Special Issue Paper

QJEP

Quarterly Journal of Experimental Psychology
1–20

DOI: 10.1177/17470218251360669
qjep.sagepub.com
Sage

Are the automatic effects of instructions modulated by instructed, context-specific strategic control?

Cai S. Longman[1], Christina U. Pfeuffer[2] and Andrea Kiesel[3]

Abstract

Automatic effects of instructions (AEIs) are typically reported with simple instructions that specify stimulus-response (S-R) mappings. Evidence in support of AEIs for instructions that specify more complex rules is less consistent. Here, we investigated whether instructions communicating context-specific strategic control routines designed to reduce evidence accumulation from the target stimulus can affect AEIs in the NEXT paradigm (Meiran et al., 2015). In each mini-block, participants were first instructed on the S-R mappings that should be implemented in the two GO trials at the end of the mini-block. In the intervening NEXT trials (0–3 trials), participants responded to each stimulus with the same (e.g. left) NEXT response. Importantly, the instructions also indicated the probability that each stimulus would be presented during the critical GO trials (e.g. 90% vs. 10%). We reasoned that this would strategically reduce the amount of information accumulated from the target stimulus prior to selecting a response, thereby reducing the magnitude of AEIs. GO performance was modulated by the context, suggesting that the strategic aspects of the instructions had been implemented. However, AEIs were broadly consistent across contexts, suggesting that the adopted strategy did not affect automatic behaviour. This pattern of results was consistent across three experiments (one preregistered), suggesting that complex instructions do not automatically trigger strategic control in dynamic environments.

Keywords

Learning; transfer; instructions; abstract representation; proactive control

Received: 21 August 2023; revised: 11 June 2025; accepted: 12 June 2025

Introduction

The assimilation and subsequent (automatic) implementation of instructions has received a lot of interest recently (for reviews, see e.g. Brass et al., 2017; Meiran et al., 2017). The most consistent finding from this body of research is that simple instructions, such as those communicating stimulus-response (S-R) mappings, are readily assimilated and (automatically) implemented, thereby modulating performance in an irrelevant task that shares the same stimuli and responses as the instructed task – automatic effects of instructions (AEIs; e.g. Liefooghe et al., 2012; Meiran et al., 2015). However, instructions can also communicate more complex rules, such as guidance on strategic control routines. For example, strategic control can be implemented by biasing preparation for a particular response or by reducing the amount of evidence accumulated prior to making a response. The central aim of the present study was to determine whether the implementation of such strategic control instructions would modulate AEIs as well as voluntary performance.

Researchers investigating AEIs tend to use one of two experimental paradigms – the NEXT paradigm (first introduced by Meiran et al., 2015; see also: Pereg et al., 2021; Pereg & Meiran, 2019) or the Inducer-Diagnostic paradigm (first introduced by Liefooghe et al., 2012; see also: Liefooghe et al., 2013; Liefooghe & De Houwer, 2018; Theeuwes et al., 2015). Although there are qualitative

[1]Division of Psychology and Social Work, University of the West of Scotland, Paisley, UK
[2]Faculty of Philosophy and Education, Catholic University of Eichstätt-Ingolstadt, Eichstätt, Germany
[3]Department of Psychology, University of Freiburg, Freiburg, Germany

Corresponding author:
Cai S. Longman, Psychology, School of Education and Social Sciences, University of the West of Scotland, High Street, Paisley PA1 2BE, UK.
Email: cai.longman@uws.ac.uk

differences between these paradigms, they broadly follow the same structure whereby the experimental session is divided into several mini-blocks, each of which includes three phases – instructions, NEXT/diagnostic and GO/inducer phases (for a systematic comparison, see Amir et al., 2022). In the first phase of each mini-block of the NEXT paradigm (used in the present study), participants receive instructions describing the rules of a unique two-choice S-R learning task that should be maintained for implementation in the GO phase (e.g. X → left-hand key press, Y → right-hand key press). On each trial in the intervening NEXT phase, one of the two stimuli is displayed on the screen, and the participant simply has to press one of the two response keys (e.g. the left-hand key) to move through the trials. The same response is used to move through the NEXT trials throughout the entire session (the NEXT response). Thus, no response selection is required during the NEXT phase. On some trials in the NEXT phase, the required response is compatible with the maintained instructions (e.g. stimulus X requires a left-hand key press in the NEXT and GO phases), and on some trials, it is not (e.g. stimulus Y requires a left-hand key press in the NEXT phase but a right-hand key press in the GO phase). The NEXT phase is usually quite short (e.g. 0–3 trials) and can include zero trials to ensure that participants prepare the instructions for immediate implementation. On each trial in the final GO phase, one of the two stimuli is displayed on the screen and the participant is required to implement the instructed task rules. In the original study (Meiran et al., 2015), the current phase was indicated by the colour of the stimulus (e.g. red = NEXT, green = GO). The central difference between the NEXT and Inducer-Diagnostic paradigm is that the diagnostic task in the latter is a two-choice task that requires response selection (e.g. judging whether the printed letter is upright or italic). Nonetheless, like the NEXT phase, the identity of the stimulus is irrelevant during the diagnostic task. The inducer task is equivalent to the GO phase in that participants are required to implement the maintained instructions that were displayed at the start of the mini-block.

The typical finding from NEXT experiments (e.g. Meiran et al., 2015; and Inducer-Diagnostic experiments: e.g. Liefooghe et al., 2012) is that participants tend to respond faster to compatible stimuli (the NEXT response is the same as the instructed GO response) than to incompatible stimuli (the NEXT response and the instructed GO response differ) during the NEXT phase, where no response selection is required – the NEXT compatibility effect. This effect is even robust on the very first trial in the NEXT phase, where participants have not been able to practice the instructed S-R mappings prior to responding. That is, even when the intervening task does not require the stimulus to be identified, performance is affected by merely instructed S-R rules. This is commonly interpreted as evidence that the maintenance of simple instructions for future implementation (e.g. in Working Memory) can result in the formation of S-R associations (or category-response associations; Longman et al., 2019) that automatically trigger, or at least prime a response on detection of the instructed stimulus, even in an irrelevant task where the stimulus need not be identified. Although AEIs have been reported many times using different paradigms, less is known about the degree to which they can be modulated by strategic control.

In their investigations, Liefooghe et al. (2012) and Meiran et al. (2015) found that a higher degree of GO/inducer task preparation was associated with larger compatibility effects in the NEXT/diagnostic task. These authors concluded that AEIs were only observed when participants prepared the GO/inducer task for subsequent performance, and that the degree of GO/inducer task preparation could predict the magnitude of AEIs. For example, in their Experiment 4, Meiran and colleagues found that the NEXT compatibility effect was largest when the GO task was maintained with a high degree of preparedness as indexed by faster RTs in the first trial of the GO phase and a reduced speeding of RTs through the GO phase (which lasted 2 or 10 trials). That is, Meiran and colleagues, like Liefooghe and colleagues, suggested that it might be possible to affect the magnitude of AEIs by strategically modulating the degree to which the task rules were actively prepared for subsequent implementation. However, strategic control of global task preparation might not be the only way to affect the magnitude of AEIs. It might be possible to strategically control the degree to which evidence is accumulated from the stimulus prior to selecting a response, and it might be possible to instruct participants to only implement the instructed task rules in a specific context.

In another prior study, Braem et al. (2017) used the Inducer-Diagnostic paradigm to investigate whether AEIs could be limited to a particular context if the participant was informed of the conditions under which the instructed S-R rules should/should not be applied. In their Experiment 2, Braem and colleagues told participants that the instructed S-R rules should only be applied when the stimulus was displayed at a specific location on the screen (above or below the central fixation – the instructed context rules) during inducer task trials (equivalent to the GO phase). Although participants did implement the instructed context rules and only performed the inducer task when the stimulus was displayed at the relevant location, they found little evidence that the context affected the magnitude of AEIs in the diagnostic task (equivalent to the NEXT phase). That is, they found AEIs of a similar magnitude in the diagnostic task irrespective of whether the stimulus was displayed at the location where the S-R rules should be implemented or where they should be withheld. They concluded that the two aspects of the instructions (the S-R rules and the context rules) were represented separately. The instructed S-R rules were implemented automatically in the irrelevant task, suggesting that the stimulus was processed to a degree that automatically triggered the instructed response, and this effect was not sensitive to the instructed context. Conversely, the instructed context rules

were only implemented in the inducer task, where they would be useful in optimising performance.

Nonetheless, it is still not clear why the implementation of complex instructions communicating context-specific strategic control routines only affected performance in the GO/inducer task and did not affect AEIs in the NEXT/diagnostic task. Likewise, it is still not known whether a strategy intended to reduce the accumulation of information from the stimulus prior to selecting a response would reduce the magnitude of AEIs in a similar way to reducing global task preparation. The aim of the present study was to address this shortfall. Specifically, our aim was to determine whether instructions communicating a strategy designed to reduce the amount of information accumulated from the stimulus in certain contexts would affect the magnitude of AEIs as well as performance in the GO/inducer task. AEIs are commonly interpreted as evidence that simple instructions maintained for subsequent implementation can automatically trigger the instructed response on detection of the instructed stimulus. We therefore reasoned that encouraging a strategy that allowed participants to reduce processing of the features of the stimulus used to identify it might reduce the likelihood that the stimulus would automatically trigger/prime the instructed response, thereby reducing the magnitude of AEIs.

To this end, we manipulated the probability with which each stimulus would appear in the GO phase of an experiment using the NEXT paradigm[1]. Importantly, this contextual information was included in the mini-block instructions, and their relevance for optimising GO performance was emphasised. In Experiments 2 and 3, we also included highly salient cues to indicate the current context throughout a mini-block in an attempt to maximise the likelihood that the contextual aspects of the instructions would be recalled more easily or would be (automatically) triggered by the context cues. Finally, in Experiment 3, we directly instructed participants to use a strategy that relied less on identifying the current stimulus and more on identifying the current phase when selecting an appropriate response. That is, we encouraged participants to optimise their performance in some contexts by minimising the amount of information accumulated about the identity of the target stimulus prior to selecting a response.

Experiment 1

In addition to the simple two-choice S-R task instructions displayed at the start of each mini-block, participants were also informed of the probability that each stimulus would appear in the GO phase (i.e. the probability that the NEXT response, as compared to the alternative response, would be the required response in the GO phase). There were three possible contexts: the NEXT response would be the required response on 90%, 50% or 10% of trials in the GO phase. That is, the contextual aspect of the instructions provided some useful information that could be implemented to strategically optimise performance in the GO phase of each mini-block. For example, in mini-blocks where the NEXT response was the required response on 90% of GO phase trials, participants could simply press the NEXT key throughout the mini-block without necessarily needing to identify the target stimulus at all. Similarly, in mini-blocks where the NEXT response was the required response on 10% of GO phase trials, participants could make a NEXT response during the NEXT phase and switch to the alternative response during the GO phase. That is, if participants used the strategic instructions, the only information required to respond with an acceptable degree of accuracy in these mini-blocks would be the identity of the current phase and not necessarily the identity of the target stimulus. Conversely, in mini-blocks where the NEXT response was the required response on 50% of GO phase trials, it was not possible to select the correct response without first identifying the target stimulus (as well as the current phase).

Thus, if response times (RTs) in the GO phase were slowest when the required response was unpredictable and fastest when the required response was highly predictable, we could conclude that participants had utilised the contextual information communicated in the instructions to optimise performance in the GO phase – that is, the implementation of the contextual instructions had modulated voluntary performance. Conversely, if RTs in the GO phase were equivalent for all contexts, then we would have to concede that participants had not utilised the contextual information from the instructions to adopt the response selection strategy described above.

Similarly, if the magnitude of the NEXT compatibility effect was reduced in mini-blocks where the required response in the GO phase was highly predictable, we could conclude that participants had utilised the contextual information communicated in the instructions to adopt a response selection strategy emphasising the identification of the current phase rather than the current stimulus – that is, the implementation of the contextual instructions had modulated involuntary performance (the automatic triggering of a planned response on detection of the instructed stimulus) by reducing the amount of evidence accumulated about the target stimulus prior to selecting a response). Conversely, if the NEXT compatibility effect was equivalent for all contexts, then we would have to concede that whether participants had adopted such a response selection strategy or not, AEIs had not been affected by the contextual aspects of the instructions.

Importantly, any evidence that the contextual aspects of the instructions had modulated *either* GO performance *or* the NEXT compatibility effect without modulating the other would suggest that strategic control of response selection criteria based on contextual information communicated via instructions only affects voluntary/involuntary performance, respectively.

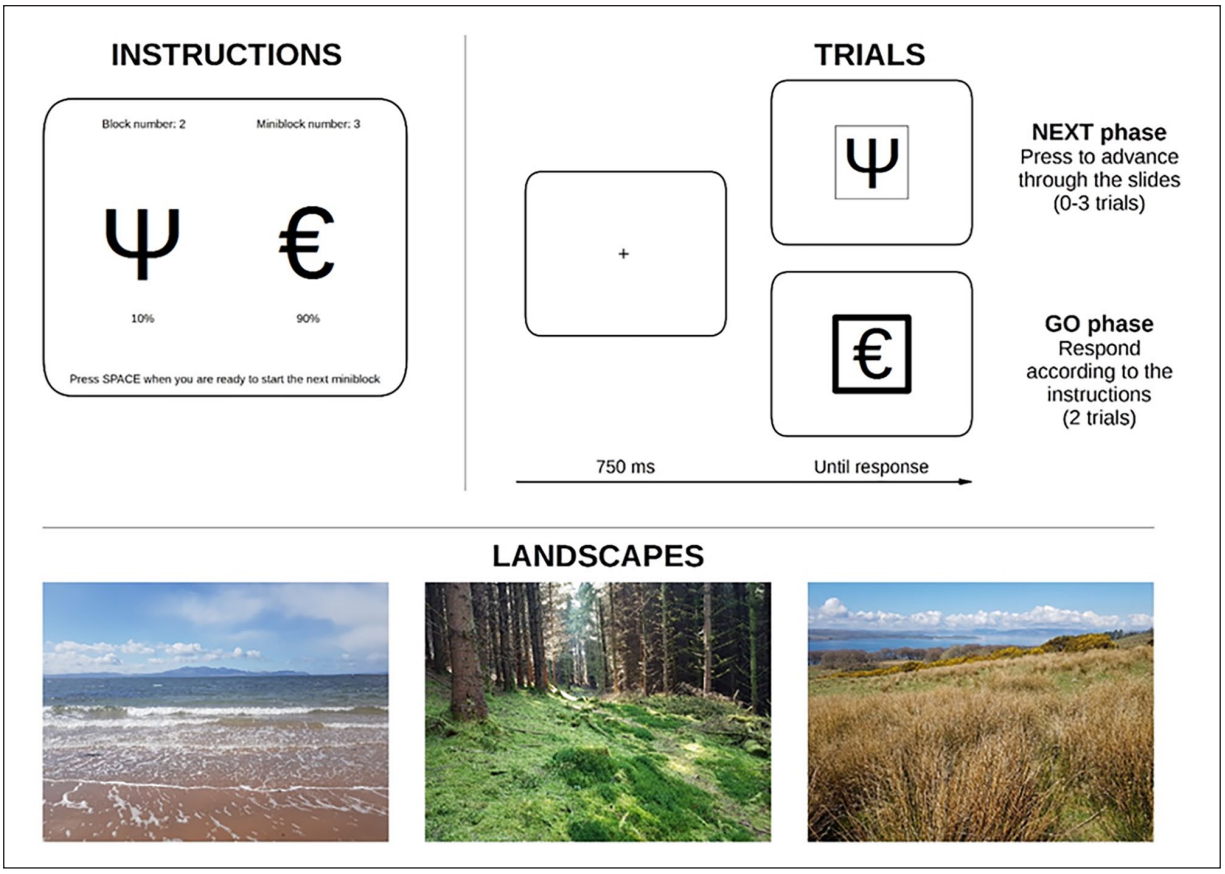

Figure 1. Mini-block instructions (left), timeline of a single trial in the NEXT and GO phases of each mini-block (right), and images of the three landscapes used in Experiments 2 and 3 (lower).

Method

Participants. The final sample after participant exclusions (see 'Analyses' section) included a total of 60 individuals (23 female, 36 male, 0 diverse, 1 prefer not to say; mean age = 26.1 years, *SD* = 9.1; 57 right-hand dominant, 3 left-hand dominant) who were recruited via 'Prolific' (an online participant pool) and were paid £7.50 for a session that lasted ~60 min. Participants identified as being fluent English speakers and at least 18 years of age and were only able to participate via a desktop or laptop PC (not a tablet or smartphone). The target sample size and exclusion criteria were decided in advance of data collection. The mean Hedges' *g* for the t-tests comparing performance on compatible versus incompatible NEXT trials for the three dependent variables in Experiment 2 of Longman et al. (2019), which used a very similar design and identical stimuli, was 0.599. Because the critical comparison in the present experiment was a modulation of this effect, we selected an N that would allow us to detect effects of roughly half the size (i.e. $g = 0.3$) at 80% power given an alpha of .05. For all experiments in the present study, we obtained approval by the local research ethics committee at the School of Education and Social Sciences, University of the West of Scotland, and participants provided informed consent after the nature and possible consequences of the studies were explained to them.

Materials and procedure. The experiment was coded using JavaScript (jsPsych; De Leeuw, 2015) and was hosted on a dedicated server at the University of Freiburg, Germany. The stimuli were black letters, numbers, and symbols selected from readily available fonts (e.g. Wingdings, Webdings, Hebrew alphabet) presented on a white background (the same stimuli were used in Longman et al.'s, 2019, Experiment 2 that used a similar design and are available to download from the Open Science Framework repository along with the jsPsych scripts used for data collection, the raw data files, and the data analysis scripts from all experiments at: https://osf.io/eudxa/). The experimental session was divided into mini-blocks, each of which included three phases (see Figure 1). Two unique stimuli were pseudorandomly selected from the entire set of 600 to be used in each mini-block. There were no repetitions of stimuli between mini-blocks.

Each mini-block started with an initial instructions phase where the two stimuli used in the remainder of the mini-block were displayed to the left/right of the centre of

the screen (one on each side). The location of each stimulus indicated the correct response for that stimulus in the GO phase (left vs. right button press – 'F' vs. 'J' key, respectively). The probability that each stimulus would be presented in the GO phase (i.e. the current context) was displayed below the relevant stimulus (see Figure 1). The probabilities were one of three possible conditions: 90% left-hand response – 10% right-hand response, 10% left-hand response – 90% right-hand response, and 50% left-hand response – 50% right-hand response as a control (pseudorandomised for each mini-block so that each condition was equiprobable throughout the session). The instructions remained visible until the participant pressed the space bar to initiate the mini-block. The mini-block then started with 'Ready. . .' presented centrally for 1000 ms before the start of the first trial. The NEXT phase had a duration of 0, 1, 2 or 3 trials. The proportion of each duration was 16.66%, 33.33%, 33.33% and 16.66%, respectively, to partially control for participants anticipating the start of the GO phase based on the number of NEXT trials performed and to encourage preparation to respond according to the GO rules from the outset (see Longman et al., 2019; Verbruggen et al., 2018). During the NEXT phase, the stimulus was presented centrally surrounded by a thin black square (line thickness = 3 pt) to indicate that the identity of the stimulus was irrelevant to the required response. To balance the number of compatible/incompatible stimuli presented on the first, second and third NEXT trial the stimulus was pseudorandomly selected to ensure that the probability of each stimulus being displayed on each NEXT trial was 50%. (Note that these counterbalancing measures were performed separately for each context before randomising the order of mini-blocks through the session). The participant was always required to press one of the response keys (counterbalanced across participants) to move through the NEXT phase trials – the NEXT response. The GO phase started immediately after the NEXT phase ended and had a duration of two trials only. In the GO phase, stimuli were pseudo-randomly selected from the two stimuli used in the current mini-block to ensure that, across the entire experiment, stimuli appeared according to the instructed probabilities for the mini-block (i.e. instructed = actual probabilities). The GO phase was identical to the NEXT phase except that the black square surrounding the stimulus was bold (line thickness = 15 pt) to indicate that the stimulus should be identified according to the instructed rules.

In both the NEXT and GO phases, the trial started with a fixation cross presented centrally for 750 ms. During the NEXT phase, the stimulus remained visible until the correct response was made (the total number of responses was recorded for analysis). During the GO phase, the stimulus remained visible until any response was made. No immediate feedback was given.

The experimental session consisted of 24 blocks, each consisting of 12 mini-blocks (a total of 96 mini-blocks per context consisting of 2–5 trials each; see above). The maximum number of observations available for analysis in each context was 40 compatible and 40 incompatible stimuli presented on the first NEXT trial. Note that performance from the second and third NEXT trials was not analysed because after the first NEXT trial, performance might be modulated by practice effects, whereas performance on the first NEXT trial can only be modulated by the instructions. GO performance (mean RT and proportion of errors, PE) for the block was presented at the end of each block until the space bar was pressed to initiate the next block. Prior to the experimental phase, there was a brief familiarisation phase consisting of 12 mini-blocks identical to the mini-blocks in the experimental phase: four mini-blocks for each context, three mini-blocks for each number of NEXT trials (0, 1, 2 or 3). Like in the experimental phase, each mini-block used unique stimuli, and the order of mini-blocks was pseudorandomised. The data from the familiarisation phase were not analysed.

In each block, there was one 'catch trial' used to ensure that participants retained the instructed probabilities of the required response in the GO phase throughout each mini-block. The catch trial could occur at the end of any mini-block within a block (pseudorandomised to ensure that the catch trial occurred equally often on each mini-block across the experimental phase). On catch trials, after the last GO response had been made for the mini-block, the two relevant stimuli were once again displayed on the left/right of the screen (as in the instructions phase for the mini-block). Participants were required to indicate the correct instructed probability that each stimulus would appear in the GO phase (i.e. the correct context) by pressing 1, 2 or 3 on a standard keyboard. Context-key mappings were displayed on screen. Catch trial accuracy was also noted in the block-end feedback.

Analyses. All data processing and analyses were performed using R (R Core Team, 2022). The critical comparison was between NEXT performance on compatible trials (stimuli where the correct response during the GO phase was the same as the required response during the NEXT phase) versus incompatible trials (stimuli where the correct response during the GO phase was different from the required response during the NEXT phase). Consistent with Longman et al. (2019), we focused only on the first NEXT trial because AEIs are largest on this trial (Meiran et al., 2015), and performance on later NEXT trials could already be modulated by practice effects. Note that this decision was made prior to data collection, but it was important to include the remaining NEXT trials in the experiment to guarantee that the duration of the NEXT phase was unpredictable. This manipulation increased the likelihood that the instructions were maintained in a state

that allows for their rapid implementation as soon as they became relevant.

Mini-blocks where an incorrect response was made during the GO phase (7.2% of mini-blocks) were omitted from all NEXT analyses because this could indicate the instructions were not processed properly. Trials with RTs <100 ms or >3000 ms were also omitted from all analyses (0.4%). Any participants for whom the above data cleaning procedures resulted in <20 (50% of) observations per cell (Context by Compatibility) or who achieved <75% accuracy on the catch trials were replaced (15 participants in total).

Consistent with Longman et al. (2019), we analysed three dependent variables. The primary dependent variable was the latency of the NEXT response (NEXT RT; equivalent to the analyses performed by Meiran et al., 2015). Because participants must make a NEXT response (see above) to progress through the NEXT phase, it is possible that NEXT RTs are slower on incompatible trials because participants press the incorrect key (i.e. the key associated with the presented stimulus in the instructed task) before pressing the NEXT key. We therefore also analysed the latency of the NEXT responses limited to those trials on which the first response was correct (Correct NEXT RT). Finally, we also analysed the proportion of errors (first given response) made on NEXT trials (NEXT PE). However, given that the error rates were quite low (mean NEXT PEs were <5% in all conditions), and the pattern of results was largely consistent between NEXT RTs and Correct NEXT RTs, we report only the analyses for NEXT RTs here. All equivalent analyses for the Correct NEXT RTs and NEXT PEs can be found in the Supplemental Materials.

A separate Context (probability that the NEXT response would be required in the GO phase for the mini-block: 90%, 50%, 10%) by Compatibility (compatible, incompatible), by Experiment half (first half, second half) repeated measures ANOVA was conducted for each dependent variable. Experiment half was included as a factor in the analyses to determine whether any of the effects were modulated by practice (and were therefore not exclusively the result of instructions). Bayes factors and effect sizes (generalised eta squared) were also calculated for all relevant effects/interactions. Bayes factors were calculated with the BayesFactor package, using the default JZS prior (.707; Morey et al., 2015). To reduce the number of model comparisons, a subtraction approach was used (see Morey et al., 2015). That is, the full model, including all effects and interactions, was compared to the full model minus the effect or interaction of interest. When this approach is used, Bayes factors <1 indicate that removing the effect/interaction reduces the fit of the model (i.e. the effect/interaction is an important contributor to the fit of the full model). Conversely, when the Bayes factor remains >1, removing the effect/interaction does not affect the fit of the model (i.e. the effect/interaction does not contribute to the fit of the model). For all Bayesian analyses, only the BF_{10} is reported (i.e. the Bayes Factor in favour of the alternative hypothesis) and are interpreted using the classifications introduced by Schönbrodt and Wagenmakers (2017).

Performance data from the GO trials were also analysed to determine whether the contextual aspects of the instructions had been implemented. As with the data from the NEXT trials, trials with RTs <100 ms or >3000 ms were omitted from all analyses, resulting in 0.25% data loss. Error trials were also omitted from the GO RT analysis. The GO RTs and proportion of errors on GO trials (GO PE) were submitted to Context by Experiment half-repeated measures ANOVAs to determine whether the contextual aspects of the instructions modulated performance in the GO phase, suggesting context-specific strategic control of voluntary performance, and whether any observed differences between Context conditions could be explained by a practice effect.

Finally, to determine whether modulations in GO performance resulting from the strategic implementation of the contextual aspects of the instructions affected the magnitude of AEIs in the NEXT phase, we performed separate Pearson correlations for each dependent variable comparing the 'Context effect' from the GO and NEXT phases over participants[2]. The Context effect was calculated by subtracting the GO performance/NEXT compatibility effect for mini-blocks where the required response in the GO phase was the NEXT response on 90%/10% of trials from the GO performance/NEXT compatibility effect for mini-blocks where the required response in the GO phase was the NEXT response on 50% of trials. We also calculated equivalent Bayes Factors for each correlation.

Results and discussion

Performance data from the NEXT and GO phases of all experiments are plotted in Figure 2. Performance data from the NEXT phase are plotted as a function of Context, Compatibility and Experiment half. Performance data from the GO phase are plotted as a function of Context and Experiment half.

GO phase. As predicted, participants responded fastest and made the fewest errors in the GO phase when the NEXT response was the required response on 90% of GO trials (mean GO RT = 570 ± 15[3] ms, mean GO PE = 3.5 ± 0.4%). GO performance was worst when the NEXT response was the required response on 50% of GO trials (mean GO RT = 614 ± 15 ms, mean GO PE = 4.9 ± 0.5%), and was intermediate when the NEXT response was the required response on 10% of GO trials (mean GO RT = 600 ± 17 ms, mean GO PE = 4.3 ± 0.4%). The repeated measures ANOVAs on these data confirmed that GO performance was

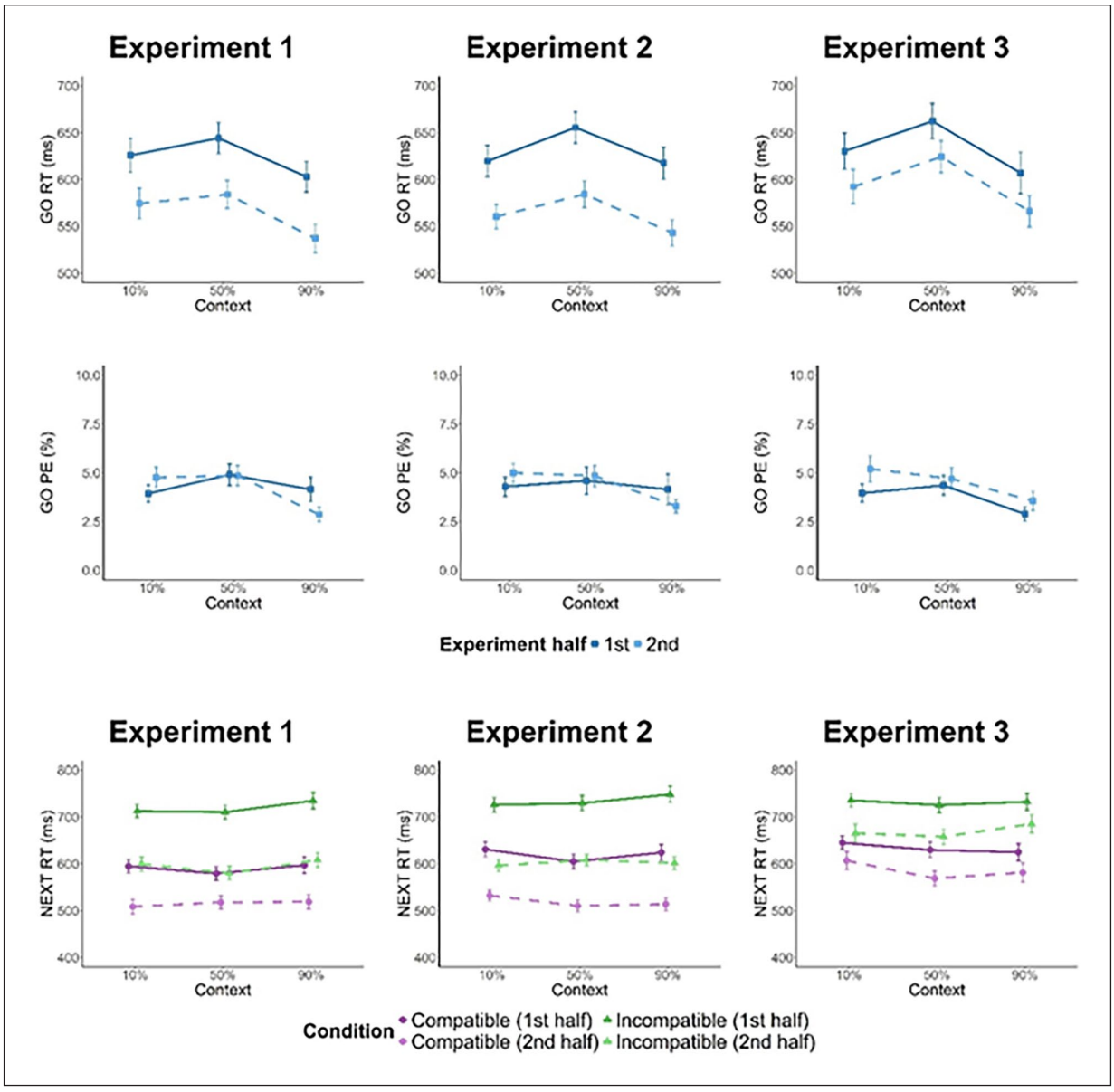

Figure 2. GO RTs, GO PEs and NEXT RTs from all experiments. Performance data from GO phase trials are plotted as a function of context (Probability that the NEXT response was the required response in the GO phase), and experiment half. NEXT RTs are plotted as a function of context, compatibility and experiment half.

Note. Error bars for GO phase data show the standard error of the mean for each context/experiment half. Error bars for NEXT RTs show the standard error of the mean difference between compatible and incompatible trials in each context/experiment half.

significantly modulated by the contextual aspects of the instructions (main effect of Context: GO RT: $F(2,118)=36.35$, $p<.001$, $BF<0.001\pm2.4\%$, $\eta^2=.148$; GO PE: $F(2,118)=8.28$, $p<.001$, $BF=0.030\pm2.7\%$, $\eta^2=.050$) suggesting that this information was used to strategically control performance in the GO phase.

Participants also made slower responses in the GO phase in the first half of the session (mean GO RT = 624 ± 17 ms) relative to the second half of the session (mean GO RT = 565 ± 15 ms; main effect of Experiment half: $F(1,59)=44.78$, $p<.001$, $BF<0.001\pm2.2\%$, $\eta^2=.311$). But response accuracy was comparable across the session (mean GO PE: first half = 4.3 ± 0.5%, second half = 4.2 ± 0.5%; main effect of Experiment half: $F(1,59)=0.24$, $p=.625$, $BF=7.493\pm2.5\%$, $\eta^2=.001$). The Context by Experiment half interaction was non-significant for GO RTs ($F(2,118)=2.16$, $p=.119$, $BF=10.134\pm4.8\%$, $\eta^2=.005$), but did reach significance for GO PEs ($F(2,118)=5.04$, $p=.008$, $BF=0.366\pm2.8\%$, $\eta^2=.030$) suggesting that the implementation of the contextual aspects of the instructions in the GO phase was modulated to a degree by practice in the GO PEs, but not the GO RTs.

Table 1. Omnibus repeated measures ANOVA results for NEXT RTs from Experiment 1.

Effect	NEXT RT					
	DF	*MSE*	*F*	*p*	*BF* (%)	η^2
Context	(2,118)	4,008.48	4.86	.009	4.193 ± 5.3	.009
Compatibility	(1,59)	24,465.73	80.99	<.001	<0.001 ± 8.0	.307
Experiment half	(1,59)	21,278.99	83.34	<.001	<0.001 ± 5.3	.284
Context × Compatibility	(2,118)	2,936.05	1.39	.253	19.150 ± 5.5	.002
Context × Experiment half	(2,118)	2,864.11	0.22	.804	37.796 ± 15.2	.000
Compatibility × Experiment half	(1,59)	4,109.52	24.94	<.001	0.007 ± 10.4	.022
Context × Compatibility × Experiment half	(2,118)	3,140.16	2.13	.123	8.012 ± 6.3	.003

Note. Equivalent Bayes factors are also reported. Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus, Bayes factors <1 indicate that the effect/interaction is an important contributor to the model.

Table 2. Pearson correlation results comparing context effect for NEXT RTs to GO performance from Experiment 1.

Context effect	NEXT RT vs. GO RT					NEXT RT vs. GO PE				
	r	*DF*	*T*	*p*	*BF*	*r*	*DF*	*t*	*p*	*BF*
50%–90%	−.111	58	−0.850	.399	0.406	.203	58	1.580	.120	0.889
50%–10%	−.214	58	−1.666	.101	1.004	.070	58	0.537	.593	0.333

Note. Equivalent Bayes factors are also reported. See main text for definition of context effect.

NEXT phase. The results from the repeated measures ANOVA on NEXT RTs from Experiment 1 are reported in Table 1.

Consistent with previous findings, the ANOVA on the NEXT RTs found that participants responded faster on compatible (mean RT = 553 ± 16 ms) relative to incompatible NEXT trials (mean RT = 658 ± 23 ms; main effect of Compatibility: $p < .001$, $BF < 0.001 \pm 8.0\%$) – a NEXT compatibility effect (i.e. an AEI). There was also some evidence that NEXT RTs were affected by the context instructions (mean NEXT RT when the NEXT response was the required response on 90% of GO trials: 615 ± 22 ms; 50% of GO trials: 597 ± 20 ms; 10% of GO trials: 605 ± 20 ms; main effect of Context: $p = .009$), but the equivalent Bayesian analysis found evidence that the effect did not materially contribute to the fit of the model ($BF = 4.193 \pm 5.3\%$) suggesting that the effect was equivocal at best. Nonetheless, the critical Context by Compatibility interaction was not significant ($p = .253$) and the Bayesian analysis provided strong evidence that removal of the interaction from the full model would not materially affect its fit ($BF = 19.150 \pm 5.5\%$) indicating that the contextual aspects of the instructions that were implemented to strategically modulate (voluntary) GO performance did not affect the magnitude of the (automatic) NEXT compatibility effect. That is, the evidence suggests that the context-specific strategy implemented by participants resulting in modulations to GO performance did not reduce automatic processing of the identity of the stimulus to a degree that reduced AEIs.

NEXT RTs were significantly slower in the first half of the experimental session (mean RT = 655 ± 23 ms) relative to the second half (mean RT = 556 ± 21 ms; main effect of Experiment half: $p < .001$, $BF < 0.001 \pm 5.3\%$) indicating that NEXT performance improved with practice. The NEXT compatibility effect was also larger in the first half of the session (mean difference = 129 ms) relative to the second half (mean difference = 81 ms; Compatibility by Experiment half interaction: $p < .001$, $BF = 0.007 \pm 10.4\%$), suggesting some influence of practice on the compatibility effect. However, neither the Context by Experiment half interaction ($p = .804$, $BF = 37.796 \pm 15.2\%$) nor the three-way interaction ($p = .123$, $BF = 8.012 \pm 6.3\%$) was significant suggesting that implementation of the contextual aspects of the instructions in the NEXT phase was not influenced by practice.

Correlations. The results of the correlations between the Context effect (see 'Analyses' section) for performance in the GO phase and the NEXT compatibility effect for Experiment 1 are reported in Table 2. None of these correlations was significant or provided better than anecdotal evidence in support of the experimental hypothesis, suggesting that any strategic differences in GO performance between contexts did not transfer to the NEXT phase, resulting in modulations to AEIs. However, it should also be noted that none of the Bayesian analyses found better than anecdotal evidence in support of the null hypothesis, either indicating that no firm conclusions can be drawn from these analyses. Nonetheless, the correlation coefficients were all ⩽.203 (including some negative correlations), suggesting that any relationship between the effect of context on GO/NEXT performance was weak at best.

Summary. Thus, the findings from Experiment 1 indicate that the contextual aspects of the instructions were used to strategically modulate performance in the GO phase. However, this strategy did not materially affect the automatic implementation of instructions in the NEXT phase (AEIs).

Experiment 2

However, the context in Experiment 1 was communicated via the instructions alone, and not by any other exogenous cues such as those used by Braem et al. (2017). Providing rich cues that indicate the current context throughout each mini-block might encourage implementation of the context-specific strategy used to modulate GO performance during the NEXT phase in at least two possible ways. First, rich contextual cues might help participants recall the current context more easily, thereby maintaining the relevant strategy in a state of heightened readiness for implementation. Keeping the strategy in a more active state might result in its implementation during the NEXT phase, thereby modulating AEIs as well as voluntary performance. Second, the contextual cues could encourage participants to blur the boundaries between the NEXT and GO phases in each mini-block. If each mini-block is perceived as a single episode rather than two separate episodes (phases), then the relevant context-specific strategy used to modulate voluntary GO performance might be more likely to be implemented during the NEXT phase, thereby modulating AEIs. Thus, the aim of Experiment 2 was to determine whether the context-specific strategy used by participants in Experiment 1 to modulate GO performance would also modulate AEIs in the NEXT phase if the current context was cued exogenously.

To this end, each context (the probability that the NEXT response would be the required response in the GO phase) was consistently bound to a particular landscape throughout the experimental session – a beach, a forest or a meadow. Thus, in addition to communicating the context via the instructions, the relevant context in each mini-block was also communicated via a background image (see Figure 1) and a corresponding audio soundscape that played throughout the mini-block. We reasoned that enriching the context cues by using realistic images/soundscapes would make them more immersive, thereby maximising the likelihood of observing any effects of the contextual aspects of the instructions on AEIs.

Method

Participants. In the final sample (see 'Analyses' section for exclusions), 60 different individuals (28 female, 31 male, 1 diverse, 0 prefer not to say; mean age = 27.7 years, SD = 8.1; 53 right-hand dominant, 7 left-hand dominant) were recruited via 'Prolific' and were paid at the same rate as participants from Experiment 1. Again, the ethics committee at the School of Education and Social Sciences, University of the West of Scotland, approved the study, and all participants provided informed consent prior to their participation.

Materials and procedure. The materials and procedure for Experiment 2 were identical to Experiment 1, with two important exceptions. First, participants were informed that each context (probability that each stimulus would be displayed in the GO phase) would be consistently bound to a single landscape throughout the session. The three landscapes were a beach, a forest and a meadow, and each was made salient by introducing photographic images for the background (see Figure 1), and corresponding audio soundscapes. The soundscapes were sourced from the BBC Sound Effects Library (https://sound-effects.bbcrewind.co.uk/) and were edited such that each audio file had a duration of 60s. All images and soundscapes can be found on the Open Science Framework data repository (https://osf.io/eudxa/). Second, to ensure that participants were able to hear the soundscapes throughout the session, a 'check' trial was added to the end of each mini-block where the participant had to indicate whether a sound (one of the three soundscapes, pseudorandomly selected) was being played or not by pressing either 'y' or 'n' on the keyboard (a sound was played on 50% of the check trials).

Analyses. As in Experiment 1, mini-blocks where an incorrect response was made during the GO phase (5.8% mini-blocks) were omitted from all NEXT analyses. Trials with RTs <100 ms or >3000 ms were also omitted from all analyses (0.26%). Any participants for whom the above data cleaning procedures resulted in <20 (50% of) observations per cell (Context by Compatibility), or who achieved $<75\%$ accuracy on the catch or check trials, were replaced (16 participants in total). As with the data from the NEXT trials, GO trials with RTs <100 ms or >3000 ms were omitted from all analyses, resulting in 0.31% data loss. Error trials were also omitted from the GO RT analysis.

Experiment 2 used identical analyses to those performed in Experiment 1. However, because the repeated measures ANOVA on NEXT RTs found a significant Context by Compatibility interaction, follow-up paired-sample *t*-tests were conducted to directly compare the magnitude of the NEXT compatibility effect between contexts. Equivalent Bayes factors and effect sizes (Hedges' *g*) are also reported for these analyses.

Results and discussion

GO phase. As in Experiment 1, the ANOVAs on GO performance found that participants responded fastest and made fewest errors in the GO phase of mini-blocks where

Table 3. Omnibus repeated measures ANOVA results for NEXT RTs from Experiment 2.

Effect	*DF*	*MSE*	*F*	*p*	*BF* (%)	η^2
Context	(2,118)	3,964.55	1.46	.237	25.648 ± 6.1	.003
Compatibility	(1,59)	17,046.23	103.08	<.001	<0.001 ± 5.9	.297
Experiment half	(1,59)	15,428.30	159.17	<.001	<0.001 ± 6.3	.372
Context × Compatibility	(2,118)	3,526.61	4.91	.009	2.783 ± 6.5	.008
Context × Experiment half	(2,118)	4,779.88	1.39	.252	12.284 ± 5.9	.003
Compatibility × Experiment half	(1,59)	4,737.40	9.32	.003	0.298 ± 5.8	.011
Context × Compatibility × Experiment half	(2,118)	4,311.71	0.08	.920	16.574 ± 6.0	<.001

Note. Equivalent Bayes factors are also reported. Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus, Bayes factors <1 indicate that the effect/interaction is an important contributor to the model.

the NEXT response was the required response on 90% of GO trials (mean GO RT = 580 ± 13 ms, mean GO PE = 3.7 ± 0.4%). GO performance was worst in mini-blocks where the NEXT response was the required response on 50% of trials (mean GO RT = 620 ± 13 ms, mean GO PE = 4.7 ± 0.5%), and was intermediate on mini-blocks where the NEXT response was the required response on 10% of trials (mean GO RT = 590 ± 13 ms, mean GO PE = 4.6 ± 0.4%). The ANOVAs on these data confirmed that GO performance was significantly modulated by the contextual aspects of the instructions (main effect of Context for GO RT: $F(2,118) = 28.00$, $p < .001$, $BF < 0.001 \pm 3.2\%$, $\eta^2 = .145$; GO PE: $F(2,118) = 3.75$, $p = .026$, $BF = 1.597 \pm 3.8\%$, $\eta^2 = .021$), indicating that the instructions were used to strategically optimise performance in the GO task.

GO RTs were also faster in the second half of the experimental session (mean GO RT = 563 ± 14 ms) relative to the first half (mean GO RT = 631 ± 17 ms; main effect of Experiment half: $F(1,59) = 79.15$, $p < .001$, $BF < 0.001 \pm 2.3\%$, $\eta^2 = 0.411$). However, GO PEs were broadly equivalent in both halves of the session (mean GO PE in the first half = 4.3 ± 0.7%; in the second half = 4.4 ± 0.5%; main effect of Experiment half: $F(1,59) < 0.01$, $p = .949$, $BF = 8.224 \pm 3.5\%$, $\eta^2 < .001$). The Context by Experiment half interaction did not reach significance for either GO RTs ($F(2,118) = 3.06$, $p = .051$, $BF = 8.099 \pm 2.5\%$, $\eta^2 = .006$) or GO PEs ($F(2,118) = 2.26$, $p = .109$, $BF = 4.235 \pm 5.2\%$, $\eta^2 = .011$), indicating that the effect of the contextual aspects of the instructions on GO performance was not modulated by practice.

NEXT phase. The results from the repeated measures ANOVA on NEXT RTs from Experiment 2 are reported in Table 3.

As in Experiment 1, the ANOVA on the NEXT RTs found that participants responded faster on compatible (mean RT = 570 ± 16 ms) relative to incompatible NEXT trials (mean RT = 669 ± 20 ms; main effect of Compatibility: $p < .001$, $BF < 0.001 \pm 5.9\%$). Although there were some small differences in NEXT RTs according to the context (mean NEXT RT when the NEXT response was the required response on 90% of GO phase trials: 621 ± 20 ms; 50% of GO phase trials: 613 ± 20 ms; 10% of GO phase trials: NEXT RT = 622 ± 18 ms), they did not reach significance (main effect of Context: $p = .237$, $BF = 25.648 \pm 6.1\%$), indicating that the contextual aspects of the instructions did not affect absolute performance in the NEXT task. Nonetheless, the critical Context by Compatibility interaction was significant ($p = .009$) indicating that the magnitude of the NEXT compatibility effect was modulated by the contextual aspects of the instructions. However, it should be noted that the equivalent Bayesian analysis found (albeit relatively weak) evidence that removal of the interaction from the model would not materially impair its fit ($BF = 2.783 \pm 6.5\%$). The additional analyses run to unpack this interaction are reported below.

NEXT RTs were significantly slower in the first half of the experimental session (mean RT = 677 ± 22 ms) relative to the second half (mean RT = 560 ± 19 ms; main effect of Experiment half: $p < .001$, $BF < 0.001 \pm 6.3\%$), indicating that NEXT performance improved with practice. The NEXT compatibility effect was also larger in the first half of the session (mean difference = 114 ms) relative to the second half (mean difference = 83 ms; Compatibility by Experiment half interaction: $p = .003$, $BF = 0.298 \pm 5.8\%$), suggesting some influence of practice on the compatibility effect. However, neither the Context by Experiment half interaction ($p = .252$, $BF = 12.284 \pm 5.9\%$) nor the three-way interaction ($p = 0.920$, $BF = 16.574 \pm 6.0\%$) was significant. This suggests that any effect of the contextual aspects of the instructions on AEIs was not influenced by practice.

The results of the additional *t*-tests comparing the magnitude of the NEXT compatibility effect across the different contexts from Experiment 2 are reported in Table 4. The NEXT compatibility effect was significantly smaller in mini-blocks where the NEXT response would be the required response on 10% of GO phase trials (incompatible – compatible difference = 80 ms) relative to mini-blocks where the NEXT response would be the required response on 50% of GO phase trials (incompatible – compatible difference = 112 ms; $p = .002$, $BF = 12.559$) and mini-blocks where the NEXT response would be the

Table 4. Results from *t*-tests comparing the NEXT compatibility effect for NEXT RTs across contexts (Probability that the NEXT response would be the required response in the GO phase) of Experiment 2.

Contrast	Difference	Lower CI	Upper CI	*DF*	*t*	*p*	*BF*	*g*
10% vs. 50%	−32	−52	−12	59	−3.18	**.002**	12.559	0.367
50% vs. 90%	5	−17	27	59	0.48	.634	0.158	0.055
10% vs. 90%	−26	−50	−3	59	−2.28	**.026**	1.530	0.299

Note. Equivalent Bayes factors are also reported. Bayes factors >1 indicate evidence in support of the experimental hypothesis, whereas Bayes factors <1 indicate evidence in support of the null hypothesis.

Table 5. Pearson correlation results comparing context effect for NEXT RTs to GO performance from Experiment 2.

Context effect	NEXT RT vs. GO RT					NEXT RT vs. GO PE				
	r	*DF*	*T*	*p*	*BF*	*r*	*DF*	*t*	*p*	*BF*
50%–90%	−.013	58	−0.102	.919	0.294	.194	58	1.503	.138	0.801
50%–10%	−.124	58	−0.954	.344	0.441	.013	58	0.100	.921	0.294

Note. Equivalent Bayes factors are also reported. See main text for definition of context effect.

required response on 90% of GO phase trials (incompatible – compatible difference = 106 ms; $p = .026$, $BF = 1.530$). However, it should be noted that the Bayesian analysis found only anecdotal evidence in support of the latter difference.

Correlations. The results of the correlations between the Context effect for performance in the GO phase and the NEXT compatibility effect for Experiment 2 are reported in Table 5. As in Experiment 1, there was little evidence that any strategic differences in GO performance between contexts transferred to the NEXT phase resulting in modulations to AEIs.

Summary. Thus, the findings from Experiment 2 confirm that the contextual aspects of the instructions were used to strategically modulate performance in the GO phase. However, contrary to Experiment 1, there was some evidence that this strategy did affect the automatic implementation of instructions in the NEXT phase (AEIs). Specifically, the NEXT compatibility effect was significantly smaller in those mini-blocks where the NEXT response was the required response on 10% of GO phase trials relative to mini-blocks where the NEXT response was the required response on 50% (and 90%) of GO phase trials. This pattern of results supports, to a degree, the notion that participants were relying less on identification of the stimulus to direct response selection in the GO phase, and that this strategy was also being implemented during the NEXT phase resulting in reduced AEIs. However, our initial prediction was that the magnitude of AEIs would be smallest in mini-blocks where the NEXT response was the required response in 90% of GO phase trials, but this was not the observed pattern. Taken alongside the conflicting results from the omnibus ANOVAs (suggesting a significant Context by Compatibility interaction) and the equivalent Bayesian analyses (which found evidence in support of the null hypothesis for this interaction), there remains a degree of doubt over the replicability of these findings and their interpretation.

Combined analysis of Experiments 1–2

To further investigate the stability of the findings from Experiment 2 that seemed to suggest some effect of the contextual aspects of the instructions on the magnitude of AEIs, the NEXT trial data from Experiments 1 and 2 were collapsed together[4]. The resulting data set was analysed using Compatibility by Context by Experiment ANOVAs and equivalent Bayesian analyses to increase the power of the initial analyses and allow a direct comparison of the patterns of results from the two experiments. Only the analysis of NEXT RTs is reported here, but the findings from the equivalent analyses using Correct NEXT RTs and NEXT PEs as the dependent variables can be found in the Supplemental Materials.

A significant Compatibility by Context interaction in these analyses would add weight to the notion that the contextual aspects of the instructions used to strategically modulate voluntary (GO) performance, also modulated AEIs. In contrast, evidence in support of the null hypothesis for this interaction from the Bayesian analyses would suggest that the pattern of results observed in Experiment 2 was less reliable.

Similarly, a significant three-way interaction would suggest that there were material differences in the effect of the contextual aspects of the instructions on AEIs between Experiments 1 and 2. Conversely, evidence in support of the null hypothesis for the three-way interaction from the Bayesian analysis would suggest that the effect of the

Table 6. Omnibus repeated measures ANOVA results for NEXT RTs from Experiments 1 and 2.

Effect	NEXT RT					
	DF	*MSE*	*F*	*p*	*BF* (%)	η^2
Experiment	(1,118)	112,172.99	0.30	.588	3.436 ± 7.7	.002
Context	(2,236)	1,993.26	5.38	.005	15.081 ± 7.5	.001
Compatibility	(1,118)	10,377.99	179.96	<.001	<0.001 ± 18.2	.109
Experiment × Context	(2,236)	1,993.26	0.96	.386	42.692 ± 6.7	<.001
Experiment × Compatibility	(1,118)	10,377.99	0.16	.688	10.073 ± 7.9	<.001
Context × Compatibility	(2,236)	1,615.66	3.02	.051	23.551 ± 6.4	.001
Experiment × Context × Compatibility	(2,236)	1,615.66	3.61	.029	10.220 ± 6.2	.001

Note. Equivalent Bayes factors are also reported. Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus, Bayes factors <1 indicate that the effect/interaction is an important contributor to the model.

contextual aspects of the instructions on AEIs observed in Experiment 2 were not significantly different to the lack of such an effect observed in Experiment 1.

Results and discussion

The results from the omnibus ANOVA combining NEXT RTs from Experiments 1 and 2 are reported in Table 6. For brevity, we only discuss the critical interactions here.

The Context by Compatibility interaction did not quite reach significance ($p = .051$), and the Bayesian analyses provided strong evidence in support of the null hypothesis ($BF = 23.551 \pm 6.4\%$). That is, any evidence from Experiment 2 indicating that the contextual aspects of the instructions affected the magnitude of AEIs was eliminated in the present analysis. This suggests that the pattern of results observed in Experiment 2 might not replicate.

The three-way interaction was significant ($p = .029$) suggesting that the Context by Compatibility interaction differed between experiments. However, the Bayesian analysis found relatively strong evidence that removal of the interaction from the full model would not materially affect its fit ($BF = 10.220 \pm 6.2\%$), suggesting that the Context by Compatibility interaction did not materially differ between experiments. Note that the effect size for this interaction in the frequentist ANOVA ($\eta^2 = .001$) suggests that any differences in the Context by Compatibility interaction were very small at best.

Taken together, the results from the combined analysis of NEXT RTs from Experiments 1 and 2 seem to suggest that any differences in the degree to which the contextual aspects of the instructions can modulate AEIs between experiments is small at best. Nonetheless, there were some conflicting results that indicate the need to replicate the study in an attempt to find more conclusive evidence in either direction.

Experiment 3

Experiment 3 was therefore a registered replication of Experiment 2 where some aspects of the procedure were optimised to maximise the likelihood that the context-specific strategy used to modulate GO performance was implemented during the NEXT phase thereby modulating AEIs. Specifically, Experiment 3 differed from Experiment 2 in two important ways.

First, participants for Experiment 3 were explicitly told how the contextual aspects of the instructions could be used to strategically modulate performance by relying more on the identity of the current phase rather than the identity of the current stimulus when selecting a response. In Experiments 1 and 2, the instructions merely stated that the contextual aspects of the instructions would be useful in optimising performance, but they did not explicitly state how. It is therefore still not known whether the strategy participants used to modulate GO performance in Experiments 1 and 2 was based on the evidence accumulated from the stimulus prior to making a response or some other strategy (e.g. biasing preparation for one response over the other in the GO phase). In Experiment 3, participants were explicitly instructed that it would be possible to achieve an acceptable level of performance in both NEXT and GO phases without necessarily relying on the identity of the stimulus to select an appropriate response in contexts where the required response in the GO phase was highly predictable. It was predicted that encouraging participants to utilise such a strategy where appropriate would reduce the magnitude of AEIs in these contexts relative to when the required response in the GO phase was unpredictable.

Second, the familiarisation phase was extended to give participants more practice with the various aspects of the paradigm before starting the experimental phase. In Experiment 2, the familiarisation phase was a single block of 12 mini-blocks and all of the different aspects of the session were introduced together. Some participants commented that they found the task difficult because there were so many different components, and they were not given an opportunity to practice them separately prior to the experimental phase. This might have affected their ability to fully assimilate all of the instructions resulting in a degree of confusion in how the contextual information

could be used to optimise performance. Contrary to this assumption, the pattern of results from the GO phase of Experiment 2 suggest that these aspects of the instructions were effectively implemented. However, we reasoned that allowing participants more time to practice the basic task before introducing the contextual aspects of the instructions, including the different location-context mappings, might help some participants to better understand how to use this information to optimise performance. Coupled with the additional instructions explicitly stating effective context-specific response selection strategies (as noted above), we predicted that extending the familiarisation phase and introducing each aspect of the session in separate blocks would help more participants to make best use of the contextual information presented in each mini-block, thereby reducing AEIs in contexts where the required response in the GO phase is highly predictable.

Specifically, if the magnitude of the NEXT compatibility effect was reduced in contexts where the required GO response was highly predictable, then we could conclude that the context-specific response selection strategy suggested in the instructions affected AEIs as well as voluntary (GO) performance. Conversely, if the magnitude of AEIs was equivalent for all contexts, then we could conclude that instructions communicating information about the context that could be used to strategically optimise voluntary (GO) performance do not affect the magnitude of AEIs (involuntary performance), even when those instructions describe an explicit response selection strategy designed to reduce the likelihood that the target stimulus would be identified prior to selecting an appropriate response.

Method

Participants. As in Experiments 1 and 2, recruitment of participants was conducted via Prolific. Participants were required to meet the following criteria to be eligible to participate: (a) fluent in English, (b) between the ages of 18 to 55, (c) normal or corrected to normal vision, (d) completed at least 20 previous submissions via Prolific with an acceptance rate of 90%. Participation was only allowed with a desktop or laptop PC (not a tablet or smartphone). Prolific's payment policy had changed since data were collected for Experiments 1 and 2. Thus, participants in Experiment 3 received £9 for completing the entire session (around 1 hr) or £2 for completing the familiarisation phase without meeting the inclusion criteria for participation in the experimental phase (see 'Materials and procedure' section for details).

Data collection was terminated when the Bayesian analyses (see below) found evidence in support of either the null or the experimental hypothesis ($BF > 6$ or < 0.167) for the critical Context by Compatibility interaction in the ANOVA on NEXT RTs. Data collection started with 60 valid data sets (see 'Analyses' section for details) and would have continued in steps of 10 valid data sets at a time (5 from each NEXT response counterbalancing group) until the termination criterion described above was met, or 150 valid data sets were collected (whichever came first). However, the relevant Bayesian analysis found conclusive evidence in support of the null hypothesis after 60 valid data sets had been collected, so data collection was terminated at this point.

From the final sample of 60 participants, 24 identified as female, 35 identified as male and 1 identified as diverse. Their mean age was 31.6 years ($SD = 9.0$). Fifty-three identified as being right-hand dominant and the remaining seven identified as being left-hand dominant.

Materials and procedure. The materials and procedure for Experiment 3 were identical to Experiment 2 with two exceptions. First, when the contextual aspects of the instructions were introduced, participants were explicitly informed of an effective strategy to use that information to optimise performance. Specifically, they were told that response selection can rely less on the identity of the current stimulus and more on the identity of the current phase in contexts where the required response in the GO phase is highly predictable. For example, an excerpt from the instructions states that 'when the required response in the critical trials (when the border is bold) is highly predictable, you can actually achieve an acceptable level of performance by only considering the weight of the border and not necessarily trying to identify the symbol'. The instructions provided some simple examples of how this might be implemented. Second, the familiarisation phase was extended to have a duration of five blocks, each containing 12 mini-blocks.

In the first familiarisation block, participants were introduced to the basic NEXT paradigm. That is, the instructions for each mini-block included only the two target stimuli displayed on the left/right of a white screen to indicate the required response in the GO phase. The stimuli in the NEXT and GO phases were also displayed on a white screen, and no mention was made of the different contexts or the associated landscapes that were introduced later.

In the second familiarisation block, the contextual aspect of the instructions was introduced. That is, the instructions for each mini-block included the probability that each of the stimuli would be displayed during the GO phase for that mini-block. As in the first familiarisation block, the instructions/stimuli were presented on a white background. Because the contextual aspects of the instructions were introduced in this familiarisation block, participants were instructed on how to make best use of this information to optimise performance by selecting responses predominantly based on the phase of the mini-block rather than the identity of the stimulus when the required response in the GO phase was highly predictable. The 'catch' trials were also introduced here.

In the third familiarisation block, the landscape images were introduced. That is, participants were informed that the landscapes were intended to help them remember which response would be most likely in the GO phase for each mini-block. The instructions/stimuli for this block were presented over a background image of one of the landscapes (depending on the context). Otherwise, this block was identical to the second familiarisation block.

In the fourth familiarisation block, the soundscapes were introduced. That is, participants were informed that the landscapes had been enriched by an accompanying soundscape to help them recall the most likely response in the GO phase. Because the soundscapes were introduced in this familiarisation block, the 'check' trials were also introduced here. This block was therefore identical to the experimental blocks that followed.

Participants were informed that the fifth and final familiarisation block was also a test. Having become familiar with every aspect of the procedure, participants were required to achieve at least 75% accuracy in every aspect of the test block (NEXT phase, GO phase, catch trial, and check trial) in order to be eligible to participate in the experimental phase. Any participants who did not meet the accuracy criterion for any aspect of the test block was automatically redirected back to the Prolific site and was paid £2 for their participation in the familiarisation phase (18 participants in total). Participants who did meet the inclusion criteria were automatically directed to the experimental phase where speeded responses, as well as accuracy, were emphasised. A further 20 blocks identical to the procedure for Experiment 2 followed.

Pre-registered analyses. As in Experiments 1 and 2, mini-blocks where an incorrect response was made during the GO phase were omitted from all NEXT analyses (7.0%). Trials with RTs <100 ms or >3000 ms were also omitted from all analyses (0.73%). Any participants for whom the above data cleaning procedures resulted in <20 observations per cell[5] (Context by Compatibility), or who achieved $<75\%$ accuracy on the catch or check trials were replaced (5 participants in total). As with the data from the NEXT trials, GO trials with RTs <100 ms or >3000 ms were omitted from all analyses, resulting in 0.59% data loss. Error trials were also omitted from the GO RT analysis.

Experiment 3 used identical analyses to those performed in Experiment 2. That is, GO RTs and GO PEs were analyzed using Context by Experiment half-repeated measures ANOVAs to determine the extent to which the contextual aspects of the instructions were used to strategically modulate GO performance.

NEXT RTs, Correct NEXT RTs and NEXT PEs were analysed using Compatibility (compatible, incompatible) by Context (the NEXT response is the required response on 90%, 50% or 10% of GO phase trials) by Experiment half (first half, second half) repeated measures ANOVAs to determine the extent to which the contextual aspects of the instructions affect AEIs. Equivalent Bayesian analyses were also performed as they were in Experiments 1 and 2. Only the NEXT RTs are reported in the manuscript. For completeness, the results from the remaining analyses are reported in the Supplemental Materials.

Although the ANOVAs did not find a significant Context by Compatibility interaction, planned paired-sample *t*-tests were conducted for each dependent variable to compare the magnitude of the NEXT compatibility effect between contexts. Bayes factors and effect sizes (Hedges' *g*) are also reported for these analyses.

Finally, Pearson correlations were computed to determine the extent to which context-specific modulations in GO performance predicted context-specific modulations in the NEXT compatibility effect.

Results

GO phase. As in Experiments 1 and 2, the ANOVAs on GO performance found that participants responded fastest and made fewest errors in the GO phase of mini-blocks where the NEXT response was the required response on 90% of GO trials (mean GO RT $= 587 \pm 20$ ms, mean GO PE $= 3.2 \pm 0.4\%$). GO performance was worst in mini-blocks where the NEXT response was the required response on 50% of trials (mean GO RT $= 643 \pm 18$ ms, mean GO PE $= 4.5 \pm 0.5\%$), and was intermediate on mini-blocks where the NEXT response was the required response on 10% of trials (mean GO RT $= 611 \pm 19$ ms, mean GO PE $= 4.6 \pm 0.6\%$). The corresponding ANOVAs on these data confirmed that GO performance was significantly modulated by the contextual aspects of the instructions (main effect of Context for GO RT: $F(2,118) = 31.77$, $p < .001$, $BF < 0.001 \pm 15.2\%$, $\eta^2 = .213$; GO PE: $F(2,118) = 6.79$, $p = .002$, $BF = 0.019 \pm 5.7\%$, $\eta^2 = .052$), indicating that the instructions were used to strategically optimise performance in the GO task.

GO RTs were also faster in the second half of the experimental session (mean GO RT $= 594 \pm 18$ ms) relative to the first half (mean GO RT $= 633 \pm 20$ ms; main effect of Experiment half: $F(1,59) = 33.66$, $p < .001$, $BF < 0.001 \pm 4.8\%$, $\eta^2 = .160$). However, participants made significantly more errors in the second half of the experimental session (mean GO PE $= 4.5 \pm 0.6\%$) relative to the first half (mean GO PE $= 3.7 \pm 0.4\%$; main effect of Experiment half: $F(1,59) = 5.95$, $p = .018$, $BF = 0.532 \pm 5.5\%$, $\eta^2 = .019$) suggesting a degree of speed-accuracy trade-off. The Context by Experiment half interaction did not reach significance for either GO RTs ($F(2,118) = 0.10$, $p = .906$, $BF = 17.907 \pm 3.7\%$, $\eta^2 < .001$) or GO PEs ($F(2,118) = 0.89$, $p = .412$, $BF = 10.218 \pm 8.4\%$, $\eta^2 = .005$) indicating that the effect of the contextual aspects of the instructions on GO performance was not modulated by practice.

Table 7. Omnibus repeated measures ANOVA results for NEXT RTs from Experiment 3.

Effect	*DF*	*MSE*	*F*	*p*	*BF* (%)	η^2
Context	(2,118)	7,757.56	2.45	.091	8.173 ± 13.6	.006
Compatibility	(1,59)	26,361.20	56.37	<.001	<0.001 ± 12.4	.201
Experiment half	(1,59)	19,381.59	27.71	<.001	<0.001 ± 12.8	.083
Context × Compatibility	(2,118)	5,916.69	2.48	.088	7.480 ± 17.7	.005
Context × Experiment half	(2,118)	5,425.40	0.99	.374	17.383 ± 12.7	.002
Compatibility × Experiment half	(1,59)	5,622.29	1.64	.205	5.498 ± 17.0	.002
Context × Compatibility × Experiment half	(2,118)	5,219.70	0.69	.504	10.932 ± 12.4	.001

Note. Equivalent Bayes factors are also reported. Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus, Bayes factors <1 indicate that the effect/interaction is an important contributor to the model.

Table 8. Results from *t*-tests comparing the NEXT compatibility effect for NEXT RTs across contexts (Probability that the NEXT response would be the required response in the GO phase) of Experiment 3.

Contrast	Difference	Lower CI	Upper CI	*DF*	*t*	*p*	*BF*	*g*
10% vs. 50%	−18	−46	10	59	−1.26	.211	0.301	0.165
50% vs. 90%	−13	−39	12	59	−1.06	.295	0.240	0.119
10% vs. 90%	−31	−62	−1	59	−2.04	.046	0.971	0.262

Note. Equivalent Bayes factors are also reported. Bayes factors >1 indicate evidence in support of the experimental hypothesis, whereas Bayes factors <1 indicate evidence in support of the null hypothesis.

Table 9. Pearson correlation results comparing context effect for NEXT RTs to GO performance from Experiment 3.

Context effect	NEXT RT vs. GO RT					NEXT RT vs. GO PE				
	r	*DF*	*T*	*p*	*BF*	*r*	*DF*	*t*	*p*	*BF*
50%–90%	−.120	58	−0.923	.360	0.430	0.255	58	2.007	.049	1.719
50%–10%	.104	58	0.794	.431	0.389	−0.164	58	−1.268	.210	0.602

Note. Equivalent Bayes factors are also reported. See main text for definition of context effect.

NEXT phase. The results from the repeated measures ANOVA on NEXT RTs from Experiment 3 are reported in Table 7.

As in Experiments 1 and 2, the ANOVA on the NEXT RTs found that participants responded faster on compatible (mean RT = 609 ± 28 ms) relative to incompatible NEXT trials (mean RT = 700 ± 32 ms; main effect of Compatibility: $p < .001$, $BF < 0.001 \pm 12.4\%$). Although there were some small differences in NEXT RTs according to the context (mean NEXT RT when the NEXT response was the required response on 90% of GO phase trials: 656 ± 33 ms; 50% of GO phase trials: 645 ± 28 ms; 10% of GO phase trials: NEXT RT = 663 ± 30 ms), they did not reach significance (main effect of Context: $p = .091$, $BF = 8.137 \pm 13.6\%$), indicating that the contextual aspects of the instructions did not affect absolute performance in the NEXT task. The critical Context by Compatibility interaction was non-significant ($p = .088$, $BF = 7.480 \pm 17.7\%$) indicating that the magnitude of the NEXT compatibility effect was broadly consistent across all contexts.

NEXT RTs were significantly slower in the first half of the experimental session (mean RT = 682 ± 30 ms) relative to the second half (mean RT = 628 ± 30 ms; main effect of Experiment half: $p < .001$, $BF < 0.001 \pm 12.8\%$), indicating that NEXT performance improved with practice. However, none of the interactions with Experiment half reached significance (all ps > 0.2, all BFs > 5.4), suggesting limited effects of practice on the other factors.

The results of the planned *t*-tests comparing the magnitude of the NEXT compatibility effect across the different contexts for Experiment 3 are reported in Table 8. None of the contrasts were significant or survived correction for multiple comparisons. The Bayesian analyses all found evidence in support of the null hypothesis, confirming that the contextual aspects of the instructions had not been assimilated to a degree that affected the magnitude of AEIs (i.e. automatic performance).

Correlations. The results of the correlations between the Context effect for performance in the GO phase and the NEXT compatibility effect for Experiment 3 are reported in Table 9. As in Experiments 1 and 2, there was little evidence that any strategic differences in GO performance between contexts transferred to the NEXT phase resulting in modulations to AEIs. Note that the positive correlation between the context effect for NEXT RTs and GO PEs for the difference in performance between mini-blocks where the required response was the NEXT response on 50%

Table 10. Omnibus repeated measures ANOVA results for NEXT RTs from Experiments 1–3.

Effect	NEXT RT					
	DF	*MSE*	*F*	*p*	*BF* (%)	η^2
Experiment	(2,177)	164,111.64	1.44	.240	2.761 ± 4.4	.014
Context	(2,354)	2,621.77	6.33	.002	6.709 ± 4.6	.001
Compatibility	(1,177)	11,312.19	230.14	<.001	<0.001 ± 5.0	.074
Experiment × Context	(4,354)	2,621.77	1.05	.380	507.352 ± 5.1	<.001
Experiment × Compatibility	(2,177)	11,312.19	0.40	.674	39.052 ± 4.4	<.001
Context × Compatibility	(2,354)	2,063.22	5.47	.005	10.243 ± 4.3	.001
Experiment × Context × Compatibility	(4,354)	2,063.22	1.64	.165	125.890 ± 4.8	<.001

Note. Equivalent Bayes factors are also reported. Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus, Bayes factors <1 indicate that the effect/interaction is an important contributor to the model.

versus 90% of GO phase trials did just reach significance ($p = .049$). However, the Bayesian analysis found only anecdotal evidence in support of the experimental hypothesis ($BF = 1.719$), suggesting that this result should be interpreted with caution.

Summary. The findings from Experiment 3 therefore more closely resembled those of Experiment 1 (where there was little evidence of the automatic implementation of the contextual aspects of the instructions) than those from Experiment 2 (where there was some indication that the contextual aspects of the instructions had been automatically implemented during the NEXT phase). That is, the magnitude of AEIs was largely unaffected by the contextual aspects of the instructions even though participants were instructed to use this information to reduce their efforts in identifying the target stimulus when the most likely required response could be determined by identifying the phase of the mini-block instead. We assumed that such a strategy would reduce the accumulation of information from the target stimulus prior to making a response, thereby reducing the magnitude of AEIs. Although the contextual aspects of the instructions were once again assimilated to a degree that modulated voluntary performance in the GO phase, there was little evidence that the additional instructions encouraging a strategic reduction in information accumulation prior to selecting a response affected the magnitude of AEIs in the NEXT phase.

Combined analysis of Experiments 1–3

To resolve the apparent conflict between the findings of Experiments 1 and 2, the NEXT trial data from Experiments 1–3 were collapsed together in a final pre-registered analysis. As with the combined analysis of data from Experiments 1 and 2, the resulting data set was analysed using Compatibility by Context by Experiment ANOVAs and equivalent Bayesian analyses to increase the power of the initial analyses and allow for a direct comparison of the patterns of results from all three experiments. Only the analysis of NEXT RTs is reported here, but the findings from the equivalent analyses using Correct NEXT RTs and NEXT PEs as the dependent variables can be found in the Supplemental Materials.

Again, the Compatibility by Context interaction and the three-way interaction are of particular interest here. Finding a significant Context by Compatibility interaction in these analyses would add weight to the notion that the contextual aspects of the instructions used to strategically modulate voluntary (GO) performance, also modulated AEIs. In contrast, evidence in support of the null hypothesis for this interaction from the Bayesian analyses would suggest that the pattern of results observed in Experiment 2 was simply an artefact and is not likely to replicate in future studies.

Similarly, a significant three-way interaction would suggest that there were material differences in the effect of the contextual aspects of the instructions on AEIs across the three Experiments. Conversely, evidence in support of the null hypothesis for the three-way interaction from the Bayesian analysis would suggest that the Context by Compatibility interaction was broadly similar across all three experiments.

Results and discussion

The results from the omnibus ANOVA combining NEXT RTs from Experiments 1–3 are reported in Table 10. For brevity, we only discuss the critical interactions here.

The Context by Compatibility interaction was significant ($p = .005$), but the Bayesian analyses provided strong evidence in support of the null hypothesis ($BF = 10.243 \pm 4.3\%$). That is, there was somewhat of a conflict between the frequentist and Bayesian analyses in that the former found evidence in support of the experimental hypothesis, but the latter found evidence in support of the null hypothesis. However, the effect size for this interaction was very small indeed ($\eta^2 < .001$), suggesting that any effect of the contextual aspects of the instructions on AEIs was minimal at best.

The three-way interaction did not reach significance ($p = .165$) and the Bayesian analysis found extreme

evidence in support of the null hypothesis ($BF = 125.890 \pm 4.8\%$), suggesting that the Context by Compatibility interaction was not materially different between the three experiments. That is, inclusion of the data from Experiment 3 in this analysis was enough to resolve the apparent conflict between the pattern of results from Experiments 1 and 2 reported above.

Although there was some conflict between the frequentist and Bayesian analyses in the present Context by Compatibility interaction, the effect size indicates that if there was any modulation of AEIs due to the contextual aspects of the instructions in the present experiments, this effect was very small at best. Taken together with the evidence in support of the null hypothesis for the three-way interaction, analysis of the combined NEXT RT data from Experiments 1–3 seem to suggest that the pattern of results in Experiment 2 was an artefact and not typical. That is, the present experiments found little evidence that instructions suggesting the context-specific strategic reduction of evidence accumulation from the target stimulus prior to selecting a response affected the magnitude of AEIs.

General discussion

The aim of the present experiments was to determine whether instructions describing context-specific strategic control routines designed to reduce the accumulation of evidence about the identity of a target stimulus prior to selecting a response would be implemented to a degree that modulated AEIs. The combined evidence from all three experiments suggests that the contextual aspects of the instructions were implemented, resulting in strategic modulations of voluntary performance in the GO phase. Specifically, performance was best when the required response in the GO phase was highly predictable and was worst when the required response was unpredictable. However, encouraging participants to select a response based on the identity of the current phase rather than the current stimulus (i.e. to reduce the amount of evidence accumulated about the identity of the stimulus prior to selecting a response) was largely not effective in modulating the magnitude of AEIs in the NEXT phase. That is, AEIs were broadly consistent across the three contexts in all three experiments.

AEIs are often interpreted as the instructed response being automatically triggered on detection of the instructed stimulus (e.g. Brass et al., 2017; Meiran et al., 2017). However, participants in the present experiments were indirectly (Experiments 1–2) or directly (Experiment 3) instructed to select an appropriate response on the basis of the current phase rather than the identity of the current stimulus. We reasoned that this would reduce the magnitude of AEIs by reducing the amount of evidence accumulated from the target stimulus prior to selecting a response. The observation that AEIs were largely unaffected by the contextual aspects of the instructions in the present experiments suggests that participants were still processing the target stimulus during the NEXT phase to a degree that triggered the instructed response. Further evidence in support of this conclusion comes from the proportion of errors made in the GO phase. If participants had used a response selection strategy that relied entirely on identifying the phase and not the stimulus, they would have made ~10% errors in the GO phase of mini-blocks where the required response was highly predictable (i.e. where the required response was the NEXT response on 10% or 90% of GO phase trials). However, the GO PEs plotted in Figure 2 indicate that participants made ~5% errors in the GO phase, and this was broadly consistent across all contexts and all experiments. This suggests that participants did not rely exclusively on the identity of the phase to select an appropriate response, even when such a strategy could have optimised their performance.

Liefooghe and De Houwer (2018) provide another example of research where AEIs were detected when there was little reason for participants to form an intention to implement the instructed S-R mappings. In their study, participants were never given the opportunity to implement the instructed S-R mappings because the inducer task merely required them to recall/recognise them (not implement them). Although participants in the present study were able to implement the instructed S-R mappings, there was no requirement to do so in two-thirds of the mini-blocks because an acceptable level of performance could have been achieved by responding according to the current phase without considering the current stimulus. Apparently, the participants from both studies were able to form a representation of the instructed S-R mappings that was effective in triggering AEIs despite being instructed that such a representation would never be implemented or would not optimise performance. This observation is especially intriguing given that AEIs are often interpreted as evidence of the automatic implementation of instructed S-R mappings maintained for future implementation. Taken together, these findings suggest that merely forming a mental representation of the relationship between a stimulus and a response is effective in producing AEIs, whether there is an intention/motivation to implement them or not.

If participants in the present study did not select a response on the basis of the current phase rather than the current stimulus, what strategy did they use instead that was effective in modulating GO performance, but not AEIs? One possibility is that they chose to perform the task as instructed (remember the S-R mappings for later implementation, press one key while moving through intervening trials, implement the S-R rules in the final phase), without attempting to implement the contextual aspects of the instructions as intended (i.e. by selecting a response based on the identity of the phase rather than the identity of the stimulus where possible). Another effective strategy might have been to simply bias preparation for one response over the other when the required response in

the GO phase was highly predictable. That is, participants might have effectively modulated GO performance by biasing preparation of a particular response rather than trying to ignore the identity of the stimulus when selecting an appropriate response (see e.g. Everaert et al., 2014). Such a strategy might explain why RTs were faster in the GO phase when the required response was highly predictable simply because a prepared (pre-activated) response might be faster to implement than a response that was not prepared in advance.

However, one might assume that advance preparation of a particular response would also affect AEIs given that Liefooghe et al. (2012) and Meiran et al. (2015) both independently found evidence that the magnitude of AEIs was (at least in part) determined by the degree of preparation for the GO task. Specifically, they found evidence that AEIs were only present when the S-R instructions were maintained for future implementation and were larger in mini-blocks where there was evidence of a high degree of preparation (e.g. faster RTs in the GO phase). This begs the question of why, if participants did use the contextual aspects of the instructions to bias preparation of a particular response in the present experiments, such a strategy was effective in modulating voluntary GO performance, but not AEIs. Indeed, the correlational analyses in the present experiments found little evidence that the context effect (difference in performance between mini-blocks where the required response was the NEXT response on 50% of GO trials relative to mini-blocks where the required response was the NEXT response on 90%/10% of GO trials) for GO performance was predictive of the context effect for AEIs. The latter observation seems to suggest that preparing (biasing) one response over another is not enough to affect AEIs despite prior findings indicating a relationship between global task preparation and AEIs (Liefooghe et al., 2013).

Further experiments would be necessary to unpick this relationship more fully. For example, by systematically manipulating the duration of the mini-block instructions screen, it would be possible to control the amount of time available to prepare for the upcoming mini-block. Any modulations of the magnitude of AEIs by the duration of the preparation interval would provide further evidence that AEIs rely on the active preparation of instructed S-R mappings for future implementation. However, such a design would only provide information about global task preparation. Another potentially fruitful avenue for future research might be to determine which components of the task and/or preparation are the drivers for modulating the magnitude of AEIs.

In Experiments 1 and 2, participants were not explicitly instructed to use a particular strategy. Instead, they were simply informed that the context might be useful in optimising GO performance. Although participants in Experiment 3 were encouraged to ignore the identity of the stimulus and select an appropriate response based on the identity of the phase, they might have also used the contextual aspects of the instructions to bias a particular response instead. The design of the current experiments cannot unequivocally distinguish between these two possible strategies. Further experiments would be necessary to confidently identify how participants were using the contextual aspects of the instructions to optimise GO performance and to determine the extent to which automatic processing of the target stimulus can be modulated by an evidence accumulation strategy designed to reduce such processing. For example, displaying the cue used to identify the current phase at a different location from the target stimulus might discourage processing of the features of the stimulus used to identify it during the NEXT phase, thereby reducing AEIs.

However, another important difference between the experiments of Liefooghe et al. (2012) and Meiran et al. (2015) on the one hand, and the present experiments on the other is that participants in the present experiments were required to rapidly change their strategy in response to regular changes in the context between mini-blocks. It is possible that randomly changing the context (and response biasing strategy) between mini-blocks was enough to reduce the effect of preparation on AEIs in the present experiments. For example, Braem et al. (2017) also found little evidence that context affected AEIs when it changed regularly through the session. However, they did find that participants were able to implement a context-specific strategy if they were given extended practice or when the context was held constant (see also Braem et al., 2022). This suggests that it might be possible to implement a single context-specific strategy that could affect AEIs. But it seems unlikely that instructions communicating complex rules (e.g. context-specific strategic control) would be automatically implemented, resulting in modulations to AEIs on a trial-by-trial basis. That is, consistent with the findings of Braem and colleagues, the present results support the notion that instructions communicating simple S-R rules and instructions communicating more complex contextual information are represented separately. While the former can result in AEIs, the latter cannot modulate the magnitude of AEIs, at least in a dynamic situation where the context changes on a regular basis (though see Jargow et al., 2022, for an example of trial-by-trial modulations of priming effects based on instructed probabilities).

Nonetheless, it should be noted that some of the present results were somewhat equivocal in this regard. Specifically, the critical Context by Compatibility interaction was significant in the ANOVAs performed on the NEXT RTs of Experiment 2 and the combined data from Experiments 1–3 (though the Bayesian analyses found evidence in support of the null hypothesis for both). However, given that the size of the effect was very small, we are reluctant to claim that the contextual aspects of the instructions were automatically implemented, or that their implementation in

a rapidly changing environment is typical. Further research is necessary to determine precisely which kinds of instructions can/cannot be automatically implemented in such environments, and whether the contextual aspects of the present experiments would have resulted in modulations to AEIs if the context was held constant throughout a block, or throughout the experimental session.

In conclusion, the present experiments found that the contextual aspects of the instructions had been implemented to a degree that modulated voluntary (GO) performance. However, there was little evidence that the implementation of this strategic control had simultaneously modulated AEIs in the NEXT phase. That is, we found evidence that instructions communicating context-specific information that was used to strategically optimise performance in the instructed task were not effective in modulating the degree to which instructions communicating S-R mappings were automatically implemented.

Data Accessibility Statement

The data and materials from the present experiment are publicly available at the Open Science Framework website: https://osf.io/eudxa/

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship and/or publication of this article: Cai Longman and Christina Pfeuffer are Guest Editors for the special issue that we hope this article will eventually be published in.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iDs

Cai S. Longman https://orcid.org/0000-0001-5564-7635
Christina U. Pfeuffer https://orcid.org/0000-0001-9394-8316
Andrea Kiesel https://orcid.org/0000-0001-5564-010X

Supplemental Material

The Supplemental Material is available at: qjep.sagepub.com

Notes

1. Note that Whitehead and Egner (2018) introduced a proportion compatibility manipulation into the NEXT paradigm to investigate whether top-down cognitive control processes could modulate the NEXT compatibility effect. However, their paradigm only manipulated the proportion of compatible/incompatible NEXT trials (not GO trials), and participants were never instructed about the manipulation. Although they found evidence that AEIs could be modulated by top-down control processes, these effects were not automatically induced by instructions that encouraged a strategy for response selection in the GO phase. That is, Whitehead and Egner (2018) did not provide evidence for automatic effects of context-specific strategic instructions.
2. We thank Jan De Houwer for suggesting this analysis.
3. Standard error of the mean.
4. We thank Jan De Houwer for suggesting this.
5. Note that the entire session for Experiment 3 had the same number of trials as Experiments 1 and 2 because of the counterbalancing protocol, but the extended familiarisation phase meant that the total number of trials available for analysis was somewhat reduced. Specifically, four additional blocks of 12 mini-blocks each were excluded from the analyses for Experiment 3 because they were part of the extended practice phase. Nonetheless, participants were still replaced if the data cleaning procedures resulted in less than 20 observations per cell. This is therefore a more conservative criterion than in the other experiments.

References

Amir, I., Peleg, L., & Meiran, N. (2022). Automatic effects of instructions: A tale of two paradigms. *Psychological Research*, *86*, 1467–1486. https://doi.org/10.1007/s00426-021-01596-1

Braem, S., Liefooghe, B., De Houwer, J., Brass, M., & Abrahamse, A. L. (2017). There are limits to the effects of task instructions: Making the automatic effects of instructions context-specific takes practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(3), 394–403. https://doi.org/10.1037/xlm0000310

Braem, S., Xu, S., Liefooghe, B., & Abrahamse, E. L. (2022). Learning when to learn: Context-specific instruction encoding. *PsyArXiv*. https://doi.org/10.31234/osf.io/7tvx3

Brass, M., Liefooghe, B., Braem, S., & De Houwer, J. (2017). Following new task instructions: Evidence for a dissociation between knowing and doing. *Neuroscience and Biobehavioral Reviews*, *81*, 16–28. https://doi.org/10.1016/j.neubiorev.2017.02.012

De Leeuw, J. R. (2015). JsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Everaert, T., Theeuwes, M., Liefooghe, B., & De Houwer, J. (2014). Automatic motor activation by mere instruction. *Cognitive, Affective, & Behavioral Neuroscience*, *14*, 1300–1309. https://doi.org/10.3758/s13415-014-0294-7

Jargow, J., Wolfensteller, U., Pfeuffer, C. U., & Ruge, H. (2022). Instructing item-specific switch probability: Expectations modulate stimulus–action priming. *Psychological Research*, *86*(7), 2195–2214. https://doi.org/10.1007/s00426-021-01641-z

Liefooghe, B., & De Houwer, J. (2018). Automatic effects of instructions do not require the intention to execute these instructions. *Journal of Cognitive Psychology*, *30*(1), 108–121. https://doi.org/10.1080/20445911.2017.1365871

Liefooghe, B., De Houwer, J., & Wenke, D. (2013). Instruction-based response activation depends on task preparation. *Psychonomic Bulletin & Review*, *20*, 481–487. https://doi.org/10.3758/s13423-013-0374-7

Liefooghe, B., Wenke, D., & De Houwer, J. (2012). Instruction-based task-rule congruency effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1325–1335. https://doi.org/10.1037/a0028148

Longman, C. S., Liefooghe, B., & Verbruggen, F. (2019). How does the (re) presentation of instructions influence their implementation? *Journal of Cognition*, *2*(1), 10. https://doi.org/10.5334/joc.63

Meiran, N., Liefooghe, B., & De Houwer, J. (2017). Powerful instructions: Automaticity without practice. *Current Directions in Psychological Science*, *26*(6), 509–514. https://doi.org/10.1177/0963721417711638

Meiran, N., Pereg, M., Kessler, Y., Cole, M. W., & Braver, T. S. (2015). The power of instructions: Proactive configuration of stimulus-response translation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *41*(3), 768–786. https://doi.org/10.1037/xlm0000063

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). *BayesFactor: Computation of Bayes factors for common designs* (Version 0.9.11-1). https://github.com/richarddmorey/bayesfactor

Pereg, M., Harpaz, D., Sabah, K., Ben-Shachar, M. S., Amir, I., Dreisbach, G., & Meiran, N. (2021). Learning the abstract general task structure in a rapidly changing task content. *Journal of Cognition*, *4*(1), 31. https://doi.org/10.5334/joc.176

Pereg, M., & Meiran, N. (2019). Rapid instructed task learning (but not automatic effects of instructions) is influenced by working memory load. *PLoS One*, *14*(6), e0217681. https://doi.org/10.1371/journal.pone.0217681

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Schönbrodt, F. D., & Wagenmakers, E. (2017). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin and Review*, *25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Theeuwes, M., De Houwer, J., Eder, A., & Liefooghe, B. (2015). Congruency effects on the basis of instructed response-effect contingencies. *Acta Psychologica*, *158*, 43–50. https://doi.org/10.1016/j.actpsy.2015.04.002

Verbruggen, F., McLaren, R., Pereg, M., & Meiran, N. (2018). Structure and implementation of novel task rules: A cross-sectional developmental study. *Psychological Science*, *29*(7), 1113–1125. https://doi.org/10.1177/0956797618755322

Whitehead, P. S., & Egner, T. (2018). Cognitive control over prospective task set interference. *Journal of Experimental Psychology: Human Perception & Performance*, *44*(5), 741–755. https://doi.org/10.1037/xhp0000493