

 Open access • Journal Article • DOI:10.1080/20445911.2017.1365871

Automatic effects of instructions do not require the intention to execute these instructions — [Source link](#)

Baptist Liefoghe, Jan De Houwer

Institutions: Ghent University

Published on: 02 Jan 2018 - Journal of cognitive psychology (Informa UK Limited)

Related papers:

- [Instruction-based task-rule congruency effects.](#)
- [The Power of Instructions: Proactive Configuration of Stimulus-Response Translation](#)
- [There is more into 'doing' than 'knowing': The function of the right inferior frontal sulcus is specific for implementing versus memorising verbal instructions.](#)
- [Following new task instructions: Evidence for a dissociation between knowing and doing](#)
- [Intelligence and the frontal lobe: the organization of goal-directed behavior.](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/automatic-effects-of-instructions-do-not-require-the-5als22cr89>

Automatic effects of instructions do not require the intention to execute these instructions.

Baptist Liefoghe & Jan De Houwer

Department of Experimental-Clinical & Health Psychology, Ghent University

Author Note

This research was supported by grant BOF16/MET_V/002 of Ghent University to Jan De Houwer.

Correspondence concerning this article should be addressed to Baptist Liefoghe, H. Dunantlaan 2, B-9000, Ghent, Belgium. E-mail: baptist.liefoghe@ugent.be

Abstract

Prior research established that newly instructed stimulus-response mappings, which have never been executed overtly before, can lead to automatic response-congruency effects. Such instruction-based congruency effects have been taken as evidence for the hypothesis that the intention to execute stimulus-response mappings results into functional associations that serve future execution. The present study challenges this hypothesis by demonstrating in a series of four experiments that maintaining instructed stimulus-response mappings for future recognition, rather than for future execution, can also lead to an instruction-based congruency effect. These findings indicate that the instruction-based congruency effect emerges even when it is very unlikely that participants form the intention to execute instructions. Alternative interpretations of the instruction-based congruency effect are discussed.

Automatic effects of instructions do not require the intention to execute these instructions.

In recent years a growing body of research focused on the assimilation of new instructions and how this leads to action. The currently prevailing view is that instructions, intended to be executed, can be implemented into a procedural representation in working memory, which guides their execution by enabling reflexive behavior (e.g., Brass, Liefoghe, Braem, & De Houwer, in press; Cohen-Kadosh & Meiran, 2007; Liefoghe, Wenke, & De Houwer, 2012; Meiran, Cole, & Braver, 2012; Wenke, Gaschler, & Nattkemper, 2007). In analogy to Exner's (1879) notion of the "prepared reflex", the implementation of instructions thus leads to a state of preparedness, which Meiran et al. (2012) labeled *intention-based reflexivity*.

The study of intention-based reflexivity mainly uses procedures in which the automatic effect of new and merely instructed Stimulus-Response (S-R) mappings is measured (e.g., Braem et al., 2017; Cohen-Kadosh & Meiran, 2007, 2009; Liefoghe et al., 2012, 2013, 2016; Meiran et al., 2015a, 2015b; Theeuwes et al., 2014; Wenke et al., 2007, 2009, 2015). For instance, Liefoghe et al. (2012) presented participants with different runs of trials on which two tasks had to be performed that shared stimuli and responses: the inducer and the diagnostic task. At the start of each run, participants received two novel arbitrary S-R mappings of the inducer task, each assigning a stimulus either to a left or a right response based on the identity of the stimulus (e.g., If "X", press left; if "Y", press right). Before executing the inducer task, several trials of the diagnostic task were performed, on which participants decided whether a stimulus was presented in *italic* or upright, again by pressing a left or right response key (e.g., upright, press left; italic, press right). After a number of trials of the diagnostic task, a probe stimulus of the inducer task was presented. Liefoghe et al. (2012) observed that performance in the diagnostic task, in terms of speed and sometimes in terms of accuracy, was better when the correct response on the diagnostic task matched with the instructions of the inducer task (e.g., "X" presented upright or "Y" presented in italic) than when the correct response on the diagnostic task did not match

with the S-R mappings of the inducer task (e.g., “Y” presented upright or “X” presented in italic). Given that (1) the diagnostic task was performed immediately after the presentation of the instructions of the inducer task, thus prior to the overt execution of these instructions and (2) the inducer task comprised novel S-R mappings on each run, the conclusion was drawn that the congruency effect observed in the diagnostic task was based on the instructed S-R mappings of the inducer task, which were never executed overtly before.

Several findings suggest that instruction-based congruency effects offer a genuine proxy of intention-based reflexivity. Wenke et al. (2009) observed that the instruction-based congruency effect disappeared when participants were frequently signaled that the inducer task would not proceed, which suggests that participants refrained to implement the instructions under such conditions. Liefoghe et al. (2012) observed an instruction-based congruency effect when the inducer task required the execution of the instructed S-R mappings (see supra), but not when the inducer task required the mere recall or recognition of the instructed S-R mappings. Liefoghe et al. (2012) proposed that when instructed S-R mappings are maintained for future recall or recognition (i.e., when no execution intention is present), they remain in a declarative format in working memory, which does not induce an instruction-based congruency effect. In contrast, when participants have the prospective intention to execute the instructed S-R mappings, a procedural representation of the instructed S-R mappings is formed, which induces an instruction-based congruency effect. This conclusion was furthermore supported by the observation that distinct brain regions are recruited during the implementation as compared to the mere memorization of instructions (Demanet et al., 2016), as well as, by research demonstrating that the instruction-based congruency effect is underlain by pre-motor activation (Everaert et al., 2014; Meiran et al., 2015b). Finally, several studies observed that the instruction-based congruency effect is a function of the degree by which the inducer task is prepared for (Deltomme, Liefoghe, & Braem, submitted; Liefoghe, De Houwer, & Wenke, 2013; Meiran et al., 2015a, Experiment 4), which also suggests its dependency on the intention to execute instructions.

The observation that the instruction-based congruency effect is a function of intention and preparation, furthermore, contrasts with findings concerning the task-rule congruency effect, based on overtly practiced S-R mappings. The task-rule congruency effect is observed when participants have to switch between two overlapping tasks (i.e., task switching; Kiesel et al., 2010; Monsell, 2003; Vandierendonck, Liefoghe, & Verbruggen, 2010). As for the instruction-based congruency effect, the task-rule congruency effect denotes the difference between congruent trials on which both tasks point toward the same response and incongruent trials on which both tasks point toward a different response. Yamaguchi & Proctor (2011) observed a task-rule congruency effect when both tasks were presented in separate blocks of trials, which suggests that this effect is present even if the intention to execute the irrelevant S-R mappings is minimal. In addition, the task-rule congruency effect has been reported to be independent of the degree by which an upcoming task is prepared for (e.g., Meiran, 1996).

Whereas most studies converge towards the hypothesis that the instruction-based congruency effect reflects intention-based reflexivity, some findings, however, suggest that the instruction-based congruency effect may also be induced when the intention to execute instructions is absent. In contrast to what was reported by Liefoghe et al. (2012), these findings show that an instruction-based congruency effect can be obtained when instructions are merely maintained for future recall. More specifically, Theeuwes, De Houwer, Eder, and Liefoghe (2015) investigated whether an instruction-based congruency effect can also be obtained on the basis of contingencies between a response and the effect this response produces in the environment (i.e., an action effect). To this end, the procedure of Liefoghe et al. (2012) was adapted, such that each run started with the instruction of two new Response-Effect contingencies (e.g., if you press the left key, the letter Q will appear on the screen; if you press the right key, the letter P will appear on the screen). In the diagnostic task, the effect-stimulus letters of the inducer instructions (i.e., “Q”, “P”) were used as targets, the orientation of which had to be judged by pressing a left or a right key. An

instruction-based congruency effect was measured by contrasting diagnostic trials on which response and target were part of the same instructed Response-Effect contingency (e.g., left and “Q”) with diagnostic trials on which response and target belonged to different instructed Response-Effect contingencies. Crucially, the inducer task required the recall of the Response-Effect contingencies instructed at the onset of the run. For instance, participants were cued either with the word “left” or “right” and had to press the corresponding key immediately, which resulted in the presentation of an effect stimulus. Next, participants had to judge whether the effect stimulus did or did not match with one of the Response-Effect contingencies that were instructed at the onset of the run (e.g., whether the letter Q actually appeared after pressing the left key). Although this type of inducer task was a recognition task and it was thus unlikely that participants had the intention to execute the instructed Response-Effect contingencies, Theeuwes et al. (2015) observed an instruction-based congruency effect in the diagnostic task.

The results of Theeuwes et al. (2015) indicate that an instruction-based congruency effect can be obtained when the inducer task simply requires the maintenance of instructions, which suggests that the intention to execute instructions is unnecessary to obtain an instruction-based congruency effect. Importantly, even though the procedure of Theeuwes et al. (2015) was fairly similar to the original procedure of Liefoghe et al. (2012), both studies yielded divergent results, with memorization leading to an instruction-based congruency effect in one case (Theeuwes et al., 2015), but not in the other (Liefoghe et al., 2012). The findings of Theeuwes et al. (2015) are, furthermore, in line with other demonstrations, which suggest that simply maintaining information in working memory is sufficient to elicit automatic response effects. This is, for instance, illustrated by research, which focuses on the representation of serial order in working memory (see Abrahamse, van Dijck, Majerus, & Fias, 2014 for a review). A common assumption in this domain is that serial information is maintained by creating a spatial representation in working memory in which serial information is ordered from left to right. Evidence for this hypothesis was reported by van

Dijck and Fias (2011), who presented participants with words that appeared serially on the screen. Participants were asked to maintain these words for an upcoming serial-recall test. In between the learning and test phase, a diagnostic task was nested, that used the to-be-maintained words as targets on which left-right responses had to be made. van Dijck and Fias (2011) observed that words, presented early in the memory sequence facilitated left-sided responses, whereas words presented at the end of the memory sequence facilitated right-sided responses. A response-congruency effect was thus present, which was based on the maintenance of serial order in working memory. These and similar findings (e.g., Abrahamse et al., 2014; De Belder et al., 2015; van Dijck, Abrahamse, Majerus, & Fias, 2013; van Dijck et al., 2014) indicate that maintaining information in working memory for future recall (i.e., without the intention to execute), is thus sufficient to induce automatic effects.

Taken together, the aforementioned findings challenge the conclusions of Liefoghe et al. (2012), who claimed that the instruction-based congruency effect is absent when the inducer task only requires the memorization of instructions for future recall. The possibility thus arises that the instruction-based congruency effect may also be present when the intention to execute instructions is absent, as it is the case for the task-rule congruency effect, based on practiced S-R mappings (Yamaguchi & Proctor, 2011). Accordingly, we reevaluated the conclusions of Liefoghe et al. (2012) by further testing the extent to which the memorization of S-R mappings allows for an instruction-based congruency effect to emerge. To this end, we focused on one of the memorization conditions reported by Liefoghe et al. (2012), namely the visual-recognition condition. In this condition, two S-R mappings were presented visually at the end of each run. Participants had to decide whether these mappings corresponded to the S-R mappings instructed at the onset of the run. When using visual recognition in the inducer task, Liefoghe et al. did not observe an instruction-based congruency effect in the diagnostic task. The present study, however, questions this null finding by considering two concerns about the study of Liefoghe et al. First, their study used a rather small sample and a restricted number of observations. This issue is considered

in Experiment 1 of the present study. Second, the visual-recognition task they used may not have been optimal to induce an instruction-based congruency effect. This issue was dealt with by changing the parameters of the inducer task across Experiments 1 to 4 of the present study.

Experiment 1

In the visual-recognition condition reported by Liefoghe et al. (2012), a small sample ($n=18$) and a restricted number of diagnostic trials (72 congruent trials and 72 incongruent trials) were used. The empirical support provided for the absence of an instruction-based congruency effect when the inducer task required visual recognition thus may not have been sufficient. This possibility is fueled by the fact that the Bayes Factor denoting the weight of evidence provided by the data in support of the null hypothesis (i.e., BF_0) was 2.94 for the reaction times (RTs) and 4.03 for the proportion of errors (PEs). Following Jeffreys (1961)¹, such evidence is anecdotal for the RTs and moderate for the PEs. The question becomes whether the null findings reported by Liefoghe et al. (2012) are indicative of a true null effect or whether the presence of an instruction-based congruency effect remained undetected due to a lack of power. This was tested in Experiment 1 by using a larger sample and tripling the number of diagnostic trials (216 congruent and 216 incongruent). As in the study of Liefoghe et al. (2012), the inducer task required participants to decide whether the S-R mappings presented at the end of each run, matched the S-R mappings instructed at the onset of that run. On half of the runs, the S-R mappings matched completely. On the other half of the runs, the response assignment was reversed. The central question was whether an instruction-based congruency effect could be observed under such conditions.

Method

¹According to Jeffreys (1961), BF scores smaller than 1 designate “no evidence”, BF scores between 1 and 3 designate “anecdotal evidence”, BF scores between 3 and 10 designate “moderate evidence”, and BF scores larger than 10 designate “strong evidence”.

Participants. Thirty-six students at Ghent University participated in return of 10 Euro. Participants were naive to the purpose of the experiment.

Materials. For the test blocks, stimuli were 144 nouns selected from a list of Dutch-naming ratings of the Snodgrass and Vanderwart' (1980) picture-set (Severens, Van Lommel, Ratinckx, & Hartsuiker, 2005). For each participant, a set of 72 pairs of stimuli was randomly constructed on the basis of this list. These pairs were randomly assigned to runs containing 4 trials (18 runs), 8 trials (18 runs), 12 trials (18 runs), or no trials (18 runs) of the diagnostic task. The use of diagnostic tasks of different lengths and the use of runs without a diagnostic task, encourage participants to actively encode and maintain the instructed contingencies during the diagnostic task (Liefoghe et al., 2012, 2013; Meiran et al., 2015a, 2015b). Each pair of stimuli was used for only one run. A practice block preceded the test blocks. This block contained 12 runs, the length of which (0, 4, 8, or 12 diagnostic trials) was determined randomly. Stimuli in this practice block were given names.

In the diagnostic task, participants indicated whether a stimulus was printed upright or in italic by pressing left (A-key) or right (P-key) on an AZERTY keyboard. The response assignment of the diagnostic task was determined randomly across participants. In each run containing a diagnostic task, half of the trials were congruent (inducer and diagnostic task point to the same response) and the other half were incongruent (inducer and diagnostic task point to different responses). For the inducer task, a pair of S-R mappings was presented. Participants decided whether these mappings matched with the pair of S-R mappings they had to memorize at the onset of each run. They pressed left (A-key) when the mappings matched and right (P-key) when they did not match. On half of the runs, the S-R mappings matched with the initial S-R mappings. The response labels of the inducer task were also presented when the probe of the inducer task appeared (see Figure 1). S-R mappings and stimuli were presented in ARIAL font, size 16.

Procedure. Participants were tested in groups of two or three. They were placed in front of personal computers attached to a 17-inch color monitor. The experiment was

programmed by using the Tscope library for C/C++ (Stevens, Lammertyn, Verbruggen, & Vandierendonck, 2006). At the start of the experiment, the overall instructions were presented and paraphrased if necessary. The experiment started with the practice block, followed by four test blocks, with a small break after each block. During each break, feedback was provided about the performance on the diagnostic task and the inducer task.

A schematic overview of a run is presented in Figure 1. Each run started with the presentation of two S-R mappings. The S-R mappings remained onscreen until participants pressed the space bar or a maximum time of 30s elapsed. The position of the S-R mapping was determined randomly, so that the instructions referring to a certain response could be presented either above or below the screen center. The first stimulus of the diagnostic task was presented 750ms after the S-R mappings were removed from the screen. Each stimulus in the diagnostic task remained onscreen until participants responded or a maximum response time of 2000ms elapsed. Depending on the run, participants performed the diagnostic task for 4, 8, 12 trials or not at all. The different types of runs were presented in a random order. The inter-trial interval in the diagnostic task was 750ms. 750ms after the last response in the diagnostic task, two S-R mappings appeared, which remained on-screen for 5000ms or until participants responded. The up-down position of these two S-R mappings again varied randomly. A new run started 1,500ms after the participants performed the inducer task. After each incorrect response, the screen turned red for 200ms. The experiment lasted approximately 45 min.

Results

For each experiment, the data of the practice block were not included in the analyses. Visual inspection of the data revealed that four participants had extremely low accuracies either on the diagnostic task (.50, .50, .51) or on the inducer task (.46). The data of these participants were excluded from the analyses. For the inducer task, the average RT was 1656ms ($SD= 380$) and the average PE was .06 ($SD= .06$).

For the diagnostic task, only diagnostic trials of runs on which the inducer task was responded to correctly were considered. For RTs, only correct trials were considered, which led to the removal of 4.81% of the total number of trials. Next, for each participant, trials with RTs more than 2.5 standard deviation above each cell mean were considered as outliers. This led to the removal of 3.01% of the total number of correct trials. Data were analyzed by using t-tests with an alpha-level of .05. Bayes Factors were additionally computed by using the R-package “BayesFactor” (Morey, Rouder, & Jamil, 2015). RTs on congruent trials ($M=586\text{ms}$, $SD=78$) did not differ significantly from RTs on incongruent trials ($M=587\text{ms}$, $SD=83$), $t < 1$, $BF_0=4.88$. PEs were significantly higher on incongruent trials ($M=.06$, $SD=.05$) than on congruent trials ($M=.04$, $SD=.04$), $t(31)=3.78$, $p < .001$, $r^2=.32$, $BF_1=46.71$. Because RTs and PEs of the diagnostic task suggested different patterns of results, both measures were combined by using the Linear Integrated Speed-Accuracy Score (LISAS, Vandierendonck, 2017)². When using this combined score, the difference between congruent and incongruent trials was significant, $t(31)=2.47$, $p < .05$, $r^2=.16$, $BF_1=2.53$.

Discussion

For the RTs, moderate evidence was obtained in support of the null hypothesis (i.e., no instruction-based congruency effect). As such, the findings of Liefoghe et al. (2012) were replicated. Strong evidence was observed for the PEs, which is at odds with the findings of Liefoghe et al. (2012). When combining RT and PE by using LISAS, anecdotal evidence in favor of the presence of an instruction-based congruency effect was obtained. The results of Experiment 1 suggest that part of the null findings reported by Liefoghe et al. (2012) may have been the result of using a small sample size and/or a restricted number of diagnostic trials. Nevertheless, the evidence obtained in favour of the hypothesis that an instruction-based congruency effect can be obtained when the inducer task requires visual recognition, was generally weak. In the following experiments, we investigate whether

²LISAS for subject i in condition j is defined as $\text{Mean_RT}_{ij} + (\text{StDev_RT}_i / \text{StDev_PE}_i) \times \text{Mean_PE}_{ij}$.

changing the parameters of the inducer task could not yield stronger evidence.

Experiment 2

In the visual-recognition task used in Experiment 1 (and in the study of Liefoghe et al., 2012), the S-R mappings presented at the end of each run were either identical to the instructed S-R mappings or they comprised reversed response assignments. This inducer task could have been completed by only encoding one of the two instructed S-R mappings. If that same S-R mapping is presented again in the probe of the inducer task, then it follows that the other S-R mapping is also correct. If that S-R mapping includes the opposite response in the probe, then it follows that the other S-R mapping is also incorrect. Possibly, when the inducer task requires visual recall, memorizing only one S-R mapping is not sufficient to obtain an instruction-based congruency effect. This hypothesis is fueled by the study of Theeuwes et al. (2015). As was already mentioned in the Introduction, an instruction-based congruency effect was observed in this study, even though the inducer task was a recall task. Importantly, in the study of Theeuwes et al. (2015) participants were encouraged to encode and maintain both instructed contingencies.

The aim of Experiment 2 was to encourage participants to maintain both instructed S-R mappings (cfr. Theeuwes et al., 2015). To this end, the inducer task used filler items. These filler stimuli were not part of the instructed S-R mappings, but could appear during the recall phase of the inducer task. For instance, at the onset of a run the S-R mappings “If car, press left” and “If rabbit, press right” were presented and at the end of the run three scenarios were possible: (a) the two instructed S-R mappings were presented again and thus both mappings were correct; (b) one of the two instructed S-R mappings was presented (e.g., “If car, press left”) together with a new S-R mapping (e.g., “If fish, press right”) and thus only one of the two S-R mappings was correct; or (c) both S-R mappings were incorrect either because the response assignment of the two instructed S-R mappings was reversed (i.e., “if car, press right” and “if rabbit, press left”) or because a new S-R mapping was

presented (e.g., “if fish, press right”) together with an instructed stimulus, which was assigned to an incorrect response (e.g., “if rabbit, press left”). Participants now had three response options to indicate each scenario, with each option occurring on 1/3 of the runs. Such inducer task cannot be completed by memorizing only one instructed S-R mapping and participants were thus encouraged to maintain both S-R mappings. In order to test whether this manipulation affected performance in the inducer task, we compared performance on the inducer task of Experiment 2 with performance on the inducer task of Experiment 1.

With respect to the diagnostic task, the question was whether stronger evidence could be obtained for the hypothesis that an instruction-based congruency effect is present when using a recall task as an inducer task. This was investigated by calculating the *BFs* of the different dependent measures. In addition, we also investigated whether this instruction-based congruency effect would differ in size compared to the effect obtained in Experiment 1. To this end, the LISAS of Experiment 2 was compared to the LISAS of Experiment 1.

Method

Thirty-nine participants were recruited, none of whom had participated in the previous experiment. Experiment 2 was identical to Experiment 1, except that the inducer task now was a three-choice recognition task. If both S-R mappings were correct, participants pressed left (A-key). If only one S-R mapping was correct, participants pressed the spacebar. If both S-R mappings were incorrect, participants pressed right (P-key). As in Experiment 1, the response labels of the inducer task were also presented on screen (see Figure 1 for an example). The filler words were also selected from the picture-naming set of Severens et al. (2005). For the practice block, the filler words were also given names.

Results

Three participants were excluded on the basis of very low accuracies either in the diagnostic task (.52, .49) or in the inducer task (.50). For the inducer task, the average RT

was 2413ms ($SD= 375$) and the average PE was .11 ($SD= .08$). Inducer-task RTs of Experiment 2 were significantly longer than the inducer-task RTs of Experiment 1, $t(66)= 8.26$, $p < .001$, $r^2= .51$, $BF_1 > 100$. Furthermore, inducer-task PEs of Experiment 2 were significantly higher compared to the inducer-task PEs of Experiment 1, $t(66)= 3.43$, $p < .001$, $r^2= .15$, $BF_1 = 29.59$.

For the diagnostic task, the overall error rate was 4.44% and 2.74% of the correct trials were identified as RT-outliers. RTs on congruent trials ($M= 560$ ms, $SD= 73$) were shorter than RTs on incongruent trials ($M= 568$ ms, $SD= 82$), $t(35)= 2.12$, $p < .05$, $r^2= .11$, $BF_1= 1.31$. The PEs were marginally higher on incongruent trials ($M= .07$, $SD= .05$) than on congruent trials ($M= .06$, $SD= .05$), $t(35)= 1.76$, $p= .09$, $r^2= .08$, $BF_1= 0.72$. For the LISAS, the difference between congruent and incongruent trials was also significant, $t(35)= 2.66$, $p < .05$, $r^2= .17$, $BF_1= 3.68$. The latter effect did not differ significantly in size from the instruction-based congruency effect - measured in LISAS - of Experiment 1, $F < 1$.

Discussion

RTs and PEs of the inducer task in Experiment 2 were significantly larger compared to Experiment 1. This difference suggests that the three-choice inducer task used in Experiment 2 was more difficult compared to the two-choice inducer task used in Experiment 1, supposedly because the inducer task in Experiment 2 was a three-choice task, which required the maintenance of both instructed S-R mappings. Nevertheless, the instruction-based congruency effect just reached significance for the RTs, whereas it was only marginally significant for the PEs. BF scores indicated that the evidence provided for the alternative hypothesis was anecdotal for the RTs and, following Jeffreys (1961), inexistent for the PEs. The combined score (LISAS) indicated that moderate evidence for an instruction-based congruency effect was obtained. The BF of the LISAS was in the same range as in Experiment 1. Encouraging participants to encode both S-R mappings did not yield stronger evidence in favor of the hypothesis that an instruction-based congruency effect

can be obtained when the inducer task requires visual recall. In addition, the between-experiment comparison indicated that the instruction-based congruency effect in Experiment 2 did not differ in size from the instruction-based congruency effect observed in Experiment 1. In Experiments 3 and 4 an additional factor is considered, namely the time available to perform the inducer task.

Experiment 3

In Experiments 1 and 2, the response deadline of the inducer task was 5000ms. When considering the overall RT in the inducer task in view of this response deadline, completing the inducer task required on average 34% of the available time in Experiment 1 and 45% of the available time in Experiment 2. It seems thus reasonable to conclude that the 5000ms response deadline was fairly lenient. Previous studies in which the inducer task required the execution of instructed S-R mappings, demonstrated that the instruction-based congruency effect in the diagnostic task is a function of the extent to which the inducer task is prepared for in advance. For instance, Liefoghe et al. (2013) assumed that imposing a strict response deadline in the inducer task would encourage participants to prepare more thoroughly for the inducer task and thus require them to maintain the instructed S-R mappings more actively during the diagnostic task. In line with their hypothesis, an instruction-based congruency effect was observed in the diagnostic task when the deadline of the inducer task was strict but not when it was lenient (see also Meiran et al., 2015; Experiment 4 for similar findings). Furthermore, Deltomme et al. (submitted) demonstrated in a correlational study that faster RTs in the inducer task are associated with a larger instruction-based congruency effect in the diagnostic task. Based on these findings, the question arises whether the degree by which the inducer task is prepared for, also modulates the instruction-based congruency effect when the inducer task requires the recall of the instructed S-R mappings, the intention to execute the instructed S-R mappings thus being absent. This hypothesis is again fueled by the aforementioned study of Theeuwes et al. (2015) in which the response deadline of the

inducer task was 1500ms. Possibly, participants in Experiments 1 and 2 were not sufficiently encouraged to prepare for the inducer task, because the response deadline of the inducer task was too lenient. Accordingly, in Experiment 3 we used a two-choice recognition task as an inducer task (cfr., Experiment 1), but now with a 1500ms deadline. Participants could in principle still complete the inducer task by encoding only one of the two instructed S-R mappings, but they only had 1500ms to respond in the inducer task. First, we tested whether imposing such deadline did indeed encourage participants to prepare for the inducer task more thoroughly. To this end, performance on the inducer task in Experiment 3 was compared to performance on the inducer task of Experiment 1. Second, we investigated whether, under such conditions, stronger evidence could be obtained for the hypothesis that an instruction-based congruency effect is present when the inducer task merely requires the recall of instructed S-R mappings. To this end, we calculated the *BFs* of the different dependent measures. Finally, we investigated whether the instruction-based congruency effect was larger when imposing a strict response deadline in the inducer task, by comparing this effect between Experiments 1 and 3.

Method

Thirty-five participants were recruited who did not participate in the previous experiments. The use of a 1500ms response deadline in the inducer task led to an additional change: whereas the S-R mappings were presented in the form of two sentences as in the previous experiments, the probe of the inducer task now consisted of two target words. One word was presented on the left side of the screen and the other word on the right side of the screen (see Figure 1). Participants had to decide whether the left-right location of the words matched with the relations described in the instructed S-R mappings (e.g., whether the word “CAR” on the left, matches with the S-R mapping “If CAR, press left”).

Results

Six participants were excluded on the basis of very low accuracies either in the diagnostic task (.61) or in the inducer task (.54, .55, .56, .54, .60). For the inducer task, the average RT was 838ms ($SD= 93$) and the average PE was .08 ($SD= .05$). Inducer-task RTs of Experiment 3 were significantly shorter than the inducer-task RTs of Experiment 1, $t(59)= 11.28, p < .001, r^2 = .68, BF_1 > 100$. In contrast, inducer-task PEs of Experiment 3 did not differ significantly from inducer-task PEs of Experiment 1, $t(59)= 1.41, p = .166, r^2 = .03, BF_0 = 1.66$.

For the diagnostic task, the overall error rate was 5.80% and 2.62% of the correct trials were identified as RT-outliers. RTs on congruent trials ($M= 573\text{ms}, SD= 60$) were shorter than RTs on incongruent trials ($M= 579\text{ms}, SD= 65$), $t(28)= 2.19, p < .05, r^2 = .15, BF_1 = 1.54$. The PEs were higher on incongruent trials ($M= .07, SD= .04$) than on congruent trials ($M= .05, SD= .03$), $t(28)= 3.33, p < .01, r^2 = .28, BF_1 = 15.26$. When combining RTs and PEs by using LISAS, the instruction-based congruency effect was again significant, $t(28)= 3.27, p < .01, r^2 = .28, BF_1 = 13.38$. The latter effect did not differ significantly in size from the instruction-based congruency effect observed in Experiment 1, $F < 1$.

Discussion

For the inducer task, the RTs were shorter compared to Experiment 1. Although the difference in PEs was not significant, evidence in favor of the null hypothesis was anecdotal. Because PEs were numerically larger in Experiment 3 compared to Experiment 1, the overall difference in inducer performance between Experiment 1 and Experiment 3 suggests that imposing a strict response deadline resulted in a speed-accuracy trade-off. It could thus be argued that participants prioritized speed above accuracy, rather than that participants prepared for the inducer task more thoroughly. Nevertheless, the LISAS of the inducer task in Experiment 3 remained significantly below the LISAS of the inducer task in Experiment 1, $t(59)= 10.47, p < .001, r^2 = .65, BF_1 > 100$. This combined measure of speed and accuracy

thus suggests that performance in the inducer task of Experiment 3 was overall better than performance in the inducer task of Experiment 1. As such, the strict response deadline that was used in Experiment 3, seemed to have succeeded in its goal.

For the diagnostic task, a significant instruction-based congruency effect was present both for the RTs and the PEs. Whereas evidence for the RTs was anecdotal, the effect observed for the PEs could be qualified as being strong (Jeffreys, 1961). When both measures were combined by using the LISAS, evidence in favour of an instruction-based congruency effect was also strong. Imposing a strict response deadline in the inducer task thus seems to increase the likelihood of observing an instruction-based congruency effect. However, the instruction-based congruency effect obtained in Experiment 3 did not differ significantly in size from the instruction-based congruency effect obtained in Experiment 1.

Experiment 4

Compared to the previous experiments, Experiment 3 yielded stronger evidence in favor of the hypothesis that an instruction-based congruency effect can be obtained when the inducer task merely requires recall. Nevertheless, the size of the instruction-based congruency effect in Experiment 3 was not significantly larger compared to the size of the instruction-based congruency effect in Experiment 1, which served as a baseline. The aim of Experiment 4, was to replicate the findings of Experiment 3 and test whether even stronger evidence could be obtained. To this end, the requirement to maintain both instructed S-R mappings and the requirement to prepare for the inducer task in advance were combined, as it was the case in the study of Theeuwes et al. (2015). To this end, we used a three-choice inducer task but now imposed a stricter response deadline. Because Experiment 2 used a three-choice inducer task with a lenient response deadline, this experiment now served as the baseline for the between-experiment comparison. First, we tested whether narrowing the response deadline changed performance in the inducer task, in such a way that is was indicative of more advance preparation. Second, we calculated the amount of evidence in

favor of an instruction-based congruency effect by using the *BFs* of the dependent measures in the diagnostic task. Finally, we investigated whether the size of the instruction-based congruency in Experiment 4 was larger than in Experiment 2.

Method

Twenty-nine new participants were recruited. Experiment 4 was identical to Experiment 2, except that the response deadline of the inducer task was now 2500ms instead of 5000ms. We opted for 2500ms because piloting sessions indicated that a response deadline of 1500ms was too hard for the participants. In line with Experiment 3, the probe of the inducer task consisted of two target words. One word was presented on the left side of the screen and one word on the right side of the screen (see Figure 1). As in Experiment 2, filler words were used such that on 1/3 of the runs both target words were on the correct location, on 1/3 of the runs both target words were on the incorrect location and on 1/3 of the runs only one of the two target words was on the correct location. The inducer task used the same response keys as in Experiment 2. Response labels of the inducer task, were again provided.

Results

Five participants were excluded on the basis of very low accuracies either in the diagnostic task (.49, .50) or in the inducer task (.47, .33, and .54). For the inducer task, the average RT was 1435ms ($SD= 156$) and the average PE was .10 ($SD= .09$). Because in Experiment 4, we varied the response deadline of the inducer task used in Experiment 2, we compared performance in the inducer task of Experiment 4 with performance in the inducer task of Experiment 2. Inducer-task RTs of Experiment 4 were significantly shorter than the inducer-task RTs of Experiment 2, $t(58)= 12.08$, $p < .001$, $r^2= .71$, $BF_1 > 100$. In contrast, inducer-task PEs of Experiment 4 did not differ significantly from inducer-task PEs of Experiment 2, $t < 1$.

For the diagnostic task, the overall error rate was 5.67% and 2.79% of the correct trials were identified as RT-outliers. RTs on congruent trials ($M= 557$ ms, $SD= 71$) were shorter

than RTs on incongruent trials ($M = 572\text{ms}$, $SD = 76$), $t(23) = 3.74$, $p < .01$, $r^2 = .38$, $BF_1 = 32.77$. The PEs were higher on incongruent trials ($M = .08$, $SD = .07$) than on congruent trials ($M = .04$, $SD = .03$), $t(23) = 2.91$, $p < .01$, $r^2 = .27$, $BF_1 = 5.90$. For the LISAS, the instruction-based congruency effect was also significant, $t(23) = 3.89$, $p < .001$, $r^2 = .39$, $BF_1 = 46.06$. The latter effect was significantly larger compared to the instruction-based congruency effect - measured in LISAS - observed in Experiment 2 in which the same three-choice inducer task was used with a 5000ms response deadline, $F(1,58) = 4.47$, $MSE = 626$, $p < .05$, $\eta_p^2 = .07$. Nevertheless, a Bayes factor ANOVA (Morey et al., 2015) revealed that the model including an interaction term (i.e., Congruency by Experiment) was preferred to a main effects model without an interaction term only by a Bayes Factor around 1.5.³ In an additional analysis, we also compared the instruction-based congruency effect obtained in Experiment 4, with the instruction-based congruency effect of Experiment 1, in which a two-choice inducer task with a 5000ms response deadline was used. The effect obtained in Experiment 4 was again larger, $F(1,54) = 4.66$, $MSE = 617$, $p < .05$, $\eta_p^2 = .08$. Yet, the interaction model was again only preferred over the main-effects model with a Bayes Factor around 1.5.

Discussion

For the inducer task, participants were significantly faster compared to Experiment 2. Both experiments did not differ in terms of PE. Imposing a strict response deadline did not result in a speed-accuracy trade-off and participants were thus genuinely encouraged to prepare for the inducer task more thoroughly compared to Experiment 2.

For the diagnostic task, a strong instruction-based congruency effect was observed for the RTs and a moderate instruction-based congruency effect for the PEs. The LISAS also offered strong evidence in favour of the alternative hypothesis, as it was the case in Experiment 3. Finally, the instruction-based congruency effect in Experiment 4 was

³Small variations occurred in the values of the BF due to Monte Carlo sampling noise. Re-running the analysis yielded BFs around 1.5.

significantly larger compared to the instruction-based congruency effects observed in Experiments 1 and 2. However, when conducting a Bayesian analysis, only anecdotal evidence was obtained favoring the interaction models over the main-effects models. Some cautiousness is thus needed when interpreting these between-experiment comparisons.

General Discussion

In the present study, we further investigated the common assumption that the instruction-based congruency effect requires the intention to execute instructions (e.g., Cohen-Kadosh & Meiran, 2007, 2009; De Houwer et al., 2005; Everaert et al., 2014; Liefoghe et al., 2012, 2013; Meiran & Cohen-Kadosh, 2012; Meiran et al., 2012, 2015a, 2015b; Wenke, et al., 2007, 2009, 2015). To this end, we reevaluated the conclusion that the instruction-based congruency effect cannot be obtained when participants maintain instructed S-R mappings for future recall or recognition (Liefoghe et al., 2012). We used an inducer task, which required visual recognition and measured the presence of an instruction-based congruency effect in the diagnostic task. Liefoghe et al. (2012) did not observe an instruction-based congruency effect with such set-up. We first argued that their null findings may have been the result of using an underpowered design. In view of this concern, a conceptual replication was conducted in Experiment 1, for which the sample size and the number of observations were substantially increased. Moderate evidence was obtained for the hypothesis that an instruction-based congruency effect emerges when the inducer task merely requires recall. Second, we reasoned that the visual-recognition task used by Liefoghe et al. (2012), and in Experiment 1 of the present study, may not have been optimal to induce an instruction-based congruency effect. Accordingly, in three follow-up experiments, we tested whether changing the parameters of the inducer task could lead to stronger evidence in favor of an instruction-based congruency effect. In Experiment 2, participants were encouraged to maintain both instructed S-R mappings by using a three-choice inducer task. In Experiment 3, participants were encouraged to prepare for the

inducer task more thoroughly by using a two-choice visual-recognition task (cfr., Experiment 1) with a strict response deadline. In both experiments moderate (Experiment 2) to strong evidence (Experiment 3) in support of the alternative hypothesis was observed. Finally, in line with the study of Theeuwes et al. (2015; Experiment 3) - which was the prime motivation of conducting the present study - the manipulations used in Experiments 2 and 3 were combined. In Experiment 4, we observed strong evidence for the hypothesis that an instruction-based congruency effect is also present when the inducer task requires recall. This effect was, furthermore, slightly larger than in Experiments 1 and 2. Taken together, our results refute the conclusions of Liefoghe et al. (2012) by demonstrating that an instruction-based congruency effect can be observed even when the inducer task requires the mere memorization of instructed S-R mappings. The present study thus challenges the widely spread assumption that the instruction-based congruency effect requires the intention to execute instructions and questions whether this effect is a proxy of instruction-based reflexivity.

Our results suggest that the instruction-based congruency effect needs to be interpreted without calling upon intention-based reflexivity. A first such interpretation is that the instruction-based congruency effect can also be obtained on the basis of a declarative representation in working memory. As mentioned in the Introduction, van Dijk and Fias (2011) observed a response compatibility effect on the basis of serial order representations in working memory. In essence, both their procedure and the procedure of Liefoghe et al. (2012, 2013) are complex-span tasks in which the trade-off is investigated between the storage of contents (serial-order information or instructed S-R mappings) and the processing of a task (i.e., the diagnostic tasks). Oberauer, Demmrich, Mayr, and Kliegl (2001; see also Oberauer, 2002) argued that the trade-off between storage and processing is a function of a cross-talk between both components of a complex-span task. More specifically, they proposed that the impact of storage on processing depends on the requirement to access the stored information during the processing component. When a processing task is

constructed in such a way that the information maintained for the storage task is accessed during the processing task, then performance in the processing task is impeded. In contrast, when the stored information is not accessed during the processing task, no trade-off between storage and processing is present. Although we omitted this feature when discussing the work of van Dijck and Fias (2011) in the Introduction (see, p. 7), it is important to note that these authors used filler items in their diagnostic task, which were not part of the memorized set of items. During the diagnostic task, participants were instructed to react only to the targets belonging to the memorized set. Accordingly, participants were required to frequently access the memorized set of items during the diagnostic task, which induced a cross-talk between the memorized information and the diagnostic task.

In the procedure of Liefoghe and colleagues (2012), cross-talk can be induced through different demands. The intention to execute the instructed S-R mappings constitutes one such demand. It is reasonable to assume that when the instructed S-R mappings need to be executed, the S-R mappings of the inducer task are accessed during the diagnostic task, because the goal of the inducer task and the diagnostic task are highly overlapping, namely to execute S-R mappings. The intention to actively maintain instructed S-R mappings for future recall or recognition may, however, also constitute a cross-talk inducing demand. Barrouillet, Bernardin, & Camos (2004) proposed that in a complex-span task, the stored information is frequently refreshed during the processing task-component, which leads to a trade-off. The presence of an instruction-based congruency effect in a memorization condition, could be considered as the result of such trade-off. The larger instruction-based congruency effect observed in Experiment 4 may suggest that when participants are encouraged to maintain both instructed S-R mappings and to prepare for the inducer task more thoroughly, the rate by which the instructed S-R mappings are refreshed during the diagnostic task is increased. Nevertheless, we remain cautious with this latter hypothesis as the between-experiment comparisons only indicated rather weak differences between Experiment 4 and Experiments 1 & 2.

Although declarative S-R mappings may be frequently accessed during the diagnostic task, the question remains how these declarative representations can lead to an instruction-based congruency effect, which is associated with automatic pre-motor activation (Everaert et al., 2014, Meiran et al., 2015b). One possibility is that such motor activation is obtained through semantic priming. More specifically, during the diagnostic task irrelevant stimuli activate the semantic concepts left or right on the basis of the declarative representations of the instructed S-R mappings. Based on these semantic concepts, the corresponding response concepts are activated that are either compatible (congruent diagnostic trials) or incompatible (incongruent diagnostic trials) with the response that is required by the diagnostic task. Support for such hypothesis comes from a study by Bundt et al. (2015), who observed that seeing the words “left” or “right”, in a context in which left-right discriminations are frequent, can lead to motor activation. We agree that such semantic-priming account is speculative. However, we also emphasize that, at present time, research on the instruction-based congruency effect is almost entirely based on tasks using left and right responses (see Wenke et al., 2007, 2009; Tibboel, Liefoghe, & De Houwer, 2016 for exceptions).

An alternative interpretation of the current findings is that the intention to recall or recognize instructed S-R mappings also leads to the formation of procedural representations, as it is the case for the intention to execute the instructed S-R mappings. As such, the assumption of Liefoghe et al. (2012) that the memorization of instructions is uniquely based on declarative representations would be wrong. Some findings in the literature do indeed suggest that simply maintaining instructions may also call upon procedural representations. Several studies demonstrated that instructions are better recalled when physically enacted than when verbally rehearsed (e.g., Allen & Waterman, 2015; Gathercole et al., 2008; Yang, Gathercole, & Allen, 2014). Accordingly, the maintenance of instructed S-R mappings may also be accomplished through the covert enactment of these instructions, which would call upon a procedural representation. Such account relates to the study of Ruge and

Wolfensteller (2010), who proposed that part of the processing of new S-R mappings relies upon motor imagery (see also, Theeuwes, Liefoghe, De Schryver, & De Houwer, in press). Such covert enactment may even occur independently of the goal of the inducer task (i.e., instruction execution or instruction recall), being automatically triggered by the processing of the action related information conveyed by the instructed S-R mappings.

The observation that the instruction-based congruency can also emerge when the intention to execute instructions is absent, is in line with the results of Yamaguchi & Proctor (2011), who observed that the task-rule congruency in task switching is present even when the intention to execute the irrelevant S-R mappings is minimal. At the present time it is assumed that a substantial difference exists between both types of effects. Whereas the task-rule congruency effect is based on S-R associations in long-term memory (e.g., Kiesel, Wendt, & Peters, 2007; Meiran & Kessler, 2008), most studies - such as the current one - suppose that the instruction-based congruency effect is based on working-memory functioning and some studies provided very direct evidence for this hypothesis by using dual-task manipulations (e.g., Cohen-Kdoshay & Meiran, 2007; Meiran et al, 2015; Meiran & Cohen-Kdoshay, 2012). Nevertheless, recent studies by Pfeuffer et al (2017, in press-a, in press-b) suggest that automatic effects of instructions may also be underlain by representations in long-term memory. Accordingly, the processes underlying the instruction-based congruency effect and the task-rule congruency effect may be less distinct than was previously assumed (e.g., Liefoghe et al., 2012).

Taken together, the core conclusion of the present study is that the instruction-based congruency effect is not confined to the intention to execute instructions and the processes underlying this effect seem more complex than the mere conception of intention-based reflexivity. Future research will be needed to pinpoint its exact nature. For now, the main message of the present study is that we need to be cautious when drawing conclusions about intention-based reflexivity on the basis of the instruction-based congruency effect.

References

- Abrahamse, E.L., van Dijck J.-P., Majerus, S., & Fias, W. (2014). Finding the answer in space: the mental whiteboard hypothesis on serial order in working memory. *Frontiers in Human Neuroscience*, *9*, 932.
- Allen, R.J., & Waterman, A.H. (2015). How does enactment affect the ability to follow instructions in working memory? *Memory and Cognition*, *43*, 555-561
- Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology : General*, *133*, 83-100.
- Braem, S., Liefoghe, B., De Houwer, J., Brass, M., & Abrahamse E.L. (2017). There are limits to the effects of task instructions: Making the automatic effects of task instructions context-specific takes practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. *43*, 394-403.
- Brass, M., Liefoghe, B., Braem, S., & De Houwer, J. (in press). Following new task instructions: Evidence for a dissociation between knowing and doing. *Neuroscience & Biobehavioral Reviews*.
- Bundt, C., Bardi, L., Abrahamse, E.L., Brass, M., & Notebaert, W. (2015). It wasn't me! Motor activation from irrelevant spatial information in the absence of a response. *Frontiers in Human Neuroscience*, *9*:539.
- Cohen-Kadosh, O., & Meiran, N. (2007). The representation of instructions in working memory leads to autonomous response activation: Evidence from the first trials in the flanker paradigm. *The Quarterly Journal of Experimental Psychology*, *60*, 1140-1154.
- Cohen-Kadosh, O., & Meiran, N. (2009). The representation of instructions operates like a prepared reflex: Flanker compatibility effects found in first trial following S-R instructions. *Experimental Psychology*, *56*, 128-133.
- De Belder, M., Abrahamse, E.L., Kerckhof, E., Fias, W., & van Dijck, J.-P. (2015). Serial position markers in space: visuospatial priming of serial order working memory

- retrieval. *PloS one*, 10, e0116469.
- Deltomme, B., Liefoghe, B., & Braem, S. (submitted). Faster execution of task instructions is associated with more instruction-based interference: Evidence from inter- and intra-individual differences.
- Demanet, J., Liefoghe, B., Hartstra, E., Wenke, D., De Houwer, J., & Brass, M. (2016). There is more into “doing” than “knowing”: The function of the right inferior frontal sulcus is specific for implementing versus memorizing verbal instructions. *Neuroimage*, 141, 207- 213.
- Everaert, T., Theeuwes, M., Liefoghe, B., & De Houwer, J. (2014). Automatic motor activation by mere instruction. *Cognitive, Affective, and Behavioral Neuroscience*, 14, 1300-1309.
- Exner, S. (1879). Physiologie der Grosshirnrinde. In L. Hermann (Ed.). *Handbuch der Physiologie*, vol. 2, part 2, pp., 189-350. Leipzig: Vogel.
- Gathercole, S.E., Durling, M., Evans, S., Jeffcock, & S. Stone (2008). Working memory abilities and children’s performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology*, 22, 1019-1037.
- Jeffreys, H. (1961). *Theory of Probability*, 3rd ed. Oxford Univ. Press.
- 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, 849.
- Kiesel, A., Wendt, M., & Peters, A. (2007). Task Switching: On the origins of response congruency effects. *Psychological Research*, 71, 117-125.
- Liefoghe, B., Degryse, J., & Theeuwes, M. (2016). Automatic effects of No-Go instructions. *Canadian Journal of Experimental Psychology*, 70, 232-241.
- Liefoghe, B., De Houwer, J., & Wenke, D. (2013). Instruction-based response activation depends on task preparation. *Psychonomic Bulletin & Review*, 20, 481-487.
- Liefoghe, B., Wenke, D., & De Houwer, J. (2012). Instruction-based task-rule congruency

- ef-fects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *38*, 1325-1335.
- Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1423-1442.
- Meiran, N., Cole, M. W., & Braver, T. S. (2012). When planning results in loss of control: intention-based reflexivity and working-memory. *Frontiers in Human Neuroscience*, *6*:104.
- Meiran, N., Pereg, M., Kessler, Y., Cole, M. W., & Braver, T. S. (2015a). The power of instructions: Proactive configuration of stimulus-response translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 768.
- Meiran, N., & Cohen-Kdoshay, O. (2012). Working memory load but not multitasking eliminates the prepared reflex: Further evidence from the adapted flanker paradigm. *Acta psychologica*, *139*, 309-313.
- Meiran, N., Pereg, M., Kessler, Y., Cole, M. W., & Braver, T. S. (2015b). Reflexive activation of newly instructed stimulus-response rules: evidence from lateralized readiness potentials in no-go trials. *Cognitive, Affective, & Behavioral Neuroscience*, *15*, 365-373.
- Meiran, N., & Kessler, Y. (2008). The task rule congruency effect in task switching reflects activated long-term. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 137-157.
- Monsell, S. (2003). Task switching. *Trends in cognitive sciences*, *7*, 134-140.
- Morey, R.D., Rouder, J.N., Jamil, T. (2015). "BayesFactor".
<http://bayesfactorpcl.r-forge.r-project.org/>
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 411-421.
- Oberauer, K., Demmrich, A., Mayr, U., & Kliegl, R. (2001). Dissociating retention and

- access in working memory: An age-comparative study of mental arithmetic. *Memory & Cognition*, *29*, 18-33.
- Pfeuffer, C.U., Moutsopoulou, K., Pfister, R., Waszak, F., & Kiesel, A. (2017). The Power of Words: On item-specific stimulus-response associations formed in the absence of action. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 328-347.
- Pfeuffer, C.U., Moutsopoulou, K., Waszak, F., & Kiesel, A. (in press-a). Multiple Priming Instances Increase the Impact of Practice-based but not Verbal Code-based Stimulus-Response Associations. *Acta Psychologica*.
- Pfeuffer, C.U., Hosp, T., Kimmig, E., Moutsopoulou, K., Waszak, F., & Kiesel, A. (in press-b). Defining Stimulus Representation in Stimulus-Response Associations Formed on the Basis of Task Execution and Verbal Codes. *Psychological Research*.
- Ruge, H., & Wolfensteller, U. (2010). Rapid formation of pragmatic rule representations in the human brain during instruction-based learning. *Cerebral Cortex*, *20*, 1656-1667.
- Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, *199*, 159-187.
- Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174-215.
- Stevens, M., Lammertyn, J., Verbruggen, F., & Vandierendonck, A. (2006). Tscope: AC library for programming cognitive experiments on the MS Windows platform. *Behavior Research Methods*, *38*, 280-286.
- Theeuwes, M., De Houwer, J., Eder, A., & Liefoghe, B. (2015). Congruency effects on the basis of instructed response-effect contingencies. *Acta Psychologica*, *158*, 43-50.
- Theeuwes, M., Liefoghe, B., De Schryver, M., & De Houwer, J. (in press). The role of motor imagery in learning via instruction. *Acta Psychologica*.
- Tibboel, H., Liefoghe, B., & De Houwer, J. (2016). Attention to future actions: the

- influence of instructed S-R versus S-S mappings on attentional control. *Psychological Research*, *80*, 905-911.
- Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods*, *49*, 653-673.
- Vandierendonck, A., Liefoghe, B. , & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference. *Psychological Bulletin*, *136*, 601-626.
- van Dijck, J.-P. & Fias, W. (2011). A working memory account for spatial-numerical associations. *Cognition*, *119*, 114-119.
- van Dijck, J.-P., Abrahamse, E.L., Acar, F., Ketels, B., & Fias, W. (2014). A working memory account of the interaction between numbers and spatial attention. *The Quarterly Journal of Experimental Psychology*, *6*, 1500-1513.
- van Dijck, J.-P., Abrahamse, E.L., Majerus, S., & Fias, W. (2013). Spatial attention interacts with serial-order retrieval from verbal working memory. *Psychological Science*, *24*, 1854-1859.
- Wenke, D., De Houwer, J., De Winne, J., & Liefoghe, B. (2015). Learning through instructions vs. learning through practice: Flanker congruency effects from instructed and applied S-R mappings. *Psychological Research*, *79*, 899-912.
- Wenke, D., Gaschler, R., & Nattkemper, D. (2007). Instruction-induced feature binding. *Psychological research*, *71*, 92-106.
- Wenke, D., Gaschler, R., Nattkemper, D., & Frensch, P. A. (2009). Strategic influences on implementing instructions for future actions. *Psychological Research*, *73*, 587-601.
- Yamaguchi, M., & Proctor, R. W. (2011). Automaticity without extensive training: The role of memory retrieval in implementation of task-defined rules. *Psychonomic Bulletin & Review*, *18*, 347-354.
- Yang T., Gathercole S.E., & Allen, R.J. (2014). Benefit of enactment over oral repetition of verbal instruction does not require additional working memory during encoding.

Psychonomic Bulletin & Review, 21, 186-192.

Figure Caption

Figure 1. Schematic overview of the different types of runs that were used in Experiments 1-4.

Figure

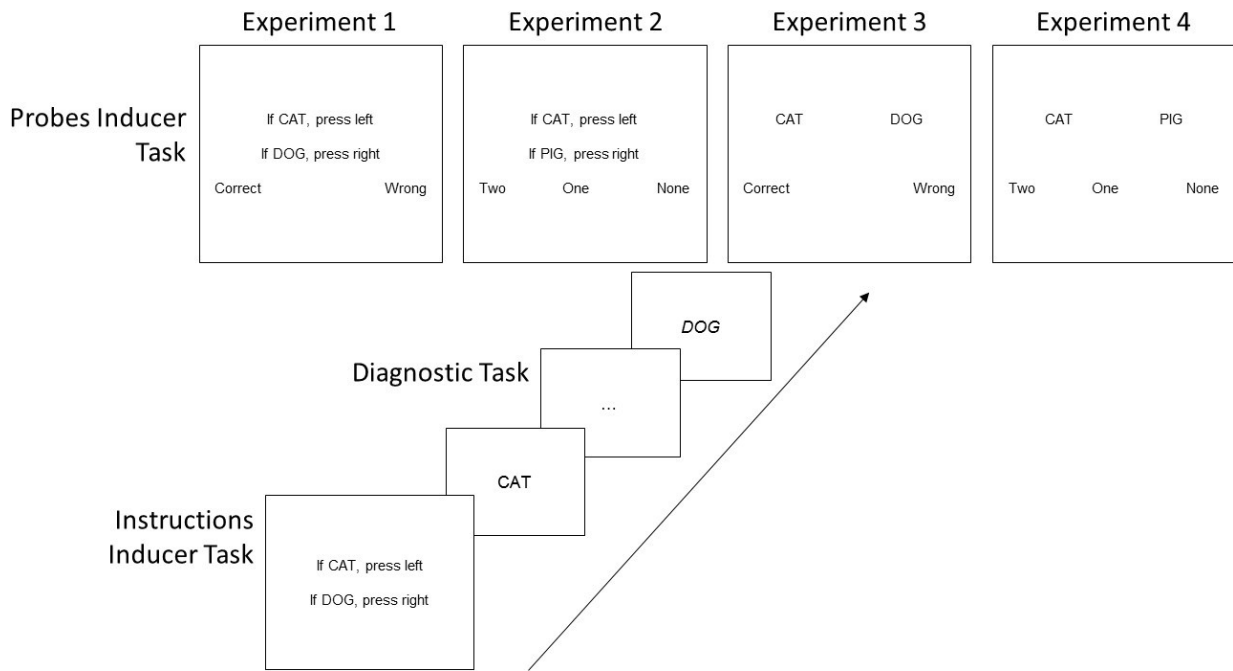


Figure 1