 1988). Secondly, costs are observed on these trials because interference from long-term memory (LTM) is especially disruptive in situations where switches from a maintenance to an updating mode are triggered (Mayr et al., 2014). This is in line with computational models of WM that have proposed that it is primarily based on a gate, which is closed during maintenance but open during updating (Frank, Loughry, & O'Reilly, 2001; O'Reilly & Frank, 2006)

When switch costs are framed this way, potential asymmetries, as for example found when switching between a strong and weak task set, can be accounted for by considering the nature of WM maintenance. Even in the maintenance mode, ongoing behavior needs to remain sensitive to potentially relevant incoming information. This efficiency of the maintenance mode has been demonstrated by studies revealing that WM can be switched to an updating mode when unexpected information is encountered (e.g. Botvinick, Braver, Barch, Carter, & Cohen, 2001; Theeuwes, Kramer, Hahn, & Irwin, 1998), but it plays out differently when switching between tasks of different strengths (Mayr et al., 2014). When performing a task switch trial, an updating process needs to occur, during which performance is highly sensitive to potential interference from the competing task on both the dominant and non-dominant task. However, when performing a

repetition trial of the dominant task, little interference from the non-dominant task occurs, and a robust maintenance mode can be supported. Given the efficiency of the maintenance mode on repetition trials of the dominant task, the performance difference with switch trials will be large. In contrast, when performing a repetition trial of the non-dominant task, competing response tendencies from the dominant task set will trigger updating attempts (Botvinick et al., 2001). As a result, the difference between maintenance and updating on repetition and switch trials, respectively, will not be as pronounced for the non-dominant task. The cost asymmetry then, is due to strong differences between the maintenance and updating mode for the dominant task, and small differences between them on the non-dominant task.

This account is broader than the task priming one in that it can explain the presence of asymmetrical switch costs in studies where no task-switch occurred (e.g. Bryck and Mayr, 2008). In line with this interpretation, it then seems that WM shielding is less efficient on external repetition trials than on internal repetition trials. Presumably, on external trials, even though irrelevant for the task at hand, the memorized figures still need to be kept accessible for upcoming internal trials. This process conflicts with the processing of the externally presented figures, which in turn triggers updating attempt. On internal trials, on the other hand, such a conflict will not be present and maintenance on trial repetitions will be more efficient. This efficient maintenance on repetition trials, then, explains the larger internal switch cost, as the difference between maintenance and updating is less pronounced for external switches.

To sum up, the switch cost asymmetry observed here could be explained in terms of interference caused by the additional activation required on the external trials. In terms of the (positive) priming account, this activation gets carried over to the next trial (positive priming) and results in a larger cost on the internal switch trials. The associative interference account, on the other hand, predicts that this activation triggers updating attempts on the external trials themselves,

which reduce the difference between external repetition and switch trials. At this point, both accounts seem equally able to explain our new findings. There is, however, a third and alternative explanation in terms of memory retrieval that fits the data equally well, which is worth considering on its own merits.

Memory retrieval costs on internal switch trials

An inherent difference between the external and internal trials that might prove instrumental in explaining the cost asymmetry, has so far remained unaddressed. On internal trials, participants have to retrieve the memorized figures, a process that is absent on external trials and that is potentially more costly than perceiving the figures. This is in line with Dark (1990)'s hypothesis of the switch cost asymmetry she observed, and as reviewed in the introduction section. She compared a single external and internal switch with a single external and internal repetition, respectively, which were cued either before (pre cue) or at the moment (simultaneous cue) a target display with a list of six items appeared. She found a larger internal than external switch cost in the simultaneous cue group, but not in the pre cue group. As the asymmetry was observed only in the group where participants could not select the memory item in advance, Dark (1990) concluded that it was due to an additional memory retrieval cost. Even though there are some important differences between our study and Dark (1990)'s, we also found that the switch cost is larger on internal than external trials.

For a memory retrieval effect to explain the asymmetry, this effect needs to be specific to the internal switch trial. This is true if two conditions are met: (a) participants activate the entire array of four figures on each internal trial and (b) they keep this array available in WM from trial-to-trial, despite not knowing whether the next trial will be a repeat trial. Whereas the current results cannot be used to determine whether (a) is true or not, we do have some evidence for point (b) –

which would imply (a) to be true as well. More specifically, in three out of four experiments (see **Table 5**), the fastest RTs are in the internal repetition condition (in Experiment 4 the RTs in that condition are similar to those in the fastest one). The error rates depict a less clear picture, but the switch cost for internal trials is in line with an internal repetition benefit as well. The efficient responding on internal repetition trials suggests that recently refreshed internal representations more accessible than perceptual representation. Though somewhat indirectly, this shows that participants likely kept the array available in WM for the next trial if it was retrieved on the current one. The presence of this costly retrieval process on the first internal (switch) trial of a series and its beneficial effect on the following (repetition) trials can account for the large switch cost observed in this condition. Combined with the absence of these effects in the external condition, this provides a plausible alternative account for the switch cost asymmetry.

Limitations and Future Directions

A limitation of the current series of experiments is that they do not allow us to differentiate unequivocally between a priming, associative interference and a memory retrieval account (see here above) for the observed asymmetric switch cost. Follow-up experiments are needed where these different theoretical accounts could be pitted against one another. For example, if this asymmetry is caused by a memory retrieval effect, a version of the experiment with long preparation interval between cue and stimuli should reduce the asymmetry. On the other hand, if priming causes the asymmetry, a long interval between response and cue should reduce it (but see e.g. Horoufchin, Philipp, & Koch, 2011 for an alternative account). Notwithstanding this limitation, we believe the present results are valuable for two main reasons. First, this asymmetry appears to be robust and resistant to slight changes in the procedure, and hence it could help to constrain existing models of attention or WM, and more specifically attention flexibility, available in the literature. In this respect, they could guide future studies on this important topic, and eventually help assess which theoretical model best accounts for the asymmetry observed during switches of attention between internal and external information. Second and more generally, the novel experimental paradigm used in this study appears especially appropriate to explore further at the behavioral and neuro-anatomical levels switches of attention between visual perception and WM.

A second limitation is that practice was not completely matched for the external and internal task, including in Experiments 2-4 where participants were first familiarized with both tasks separately, before being confronted with their combination within the same block. This resulted from the tradeoff between making the two main conditions most similar at the presentation level and the risk of having participants using a similar cognitive process for both of them (i.e., using mainly memory to help process the external stimuli). More specifically, practice could be deemed specific for the internal task (i.e. matching stimulus to a location) while it was presumably

more general for the external task (i.e. deciding whether a stimulus was part of the set). As a result, we cannot formally exclude the possibility that differences in (relevant) practice between the two modalities might partly explain the asymmetry found between them found in our study. However, because this asymmetry appeared to be resistant to a main change in the amount of practice introduced with the external task while practice with the internal task remained constant throughout (compare results of Experiments 1 and 2), we are confident that an imbalance in the practice between the two modalities alone is unlikely to explain these results.

Finally, though this initial research provides robust evidence for the existence of an asymmetry when switching between external and internal attention, we should be cautious in (over)generalizing this effect to other situations or contexts where alternations between these two modalities occur across successive trials. More specifically, further research is needed is determine whether this finding can generalize to attention switching when different types of stimuli are used.

Conclusions

Using a novel experimental design, we report evidence for the existence of a cost when switching between external and internal attention, which can be explained in terms of a supervisory attention control mechanism for external and internal information. Furthermore, this cost was clearly larger when switching away from external towards internal attention than the reverse direction. We discussed this pattern of results in terms of priming, associative interference and/or LTM retrieval. Additional empirical research is needed to assess which of these competing accounts could best explain this asymmetric switch cost found between visual perception and memory, thereby allowing to constrain further existing cognitive models of attention (flexibility). Moreover, because this paradigm is straightforward and it produces a robust asymmetric switch cost, it could easily be used in future studies to assess possible variations in attention flexibility as a function of specific methodological changes or dispositions.

Acknowledgements

References

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. *Attention and Performance XV: Conscious and Nonconscious Information Processing.*, (January 1994), 421–452. https://doi.org/10.1126/science.1134404

Baddeley, A. (1986). *Working Memory. Oxford Psychology Series*. New York, US: Clarendon Press/Oxford University Press. Retrieved from https://psycnet.apa.org/record/1986-98526-000

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11488380

Bryck, R. L., & Mayr, U. (2008). Task selection cost asymmetry without task switching. *Psychonomic Bulletin & Review*, *15*(1), 128–134. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18605492

Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Sciences*, *11*(7), 290–298. https://doi.org/10.1016/j.tics.2007.05.004

Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Carlson, R. A., Wenger, J. L., & Sullivan, M. A. (1993). Coordinating information from perception and working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *19* 3(3), 531–548. https://doi.org/10.1037/0096-1523.19.3.531

Chun, M.M., Golomb, J., & Turk-Browne, N. B. (2011). A Taxonomy of External and Internal

Attention. *Annual Review of Psychology*, *62*(1), 73–101.

https://doi.org/10.1146/annurev.psych.093008.100427

Chun, Marvin M., & Johnson, M. K. (2011). Memory: Enduring traces of perceptual and

reflective attention. *Neuron*, *72*(4), 520–535. https://doi.org/10.1016/j.neuron.2011.10.026

Dark, V. J. (1990). Switching between memory and perception: moving attention or memory

retrieval? *Memory & Cognition*, *18*(2), 119–127. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/2319955

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and

Language*, *59*(4), 447–456. https://doi.org/10.1016/j.jml.2007.11.004

Endo, N., Saiki, J., Nakao, Y., & Saito, H. (2003). Perceptual judgments of novel contour shapes

and hierarchical descriptions of geometrical properties. *Shinrigaku Kenkyu : The Japanese

Journal of Psychology*, *74*(4), 346–353. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/14708480

Frank, M. J., Loughry, B., & Reilly, R. C. O. (2001). Interactions between frontal cortex and

basal ganglia in working memory : A computational model, *1*(2), 137–160.

Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2005). Involvement of rostral prefrontal cortex in

selection between stimulus-oriented and stimulus-independent thought. *European Journal of

Neuroscience*, *21*(5), 1423–1431. https://doi.org/10.1111/j.1460-9568.2005.03981.x

Gilbert, S. J., & Shallice, T. (2002). Task Switching: A PDP Model. *Cognitive Psychology*, *44*(3),

297–337. https://doi.org/10.1006/cogp.2001.0770

Horoufchin, H., Philipp, A. M., & Koch, I. (2011). The dissipating task-repetition benefit in cued

task switching: Task-set decay or temporal distinctiveness? *Journal of Experimental

Psychology: Human Perception and Performance*, *37*(2), 455–472.

https://doi.org/10.1037/a0020557

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, *14*(89), 81. Retrieved from

https://psycnet.apa.org/record/1928-00322-001

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I.

(2010). Control and interference in task switching--a review. *Psychological Bulletin*, *136*(5),

849–874. https://doi.org/10.1037/a0019842

Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: toward an integrative

account of internal and external selection processes. *Psychonomic Bulletin & Review*, *20*(2).

https://doi.org/10.3758/s13423-012-0359-y

Kosslyn, S. M., Behrmann, M., & Jeannerod, M. (1995). The cognitive neuroscience of mental

imagery. *Neuropsychologia*, *33*(11), 1335–1344. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/8584172

Lennie, P. (2003). The cost of cortical computation. *Current Biology : CB*, *13*(6), 493–497.

Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12646132

Lenth, R. (2018). Estimated Marginal Means, aka Least-Squares Means.

https://doi.org/10.1080/00031305.1980.10483031

Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed

models to analyse reaction time data. *Frontiers in Psychology*, *6*(August), 1–16.

https://doi.org/10.3389/fpsyg.2015.01171

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in

Cognitive Sciences*. https://doi.org/10.1016/j.tics.2005.04.010

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of

Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037//0278-

7393.26.5.1124

Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/0278-7393.29.3.362

Mayr, U., Kuhns, D., & Hubbard, J. (2014). Long-term memory and the control of attentional control. *Cognitive Psychology*, *72*, 1–26. https://doi.org/10.1016/j.cogpsych.2014.02.001

Meiran, N. (1996). Reconfiguration of Processing Mode Prior to Task Performance. *Journal of Experimental Psychology*, *00*(6), 1423–1442. Retrieved from http://www.bgu.ac.il/~nmeiran/index_files/Meiran 1996.pdf

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component Processes in Task Switching. *Cognitive Psychology*, *41*(3), 211–253. https://doi.org/10.1006/cogp.2000.0736

Mesulam, M. (1998). From sensation to cognition. *Brain*, *121*, 1013–1052.

Meuter, R. F. I., & Allport, A. (1999). *Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection*. Retrieved from http://www.idealibrary.com

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 in the random task cuing paradigm. *Acta Psychologica*, *115*(2–3), 123–142. https://doi.org/10.1016/j.actpsy.2003.12.004

Monsell, S, Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: is it easier to switch to the weaker task? *Psychological Research*, *63*(3–4), 250–264. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11004879

Monsell, Stephen. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing Information during Working Memory : Beyond Sustained Internal Attention. *Trends in Cognitive Sciences*, *xx*, 1–13.

https://doi.org/10.1016/j.tics.2017.03.010

Nanay, B. (2015). Perceptual content and the content of mental imagery. *Philosophical Studies*,

(October 2014), 1723–1736. https://doi.org/10.1007/s11098-014-0392-y

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in*

Neuroinformatics, *2*, 10. https://doi.org/10.3389/neuro.11.010.2008

Reilly, R. C. O., & Frank, M. J. (2006). Making Working Memory Work : A Computational

Model of Learning in the Prefrontal Cortex and Basal Ganglia, *328*, 283–328.

Rogers, R., & Monsell, S. (1995). Cost of a predictable switch between simple cognitive tasks.

Journal of Experimental Psychology. General. https://doi.org/10.1037/0096-3445.124.2.207

Rousselet, G. A., Pernet, C. R., & Wilcox, R. R. (2017). Beyond differences in means: robust

graphical methods to compare two groups in neuroscience. *European Journal of*

Neuroscience, *46*(2), 1738–1748. https://doi.org/10.1111/ejn.13610

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in

task switching. *Journal of Experimental Psychology. Human Perception and Performance*,

27(4), 763–797. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11518143

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the

future: the prospective brain. *Nature Reviews Neuroscience*, *8*(9), 657–661.

https://doi.org/10.1038/nrn2213

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for

dynamic events. *Perception*, *28*, 1059–1074. Retrieved from

http://www.chabris.com/Simons1999.pdf

Sohn, M. H., & Anderson, J. R. (2001). Task preparation and task repetition: two-component

model of task switching. *Journal of Experimental Psychology. General*, *130*(4), 764–778.

Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11757879

Speelman, C. P., & McGann, M. (2013). How Mean is the Mean? *Frontiers in Psychology*, *4*, 451. https://doi.org/10.3389/fpsyg.2013.00451

Stroop, R. J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology. Human Perception and Performance*, *13*(1), 89–103. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2951490

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our Eyes do Not Always Go Where we Want Them to Go: Capture of the Eyes by New Objects. *Psychological Science*, *9*(5), 379–385. https://doi.org/10.1111/1467-9280.00071

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*. https://doi.org/10.1037/a0019791

Verbruggen, F., Liefooghe, B., Vandierendonck, A., & Demanet, J. (2007). Short cue presentations encourage advance task preparation: A recipe to diminish the residual switch cost. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(2), 342–356. https://doi.org/10.1037/0278-7393.33.2.342

Verschooren, S., Schindler, S., Raedt, R. de, & Pourtois, G. (2019). Switching attention from internal to external information processing: a review of the literature and empirical support of the resource sharing account. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/S13423-019-01568-Y

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology*. https://doi.org/10.1016/S0010-0285(02)00520-0

Weber, R. J., Burt, D. B., & Noll, N. C. (1986). Attention switching between perception and memory. *Memory & Cognition*, *14*(3), 238–245. https://doi.org/10.3758/BF03197699

Wickham, H. (2016). Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. Retrieved from https://cran.r-project.org/web/packages/ggplot2/index.html

Wickham, H. (2017). Tidyverse: Easily Install and Load the "Tidyverse" [R package tidyverse version 1.2.1]. Retrieved from https://cran.r-project.org/web/packages/tidyverse/index.html

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychological Research*, *63*(3–4), 212–233. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11004877

Yeung, N., & Monsell, S. (2003a). Switching Between Tasks of Unequal Familiarity: The Role of Stimulus-Attribute and Response-Set Selection. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 455–469. https://doi.org/10.1037/0096-1523.29.2.455

Yeung, N., & Monsell, S. (2003b). The effects of recent practice on task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 919–936.