Journal of Experimental Psychology: General
2007, Vol. 136, No. 1, 43–63

0096-3445/07/$12.00 DOI: 10.1037/0096-3445.136.1.43

Modes of Executive Control in Sequence Learning: From Stimulus-Based to Plan-Based Control

Elisabet Tubau
Universitat de Barcelona

Bernhard Hommel
Leiden Universiteit

Joan López-Moliner
Universitat de Barcelona

The authors argue that human sequential learning is often but not always characterized by a shift from stimulus- to plan-based action control. To diagnose this shift, they manipulated the frequency of 1st-order transitions in a repeated manual left–right sequence, assuming that performance is sensitive to frequency-induced biases under stimulus- but not plan-based control. Indeed, frequency biases tended to disappear with practice, but only for explicit learners. This tendency was facilitated by visual–verbal target stimuli, response-contingent sounds, and intentional instructions and hampered by auditory (but not visual) noise. Findings are interpreted within an event-coding model of action control, which holds that plans for sequences of discrete actions are coded phonetically, integrating order and relative timing. The model distinguishes between plan acquisition, linked to explicit knowledge, and plan execution, linked to the action control mode.

Keywords: sequential learning, executive control, individual differences, implicit and explicit knowledge, event-coding theory

Most everyday behavior consists of sequential acts, that is, of ordered sequences of more or less elementary behavioral units —just think of getting out of bed or preparing a cup of coffee. How do people acquire such behavioral sequences? According to William James (1890), perceiving the sensory effects of a given behavioral unit (such as the feeling of having assumed a seated position) may trigger the next unit (standing up from the bed) and so forth, until the sequence comes to its end. If so, sequential learning could be viewed as perceptual learning and consist of acquiring associations between the sensory cues that trigger the corresponding action units. However, Hugo Münsterberg (1892) pointed out a serious problem with this associative approach. Creating an association between one stimulus code and another, Münsterberg reasoned, eliminates the directional element that is needed to control behavioral sequences that are directed in time. That is, associating S1, the first stimulus of a sequence, with S2, the second, would as likely lead to the automatic retrieval of S2 if S1 is encountered as S1 would be retrieved if S2 is encountered: Learning ordered sequences would thus be impossible. Accordingly, behavioral sequences must rely on motor learning, that is, on the acquisition of a motor program. Münsterberg failed to explain why the temporal order problem does not also apply to motor learning. However, by introducing the idea that sequence production may rely on the acquisition of a program, that is, of a cognitive structure that governs the execution of behavioral sequences, he provided an interesting theoretical alternative to the James's chaining theory.

We were motivated to conduct the present study by the idea that James's stimulus-driven account of sequential learning and Münsterberg's program hypothesis may not represent mutually exclusive approaches. Hence, we do not assume that one account is right and the other is wrong. Instead, we attempt to integrate these two accounts by introducing and developing the concept of two alternative modes of executive control in a sequential task, *stimulus-based control* and *plan-based control*, which, under some conditions, can be strategically chosen.

The stimulus-based control mode represents a case of a "prepared reflex" (Hommel, 2000), where the cognitive system is prepared to respond to particular, typically highly response-compatible stimuli in a more or less automatic fashion. Accordingly, not much of the sequence is actually learned (in contrast to James's original chaining idea). What is learned is an efficient strategy to delegate control to external information. In contrast, the plan-based control mode relies on the construction of an action plan (Hommel, 2003; Luria, 1961; Miller, Galanter, & Pribram, 1960), which we assume to consist of ordered sequences of representations of action effects (Elsner & Hommel, 2001; Hommel, 1996). Plan-based control implies that plan-related representations (i.e., action-triggering signals) are internally generated (i.e., by means of inner speech; Luria, 1961; Vygotsky, 1934/1986; Zelazo, 1999). Hence, as we go along, we present theoretical and empirical reasons to motivate the idea that the construction and implemen-

Elisabet Tubau and Joan López-Moliner, Departament de Psicologia Bàsica, Universitat de Barcelona, Barcelona, Spain; Bernhard Hommel, Department of Psychology, Leiden Universiteit, Leiden, the Netherlands.

We are indebted to Scott Sinnett for his valuable comments and for improving the English.

Correspondence concerning this article should be addressed to Elisabet Tubau, Departament de Psicologia Bàsica, Universitat de Barcelona, Pg. De la Vall d'Hebron, 171, 08035, Barcelona, Spain. E-mail: etubau@ub.edu

tation of sequential plans is strongly related to and may sometimes even rely on phonetic encoding of the goals or instructions to respond. But before going into the details of our empirical strategy, we first address our distinction between stimulus-based and plan-based control modes.

Stimulus Control Versus Response Control

Whereas researchers in earlier studies were mainly concerned with deciding whether sequential learning is better characterized as stimulus or as response learning, more recent approaches have suggested that sequence learning involves a shift from perceptual to internal control on the basis of some internal representation or motor program (Hoffmann & Koch, 1997; Nattkemper & Prinz, 1997). Indeed, an increasing number of studies converge on the conclusion that advanced sequential learning is largely response based, although not muscle specific. For instance, Grafton, Hazeltine, and Ivry's (1998) participants successfully transferred sequential knowledge from a series of finger movements to a series of arm movements. Along the same line, Willingham, Wells, Farrell, and Stemwedel (2000) observed transfer between two series of keypresses only if the sequence of response locations was maintained, whereas the sequence of finger movements was not important (for a similar point, see Koch & Hoffmann, 2000). Furthermore, Hazeltine, Ivry, and Chan (1999; as cited in Keele, Ivry, Mayr, Hazeltine, & Heuer, 2003) found integration between two concurrently learned sequential tasks if the tasks shared a common response mode but not if they shared a common input modality. This suggests that sequential learning is mainly based on response codes, which is in agreement with the observation that the transfer of learning is hampered by a change of responses but not by the stimuli that signal them (Nattkemper & Prinz, 1997; Willingham, 1999; but see Willingham et al., 1989, for different results).

Even though some agreement has been achieved that advanced learning of sequences commonly involves response-related information, most approaches are vague with regard to the concrete characteristics of this information. For instance, Hazeltine (2002) was able to show that a change in the sequence of responses does not hamper performance if the environmental consequences of the new responses remain the same as in the training phase. As Hazeltine pointed out, sequence learning may thus be neither stimulus nor response based but what one may call goal based: Participants learn to produce a series of changes in the environment regardless of the particular triggered movements. Obviously, this only works if there is some association between the stored representations of those changes and the motor actions producing them. Indeed, converging evidence shows that actions are cognitively represented in terms of their effects (James, 1890; Lotze, 1852; see Hommel, 1996), which implies that motor responses are associated with and cognitively addressed via codes of action effects (Elsner & Hommel, 2001; Hommel, 1996). If so, it makes sense to assume that response-related sequence learning consists of integrating and organizing action-effect codes into structured action plans (Hazeltine, 2002; Stöcker & Hoffmann, 2004; Stöcker, Hoffmann, & Sebald, 2003).

Recent brain-imaging studies support such a plan-based view by revealing close associations between sequence learning and the activation of brain areas that are known to be involved in action planning, such as prefrontal, premotor, and supplementary motor cortex and cerebellum (e.g., Grafton, Hazeltine, & Ivry, 1995; Hazeltine, Grafton, & Ivry, 1997; Jueptner, Frith, Brooks, Frackowiak, & Passingham, 1997; Juptner, Stephan, et al., 1997). A recent study even found support for a shift from perceptual to plan-based control: Whereas superior parietal and occipital cortical regions were intensely involved in visual sequence learning during early stages and low performance, later stages of acquisition and higher levels of performance were characterized by stronger recruitment of prefrontal and mediotemporal regions (Müller, Kleinhans, Pierce, Kemmotsu, & Courchesne, 2002), which are more strongly related to sequential planning.

Learning Systems Versus Control Modes

Since the early work of Nissen and Bullemer (1987), many studies have proposed a distinction between two different systems for sequence learning (e.g., Curran & Keele, 1993; Destrebecqz & Cleermans, 2001; Keele et al., 2003; Mayr, 1996; Reber & Squire, 1994, 1998; but see also Perruchet & Amorim, 1992, or Shanks & Johnstone, 1999, for alternative interpretations). For example, Curran and Keele (1993) were among the first to suggest the existence of multiple types of learning. They showed that acquiring second- or higher order transitions is blocked—in terms of both acquisition and use—by performing a demanding secondary task, whereas learning first-order transitions is unaffected. Even though some failures to replicate have been reported (Frensch, Lin, & Buchner, 1998), a follow-up study has confirmed that the dual-task technique does indeed work to dissociate two types or patterns of learning (Frensch, Wenke, & Rünger, 1999). It is interesting to note that the learning type that is more sensitive to attentional manipulations seems to be closely related to the emergence of explicit declarative knowledge about the sequence (Curran & Keele, 1993). This fits well with the findings of imaging studies (i.e., Grafton et al., 1995; Hazeltine et al., 1997) comparing sequence learning under single- and dual-task conditions. Learning under single-task conditions was found to be associated with activation in inferior areas of the parietal occipital lobe and in the temporal lobe, whereas learning under dual-task conditions was associated with activation in, among others, parietal areas known to be involved in spatial coding and visually guided action (Jeannerod, 1997; Ungerleider & Mishkin, 1982).

According to Milner and Goodale (1995), the ventral pathway processes information that is accessible to verbal report, whereas the dorsal pathway directly feeds into movement control. Recently, Keele et al. (2003) extended and adapted this distinction to sequence learning. In their proposal, one learning system is assumed to be more ventral, including occipital and temporal areas as well as lateral prefrontal and premotor regions. The other learning system, more dorsal, comprises parietal and supplementary motor areas. The ventral system supports (but does not enforce or require) the development of explicit or declarative knowledge about sequences (as also suggested by Willingham, 1998) and is capable of cross-dimensional association, that is, of connecting sequences of codes defined in different representational subsystems. It is assumed to operate with highly processed, categorized stimulus information and supports the development of hierarchically structured action plans needed for the learning of sequences with ambiguous, context-dependent interelement transitions (Cohen et al., 1990; Keele, Cohen, & Ivry, 1990). In contrast, the dorsal system does not support explicit learning and is strictly intradimensional; hence, it can only connect codes defined in the same dimension, such as tasks with spatial stimuli and responses. It operates on relatively raw, uninterpreted stimulus codes (cf. Milner & Goodale, 1995) and is not capable of relational, context-sensitive sequential coding. It is claimed that both systems pick up contingencies automatically, but the cross-dimensional activities of the ventral system are likely to suffer from informational cross-talk in the presence of other, secondary tasks or interfering stimuli. Also, ventrally mediated learning may be more sensitive to attentional constraints (i.e., whether the to-be-associated events are defined on a dimension that is attended or not; Jiménez & Méndez, 1999).

Support for such a distinction comes also from Mayr and colleagues (Helmuth, Mayr, & Daum, 2000; Mayr, 1996), who showed that the sequences of responses to visual objects and to the spatial location in which they appear can be learned separately and, it seems, concurrently. If so, one would assume that learning takes place in different, nonoverlapping pathways: one more specialized to spatial information and one to nonspatial information. Tubau and López-Moliner (2004) also presented data consistent with this conclusion. These authors compared the learning of response sequences triggered by either spatial locations or location symbols. Performance on symbolic stimuli seemed to be more dependent on explicit knowledge, whereas spatial stimuli yielded much better implicit learning while at the same time producing more superficial learning based on first-order transitions (Soetens, Melis, & Notebaert, 2004; see below). Thus, only the learning system occupied with the symbolic sequence seemed to allow the acquisition of higher order transitions or the parsing of the sequence into larger chunks.

However, Keele et al.'s (2003) dorsal–ventral account is not the only viable interpretation of the available evidence. For instance, the shifts of activation observed in brain studies (Grafton et al., 1995; Hazeltine et al., 1997; Jenkins, Brooks, Nixon, Frackowiak, & Passingham, 1994; Jueptner, Frith, et al., 1997; Jueptner, Stephan, et al., 1997; Müller et al., 2002; Toni, Krams, Turner, & Passingham, 1998) may simply reflect a transition of control from external, stimulus-based control to a more internal, action-related or plan-based representation—a shift that may not need to fully comply with the ventral–dorsal distinction. Indeed, some empirical observations are rather difficult to make sense of in purely ventral–dorsal terms. For instance, Tubau and López-Moliner (2004) signaled responses by means of location symbols presented at randomly chosen left and right screen locations. Nonexplicit learners showed a Simon effect (Simon & Rudell, 1967), that is, better performance when stimulus and response locations were in agreement. However, explicit learners were able to overcome the Simon effect after some practice. Along the lines of Keele et al. (2003), one may attribute this latter effect to a transition of control from the dorsal to the ventral system, thereby gating out the possibly dorsal location information that produces the Simon effect. Unfortunately, however, a number of findings demonstrate that the Simon effect is not a dorsal phenomenon, at least not according to the conceptualization of Milner and Goodale (1995): The Simon effect is sensitive to instructions that change the spatial meaning of the responses (Hommel, 1993), and it can be produced by using induced motion and other illusory spatial cues (Kerzel, Hommel, & Bekkering, 2001) or by using spatial features of memorized objects (Hommel, 2002). Thus, a shift from dorsal to ventral pathways cannot explain the disappearance of the otherwise robust Simon effect. As an alternative, we attribute this observation to a shift of the control mode from what one may call *online* or *stimulus control* to a proactive *offline* or *plan control* mode that is driven by an internally generated action plan.

Frequency Learning Versus Plan Construction

As mentioned above, Curran and Keele's (1993) nonattentional form of learning allows for only the learning of dependencies between contiguous elements (e.g., unique transitions), an assumption that is consistent with observations in studies on implicit tone sequence learning (e.g., Saffran, 2002) and artificial grammar learning (e.g., Mathews et al., 1989). Thus, nonattentional, automatic, or implicit learning seems to be mainly restricted to first-order frequency information, a limitation also found in studies with nonhuman mammals (e.g., Newport & Aslin, 2004; Toro & Trobalón, 2005). This hypothesis emphasizes the importance of the transitions between consecutive stimuli and relates implicit sequential learning to sequential effects. In two-choice random reaction time (RT) tasks, response latencies to repetitions are faster than to alternations when the response–stimulus interval (RSI) is short (e.g., Bertelson, 1965; Kirby, 1976; Notebaert, Soetens, & Melis, 2001). This automatic facilitation, supposedly linked to a priming mechanism, is known as the *repetition effect*. Soetens et al. (2004) showed that implicit sequence learning is strongly related to this sequential effect. In their experiments, each of the two responses corresponded to two different stimuli, so that the stimuli could follow a probabilistic sequential pattern when the response sequence was random. Practice with this sequence had an effect similar to that observed in purely random two-choice tasks (i.e., the repetition effect tended to disappear). The authors concluded that the benefits of learning in tasks with structured as compared with random sequences are due to the frequency of occurrences of particular stimulus sequences (Soetens et al., 2004; see also Cleeremans & McClelland, 1991). Consistent with this frequency-based reasoning, the repetition benefit turns into an alternation effect if alternations are more frequent than repetitions (e.g., Hoffmann, Martin, & Schilling, 2003; Tubau & López-Moliner, 2004).

We consider the sensitivity to frequency information, especially that related to first-order transitions, to be informative about the current mode of control in sequence performance. For example, participants in the Tubau and López-Moliner (2004) study were required to learn a sequence in which alternations were three times as frequent as repetitions. Although participants who did not acquire explicit knowledge were strongly biased by this manipulation (i.e., performance was better for alternations than for repetitions), explicit learners were not only much faster than nonexplicit learners but they also did equally well on repetition and alternation responses. We attribute this pattern to a difference in the mode of control: stimulus-based control (in nonexplicit learners) and plan-based control (in explicit learners). Assuming that the emergence of explicit knowledge reflects that the plan (as a representation of the order of the actions) has been acquired, this result suggests that being able to execute the learned plan reduces or eliminates the impact of local, stimulus-related contingencies, probably because of its hierarchical structure.[1] If so, the Tubau and López-Moliner (2004) study suggests two conclusions. First, the presence or absence of first-order transitional frequency effects possibly indicates the operation of a stimulus-based versus plan-based control mode. Second, explicit knowledge seems to be strongly associated with the acquisition of an action plan. The present study was designed to further explore these relationships and to test whether the distinction between stimulus- and plan-based control modes sheds light on and provides a deeper insight into situational, strategic, and task-related factors in the acquisition of sequential skills.

Overview of Experiments

To test our stimulus–plan control framework, we manipulated factors that we thought might promote a change in the locus of control. An obvious factor is the task instruction, which can be *incidental* (i.e., not mentioning the explicit acquisition of sequential regularities as a task goal) or *intentional* (i.e., referring to explicit acquisition as a task goal). In the case of deterministic sequences, intention to learn has been observed to increase both the amount of explicit knowledge and the response speed (i.e., Eliassen, Souza, & Sanes, 2001; Frensch & Miner, 1994; Tubau & López-Moliner, 2004), suggesting an enhancing role of intention for plan formation. Therefore, if plan-based control is reflected in the disappearance of first-order transitional frequency effects and in the development of explicit knowledge, intentional instructions should more likely produce this plan-related pattern.

A second factor that we manipulated is the stimulus mode: visual (spatial and verbal) and auditory. As we elaborate in the General Discussion, there are reasons to believe that the creation of complex action plans and of hierarchical event representations in particular is mediated or at least strongly facilitated by inner speech (Barkley, 1997; Luria, 1961; Meacham, 1984; Vygotsky, 1934/1986; Zelazo, 1999)—an assumption that is consistent with the idea that phonetic (i.e., sound-related) coding is associated with plan-based control (cf. Goschke, 2000), goal activation (Miyake, Emerson, Padilla, & Ahn, 2004), or serial order control in task switching (Bryck & Mayr, 2005). If so, one would expect that visual–verbal stimuli, auditory stimuli, or other stimuli that can be easily coded phonetically propagate a plan-based control mode, thereby working against first-order frequency-related effects. In contrast, spatial, highly response-compatible stimuli should propagate a stimulus-based control mode, which should be visible in stable transitional frequency effects even in advanced learning. However, if control modes can strategically develop, one would expect that stimulus-induced effects interact with instruction. That is, even if spatial, response-compatible stimuli suggest a stimulus-based control mode, intentional instructions should increase chances for plan development.

To test these assumptions, we compared sequential learning with spatial, highly response-compatible stimuli (Experiment 1) and learning with symbolic, verbal stimuli (Experiment 2). We hypothesized that the latter would induce plan control independent of the type of instruction, whereas the former would do so only with an intention to learn—in contrast, incidental instruction should induce stimulus control. Moreover, if plan control is based on phonetically coded triggering signals, irrelevant sounds should work against plan control and force participants to base their responses mainly on the spatial information provided online (Experiments 3 and 4A). In contrast, irrelevant sounds presented after the response (Experiment 4B) or nonauditory noise (i.e., random locations of visual symbolic stimuli; Experiment 5C) should not have any effect on plan execution. However, if the connection between phonetic coding and plan control is as tight as we suggest, facilitating this type of encoding should enhance the development of a plan, even in the absence of the intention to learn (Experiments 5A and 5B). As argued above, we consider the emergence of explicit knowledge to indicate that a plan was acquired (at least that the order of triggering signals has been correctly represented). Consequently, we expected explicit learners to show evidence of plan-based control (faster responses, unbiased by first-order transition frequency and sensitive to plan interference) and nonexplicit learners to show evidence of stimulus-based control (slower responses, biased by transitional frequencies and insensitive to plan interference).

Experiment 1

In Experiment 1, we analyzed the effect of instruction (incidental vs. intentional) on learning a sequence of responses to spatial stimuli. Responses were mapped onto stimuli in a spatially compatible fashion, a condition that can be expected to support stimulus-based control: Stimuli trigger spatially corresponding responses even when location is task irrelevant (Simon & Rudell, 1967); thus, this effect should be exaggerated when location is relevant. Accordingly, staying in a stimulus-based control mode would be the most natural strategy to deal with the task in Experiment 1, allowing one to pick up local stimulus and stimulus-transition frequencies, which should result in frequency-based effects. To induce such effects, we used the same sequence as was used in Tubau and López-Moliner (2004), in which alternations of "left" (L) and "right" (R) responses were much more frequent than repetitions (the specific sequential pattern was RLRRLLRL; see also Eliassen et al., 2001). Even though stimulus-based control may be the more obvious strategy, we assume that participants can switch to plan-based control because of task requirements. As we assume that the acquisition of explicit knowledge is strongly correlated with plan-based control, we predicted that intentional instructions—which ask for the acquisition of explicit knowledge—would likely induce a plan-based control mode, which again should work against frequency-based effects.

Method

Stimuli and apparatus. A serial reaction time task with spatially defined stimuli and responses was used. The stimuli were presented using an IBM-compatible PC. The letter *X* appeared to either the left or the right of the center of the screen according to the sequence RLRRLLRL, and it was on the screen until the participant responded. The *X* was 0.7 cm high on a black back-

[1] Of course, it may be that stimulus-based control only allows for the acquisition of implicit information, whereas plan-based control allows for the acquisition of explicit information. At this point, we only mean to refer to a correlation between the mode of control and the reportability of the acquired knowledge without implying any particular causal direction. We return to this issue in the General Discussion.

ground, and it appeared 3.0 cm to the left or right of the screen's center. The left and right buttons of an external response box were operated with left and right index fingers, respectively.

Design. Participants received either incidental or intentional instructions. The Knowledge factor was created post hoc from the analyses of the verbal protocols given by the participants (see the *Procedure* section); hence, it varied between participants. Accordingly, the design comprised four independent groups that differed from each other with regard to the instructions and Knowledge. The within-subject factors were block of trials and transition frequency or, in short, frequency (frequent = alternations, infrequent = repetitions).

Participants. A total of 35 students from the University of Barcelona, Spain, participated in this experiment. Participants randomly received either incidental (17 participants) or intentional (18 participants) instructions.

Procedure. Participants were instructed to respond by pressing the button corresponding to the location of the letter *X* as quickly and as accurately as possible. For participants with incidental instructions, the experiment was introduced as one exploring the effect of training on RT. Those who received intentional instructions were informed about the existence of a repeating sequence of locations and their goal was, in addition to responding as quickly and as accurately as possible, to try to discover the structure of the sequence. After reading the instructions, participants were presented with a practice block of 24 random trials. They then had to work through the repeating sequence for 15 blocks of 48 trials each (each block contained six repetitions of the location sequence introduced above: RLRRLLRL). A final block of 48 random trials was used to test the amount of sequence learning. Repetitions and alternations were equally frequent in the random initial and final blocks. There was a pause of 5 s between blocks, and the RSI was 250 ms. Finally, all participants answered a questionnaire that tested their explicit knowledge. The first question was "Have you noticed any repeated location sequence?" In cases of an affirmative answer, there was the additional question: "Can you please write down the sequence of locations?"

Results and Discussion

As we previously mentioned, we created the post hoc factor Knowledge. Participants were assigned to the explicit groups when they reproduced the complete sequence (elements and order) correctly. As the sequence was presented continuously within a block, any starting point was considered correct (i.e. RRLLRLRL; RLRLRRLL; RLRRLLRL). Otherwise, participants were included in the nonexplicit groups. Explicit learners in the incidental and intentional conditions were 47% (8 out of 17) and 78% (14 out of 18), respectively; this difference was close to the significance criterion, $\chi^2(1, N = 35) = 3.53$, $p = .06$. Trials with incorrect responses (3%) or with RTs longer than 1,000 ms were not included in the analysis. An analysis of variance (ANOVA) with block (15 sequence blocks) and transitional frequency (high alternation and low repetition) as within-participant variables and instructions (intentional vs. incidental) and knowledge (explicit vs. nonexplicit) as between-participant variables showed that the main factors of block, frequency, and knowledge and the interactions Block × Knowledge and Frequency × Knowledge were significant (see Table 1). As shown in Figure 1, explicit learners were faster than nonexplicit ones (see also Table 2). The triple (Block × Instruction × Knowledge) interaction was also reliable; intentional–explicit learners tended to be faster than incidental–explicit ones. The difference between both explicit groups was close to the significance criterion, $F(1, 20) = 3.77$, $p = .06$, $\eta^2 = .16$. Block was not significant for nonexplicit participants.

Figure 1 and Table 1 also show that the transitional-frequency manipulation was highly significant (in general, responses were much faster for alternations than for repetitions), especially in the case of nonexplicit groups. For explicit learners, a significant Frequency × Instruction interaction emerged; the transitional fre-

Table 1
Results of the Analyses of Variance in Experiments 1, 2, and 3

	Experiment 1			Experiment 2			Experiment 3		
Factor	*F*	η^2	*df*	*F*	η^2	*df*	*F*	η^2	*df*
Block	10.27**	.25	14, 434	4.41**	.12	14, 448	8.62**	.22	14, 434
Block × Instruction (I)	1.66†	.05	14, 434	2.07*	.07	14, 448	*ns*		14, 434
Block × Knowledge (K)	7.79**	.20	14, 434	4.49**	.12	14, 448	1.91*	.06	14, 434
Block × I × K	1.89*	.06	14, 434	*ns*		14, 448	*ns*		
Frequency	91.89**	.75	1, 31	11.24*	.26	1, 32	53.48**	.63	1, 31
Frequency × K	19.73**	.39	1, 31	11.19**	.28	1, 32	*ns*		1, 31
Frequency × I × K	4.59*	.13	1, 31	*ns*		1, 32	*ns*		1, 31
Knowledge	10.68*	.26	1, 31	12.87**	.29	1, 32	*ns*		1, 31
Explicit groups									
Block	22.17**	.53	14, 280	7.96**	.29	14, 266	7.20**	.30	14, 238
Block × I	3.83**	.16	14, 280	2.45*	.11	14, 266	*ns*		14, 238
Frequency	16.04**	.44	1, 20	*ns*		1, 19	36.42**	.68	1, 17
Frequency × I	4.93*	.20	1, 20	*ns*		1, 19	*ns*		1, 17
Nonexplicit groups									
Block	*ns*		14, 154	*ns*		14, 182	3.31**	.19	14, 196
Frequency	110**	.91	1, 11	13.51*	.51	1, 113	22.86**	.62	1, 14

Note. Only significant results in any of the experiments are displayed.
† $p = .06$. * $p < .05$. ** $p < .001$.

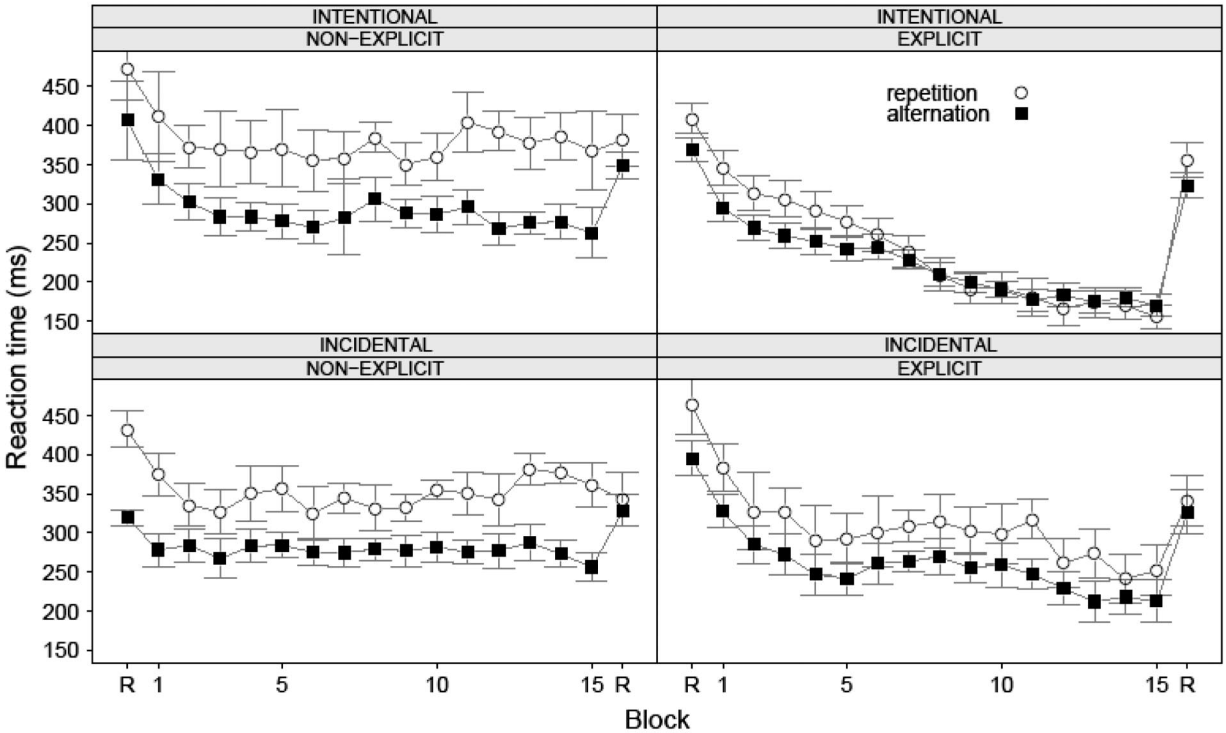

Figure 1. Results from Experiment 1, which show reaction time as a function of learning block split by transition frequency (low repetition and high alternation). Data are presented separately for type of instructions (intentional instructions in the top panels, incidental instructions in the bottom panels) and knowledge (nonexplicit knowledge in the left panels, explicit knowledge in the right panels). R = random blocks. Error bars denote 1 standard error.

quency effect was reliable for incidental–explicit learners, $F(1, 14) = 11.19$, $p < .05$, $\eta^2 = .61$, but not for intentional–explicit ones, $F < 1$.

Regarding the amount of learning (RT last random block − RT last sequence block), an ANOVA with frequency as a within-participants variable and knowledge and instructions as between-participant variables yielded similar results: There was a greater amount of learning for alternations than for repetitions, $F(1, 31) = 6.97$, $p < .05$, $\eta^2 = .18$, and a significant Frequency × Knowledge interaction, $F(1, 31) = 12.14$, $p < .005$, $\eta^2 = .28$. The effect of frequency was reliable for nonexplicit learners, $F(1, 11) = 10.96$, $p < .01$, $\eta^2 = .50$, but not for explicit ones, $F < 1$. However, in the latter case, the Frequency × Instruction interaction was significant, $F(1, 20) = 6.26$, $p < .05$, $\eta^2 = .24$. Intentional–explicit learners showed the opposite pattern (more learning for repetitions than for alternations), $F(1, 13) = 6.53$, $p < .05$, $\eta^2 = .33$. Knowledge was also highly significant, $F(1, 31) = 18.41$, $p < .001$, $\eta^2 = .37$. As shown in Figure 1, the amount of learning was much greater for explicit learners than for nonexplicit ones.

Hence, intention to learn seemed to facilitate the shift to the plan-based control mode; only explicit learners in the intentional condition switched to a control mode that does not rely on records

Table 2
Reaction Time Means, in Milliseconds, of the Last Sequence Block Split by Group and Type of Transition in Experiments 1, 2, and 3

Group and transition	Experiment 1				Experiment 2				Experiment 3			
	Repetition		Alternation		Repetition		Alternation		Repetition		Alternation	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Incidental												
Nonexplicit	361	33	256	20	470	27	415	24	337	32	280	22
Explicit	252	35	213	21	325	32	343	28	275	43	226	29
Intentional												
Nonexplicit	368	49	263	30	397	44	363	40	349	47	289	32
Explicit	156	26	170	16	280	25	293	22	249	29	218	20

of local transitions, overcoming the transitional frequency effect. Moreover, explicit knowledge correlated with evidence for plan-based control in the intentional condition (78% of the participants showed both explicit knowledge and plan-based control). In contrast, with incidental instructions, participants seemed to learn only first-order transitions, even in the case of explicit learners. Although 47% of them also reported the correct sequence, they were still affected by the transitional frequency. One possible account for this pattern is that incidental instructions delay both the emergence of explicit knowledge and the formation of an action plan needed for optimal response control. If so, Experiment 1 was long enough for some participants in the incidental condition to acquire the knowledge needed to correctly report the sequence but not to fully create the action plan that eliminates the frequency effect. Thus, as we discuss below, this result suggests that having access to an explicit (i.e., verbal) representation of the plan does not imply that the plan can be optimally executed.

Experiment 2

In Experiment 1, we used spatially compatible stimulus–response pairs, which allowed for both stimulus- and plan-based control. Accordingly, participants had two options: to prefer a less demanding stimulus-driven strategy or to trade it for the more demanding (but more efficient for overcoming stimulus-induced biases) plan-related strategy if the instruction made this an obvious choice. If one considers the use of symbolic stimuli, such as letters, stimulus–response compatibility would be reduced and the stimuli (and their corresponding responses) would more likely be verbally (i.e., phonetically) coded, thereby propagating plan-based control. If so, we would expect a less dramatic impact of instruction; that is, the transitional frequency effect should be eliminated under both incidental and intentional instructions. We tested this expectation in Experiment 2 by replacing the spatial stimuli used in Experiment 1 with symbolic letters that indicated left and right locations (i.e., the first letters of directional words).

Method

The same task as in Experiment 1 was used, only that "left" and "right" responses were signaled by the centrally presented first letters of location words: *D* (for *dreta*, Catalan for *right*) and *E* (for *esquerra*, Catalan for *left*). A total of 36 students from the University of Barcelona participated in this experiment. They were randomly presented with incidental (19 participants) and intentional (17 participants) instructions.

Results and Discussion

Participants were assigned to the explicit and nonexplicit groups using the same criteria as in Experiment 1. Intentional instructions again facilitated the emergence of explicit knowledge: the percentage of explicit learners was 42% (8 out of 19) and 76% (13 out of 17) under incidental and intentional instructions, respectively, $\chi^2(1, N = 36) = 4.36, p < .05$. The RT data were treated as were the data in Experiment 1. An ANOVA with the same factors as in Experiment 1 showed similar effects of block, knowledge, Block × Knowledge, and Block × Instruction (see Table 1). Block was only reliable in the intentional condition for explicit learners, $F(14, 210) = 3.40, p < .001, \eta^2 = .18$ (see also Table 2). The main effect of frequency, although having less impact than it had in Experiment 1, was also significant. However, whereas frequency was significant for nonexplicit learners, it was not for explicit ones, regardless of instructions (see Table 1 and Figure 2). The ANOVA with the amount of learning replicated the findings of Experiment 1: Explicit learners learned more than nonexplicit ones, $F(1, 32) = 18.95, p < .001, \eta^2 = .37$, and both frequency and Knowledge × Frequency were significant: For frequency, $F(1, 32) = 4.33, p < .05, \eta^2 = .12$; for Knowledge × Frequency, $F(1, 32) = 13.78, p < .01, \eta^2 = .30$. Frequency was only reliable for nonexplicit learners, $F(1, 13) = 13.49, p < .01, \eta^2 = .51$.

The outcome of Experiment 2 is clear. First, the predicted elimination of the transitional frequency effect was found independent of instructions. This is consistent with the assumption that using symbolic stimuli enhances verbal coding and plan development, irrespective of the motivation of participants to do so. Second, the transitional frequency effect was eliminated only for participants who were able to report the correct sequence, supporting the hypothesis that verbal or phonological coding played an important role in response control. However, intentional instructions had a similar effect as observed in Experiment 1, enhancing plan learning and speeding up the time course of creating it (the percentages of explicit learners in each learning condition were similar to the ones observed in Experiment 1). Hence, as is elaborated in the General Discussion, intentional instructions were not only relevant for enhancing a phonological encoding, they also seemed to activate specific rehearsal mechanisms for improving participants' memory of the sequence.

Experiment 3

Experiments 1 and 2 converge on the conclusion that eliminating the first-order transition frequency effect requires conditions that support the phonological encoding of the sequence and, thereby, the emergence of explicit knowledge. Given that we used different types of stimuli and different types of stimulus–response relations in the two experiments—spatial stimuli that are strongly response compatible in Experiment 1 but symbolic stimuli that are not as compatible in Experiment 2—the correlation between explicit knowledge and the processes that underlie the disappearance of the frequency effect does not depend on particular stimuli or stimulus–response relationships. Yet, in Experiment 1, the frequency effect was eliminated only if knowledge and intention were combined, whereas an instructional effect of this sort was not observed in Experiment 2. We attribute this discrepancy to the stimulus material and the stimulus–response relation it implies: The obvious and direct mapping in Experiment 1 afforded both stimulus-based and plan-based control, whereas the symbolic material of Experiment 2 only allowed for plan-based control. Hence, explicit learners, even with incidental instructions, formed a successful plan.

This scenario would be supported if it could be demonstrated that control choices vis-à-vis direct stimulus–response mappings could also go the other way: that is, if conditions could be found under which spatial stimulus–response pairings are processed in a way that does not eliminate frequency effects. In Experiment 3, we attempted to create such a condition. We used the same task that successfully eliminated the frequency effect in Experiment 1 but

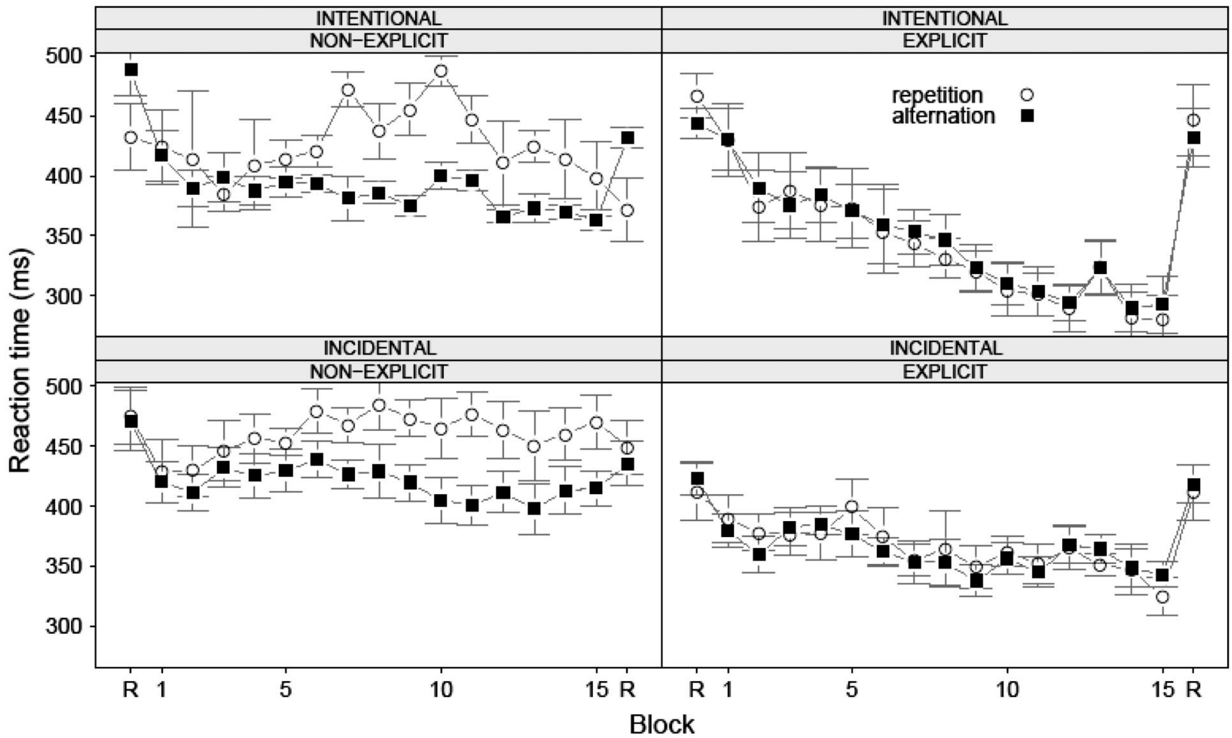

Figure 2. Results from Experiment 2, which show reaction time as a function of learning block split by transition frequency (low repetition and high alternation). Data are presented separately for type of instructions (intentional instructions in the top panels, incidental instructions in the bottom panels) and knowledge (nonexplicit knowledge in the left panels, explicit knowledge in the right panels). R = random blocks. Error bars denote 1 standard error.

added a slight modification that was intended to hamper verbal coding and, hence, plan-based control. To achieve that, we replaced the neutral letter *X*, which marked the stimulus locations in Experiment 1, with the first letters of location words that in the participants' native language denoted the locations left and right. As these letters varied randomly, they were of no use or informative value and could thus be safely ignored. However, from the literature on the Stroop effect, it is known that when the first letter of a written word relates to possible but incorrect response alternatives, it is likely to create response conflict (Rayner & Springer, 1986). This conflict would apply to half of all trials. Obviously, such a semantically based conflict can be avoided only if the stimuli are coded in a strictly spatial (but not verbal) fashion, which again should propagate stimulus-based control. This focus on spatial stimulus codes should allow participants to prevent any response conflict induced by the letter, but handing over control to the stimulus should guarantee a transitional frequency effect in all conditions, irrespective of instructions and knowledge.

Method

The same task as in Experiment 1 was used. However, instead of the neutral *X*s used to signal spatial location, we used the directional letters *E* and *D* (which were task irrelevant). Letters and locations were combined randomly, so that some combinations were semantically congruent (*E* on the left or *D* on right) whereas others were incongruent (*D* on the left or *E* on right). A total of 40 students from the University of Barcelona participated for extra course credit. Eighteen participants received incidental instructions and the other 22 received intentional instructions. All participants were explicitly instructed to ignore the specific letter and to attend only to its location.

Results and Discussion

Data from 1 participant in the incidental instruction group and from 4 participants in the intentional group were eliminated because of high error levels in incongruent trials (more than 70%, suggesting that these participants were responding to the symbol most of the time). The remaining data were analyzed as in the previous experiments. Intentional instructions again enhanced the emergence of explicit knowledge; the percentages of explicit learners were 72% (13 out of 18) and 35% (6 out of 17) in the intentional and incidental conditions, respectively; $\chi^2(1, N = 35) = 4.80$, $p < .05$. As shown in Figure 3, in general, RTs decreased through the sequence blocks, especially in explicit learners, and responses to alternations were faster than to repetitions (results of an ANOVA with the RT data and with the same design as previous experiments are presented in Table 1). In contrast to Experiments 1 and 2, the frequency effect had an important impact in all the groups, regardless of knowledge and type of instruction. That is, no experimental group was able to overcome the transi-

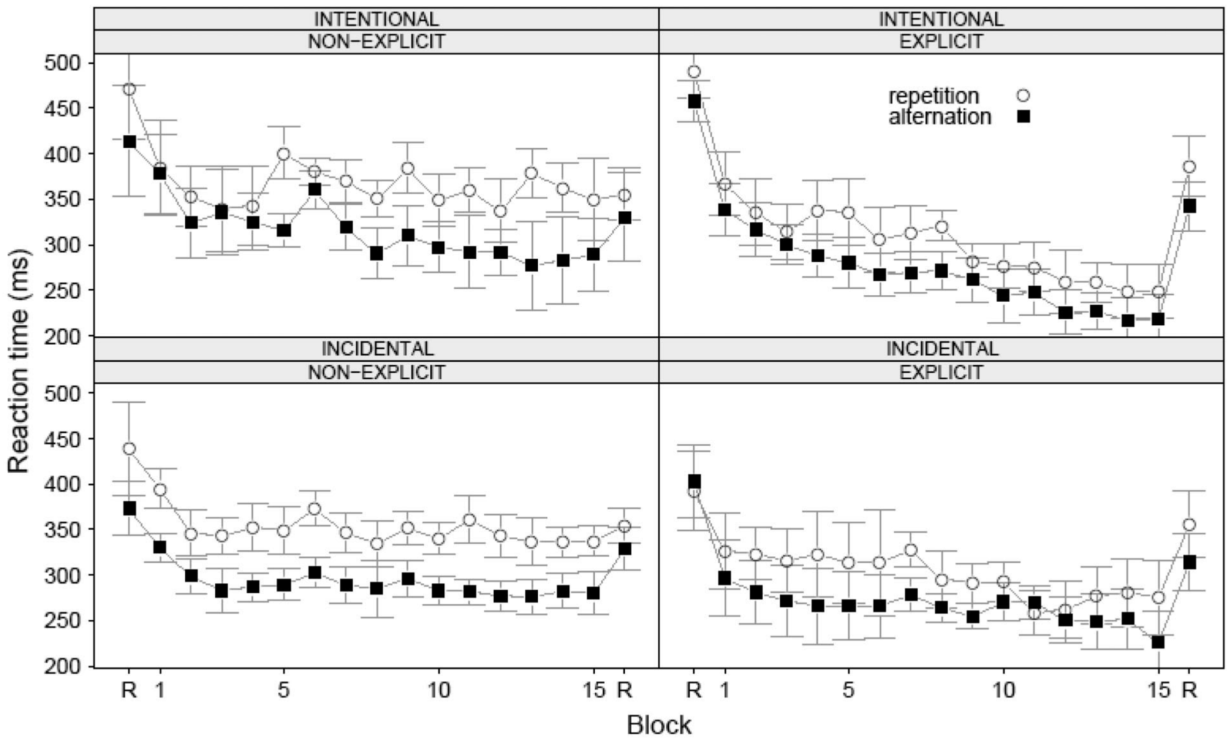

Figure 3. Results of Experiment 3, which show reaction time as a function of learning block split by transition frequency (low repetition and high alternation). Data are presented separately for type of instructions (intentional instructions in the top panels, incidental instructions in the bottom panels) and knowledge (nonexplicit knowledge in the left panels, explicit knowledge in the right panels). R = random blocks. Error bars denote 1 standard error.

tional frequency effect, which we take to reflect the absence of plan-based control. Furthermore, the global effect of knowledge, although in the same direction as in previous experiments, was only close to the significance criterion, $F(1, 31) = 3.09$, $p = .09$, $\eta^2 = .09$ (see Table 2). However, consistent with previous experiments, knowledge was clearly significant in the ANOVA with the amount of learning (RT random block − RT last sequence block), $F(1, 31) = 11.22$, $p < .01$, $\eta^2 = .27$.

Differing from previous experiments, block was significant for both explicit and nonexplicit learners. As shown in Figure 3, the effect of block in the case of nonexplicit learners was due to the difference between the first and the second sequence blocks; partial comparisons between blocks showed that only this difference was significant, $F(1, 14) = 8.52$, $p = .05$, $\eta^2 = .38$. Such improvement might reflect an initial effort to attend only to the spatial location of the symbol, ignoring its irrelevant meaning.

Regarding this point, we also analyzed effects of the symbolic congruency between stimulus location, the relevant stimulus feature, and the connotation of the stimulus letter, the irrelevant feature. To achieve this, we divided trials into those with congruent letter–location combinations and those with incongruent combinations, thus creating an additional factor. As expected, an ANOVA of RTs did not reveal any evidence of a congruency effect or an interaction of congruency with another factor. This confirms our expectation that participants switched to a control mode that avoids the impact of congruency (verbal coding), thus inhibiting action-plan formation and/or use. The impact of transitional frequency in all the groups also suggests that participants were using a stimulus-based control mode. However, did random symbols prevent plan acquisition or did they interfere with the precise control of the response timing?

As we previously commented, intention to learn supported the development of explicit knowledge to the same degree as in our previous experiments (the percentage of explicit learners in the intentional condition was twice as large as that in the incidental condition; 72% vs. 35%, respectively). Assuming that intention promoted the phonological encoding of the stimulus and accordingly enhanced plan acquisition, presenting random letters as spatial signals did not interfere with this process. But if plans were created and performance was hampered by those stimuli nevertheless, this would likely be related to interference with plan execution: Plans were available but were not or could not be carried out. Therefore, random symbols appeared to hamper performance not because of their semantic conflict with verbal–explicit encoding (explicit knowledge was not prevented) but because of their interference with some internal signal controlling the precise response timing. If, as we argue, the phonological representation of the plan is the reason for such interference, any auditory noise presented with the stimuli (i.e., irrelevant tones) should also interfere with the execution of the plan. However, enhancing the phonological encoding of the spatial sequence (i.e., by means of auditory action effects) would facilitate plan-based control. In the following experiments, we tested these hypotheses.

Experiment 4A

In Experiment 3, we found that intentional instructions produced about the same percentage of explicit learners as was produced in previous noninterfering conditions. To explain this outcome, we distinguished between the acquisition and the execution of an action plan and suggested that verbal interference did not prevent plan acquisition but plan execution only. Execution is hampered, so we assume, because making use of action plans requires or is at least facilitated by phonetic codes, with which irrelevant verbal stimuli would have interfered. However, alternative interpretations are possible. Even though the distractors in Experiment 3 worked as expected, they provided not only verbal noise but also conflicting information, that is, information that was likely to induce stimulus and/or response conflict. Experiment 4A was carried out to test this possible explanation, in which case a more complex interpretation than the one we just offered would be necessary. We replaced the Stroop-like distractors used in Experiment 3 with random tones, which do not elicit a response bias but nevertheless introduce noise in a presumably plan-related format.

Evidence from sequence-learning experiments suggests that irrelevant random tones presented before the response hamper sequence learning to a greater extent than do the same tones presented after the response (Hsiao & Reber, 2001; Schmidtke & Heuer, 1997). Comparable effects of irrelevant tones in nonlearning tasks have been interpreted as interference with response timing (Burgess & Hitch, 1999; Henson, Hartley, & Burgess, 2003) and, hence, with controlling the execution of the planned action (see also Saito & Baddeley, 2004). Even though speech sounds often produce more pronounced interference, simple tones are also effective, provided that they are variable (Jones & Macken, 1993). In view of these findings, we hypothesized that tones that randomly vary in pitch, presented together with the stimulus, would interfere with the execution of acquired action plans. To test if only plan execution was prevented by the tones, we did not present the random tones in the last five sequence blocks. If our assumption is correct, this manipulation would allow explicit learners (who should have acquired the plan) to immediately improve their performance by switching to plan-based control—which again should eliminate the frequency bias.

Method

The same task as in Experiment 1 was used. A randomly selected 100-ms high- or low-pitched tone (440 Hz and 200 Hz, respectively) was presented together with the onset of the spatial stimulus. Participants were given intentional instructions and were advised to try to ignore the tones and to attend and respond to only the location of the *X*. The random tones were presented until Block 11, after which no tones were presented (Blocks 12–16). Block 17 was a random block.[2] A total of 14 students from the University of Barcelona participated for extra course credit.

Results and Discussion

Participants were assigned to the explicit and nonexplicit groups the same way as were participants in previous experiments. The percentage of explicit learners was 64% (9 out of 14). The RT data were also treated as were data in the previous experiments. The ANOVA with the first 11 blocks yielded an effect of block that approached significance, $F(10, 120) = 1.86$, $p = .06$, $\eta^2 = .13$, and significant effects of frequency, $F(1, 12) = 52.20$, $p < .001$, $\eta^2 = .81$, and knowledge, $F(1, 12) = 6.88$, $p < .001$, $\eta^2 = .36$. None of the interactions were significant. Hence, although explicit learners were faster than nonexplicit learners (see Figure 4A), they were equally affected by the transitional frequency.

In contrast, the ANOVA with the last blocks without tones (Blocks 12–16) showed not only a significant effect of frequency, $F(1, 12) = 11.37$, $p = .01$, $\eta^2 = .49$, but also a reliable Frequency $\times$ Knowledge interaction, $F(1, 12) = 8.93$, $p < .01$, $\eta^2 = .43$. As can be seen in Figure 4A, explicit learners overcame the frequency effect in the noninterfering blocks. Knowledge was highly significant, $F(1, 12) = 82.99$, $p < .001$, $\eta^2 = .87$; explicit learners were much faster than nonexplicit learners in these last blocks (in the last sequence block, explicit learners were, on average, 222 ms faster than nonexplicit learners). Indeed, only the RTs of explicit learners tended to decrease: Block was significant for the explicit-learner group, $F(4, 32) = 7.04$, $p < .001$, $\eta^2 = .47$, but not for the nonexplicit one, $F < 1$. Furthermore, nonexplicit learners were much slower in the no-tone blocks than in the tone blocks; the mean RT of nonexplicit learners in the last tone block (Block 11) was 252 ms (for alternations, $M = 208$, and for repetitions, $M = 294$) compared with 357 ms (for alternations, $M = 317$, and for repetitions, $M = 396$) in the final no-tone block (Block 16). Differences between these blocks were significant, $F(1, 4) = 13.24$, $p < .05$, $\eta^2 = .77$.

Similar to the participants in Experiment 3 working with random symbols, the explicit learners in the current experiment showed the transitional frequency effect when random tones were presented. In the no-tone blocks, however, the frequency effect was overcome by explicit learners, suggesting that the interference was specifically related to the execution of the acquired plan. However, the pattern of nonexplicit learners (demonstrating a frequency effect in both tone and no-tone blocks and slower RTs in the no-tone blocks) suggests a clear stimulus-based control. If no plan has been acquired, the frequency bias cannot be overcome. Furthermore, nonexplicit learners' RTs were more sensitive to the physical dimension of the stimuli, being much faster after a sound, although irrelevant, than after a visual stimulus, as observed in nonlearning situations (i.e., Sanders, 1998). In summary, Experiment 4A demonstrates that even simple sounds can interfere with plan execution, which rules out the hypothesis that the interfering effects of distractors in Experiment 3 were due to the fact that they provided conflicting information. Accordingly, the general hampering effect of irrelevant random sounds on sequence learning (e.g., Hsiao & Reber, 2001; Schmidtke & Heuer, 1997) appears to reflect a specific interference of plan execution.

[2] To be consistent with the previous experiments, we always presented a final random block, which is displayed in the figures. Nevertheless, as analyses with amount of learning (RT random block – RT last sequence block) were quite redundant with the block and frequency effects analyses, we decided to report only these effects in the rest of the experiments.

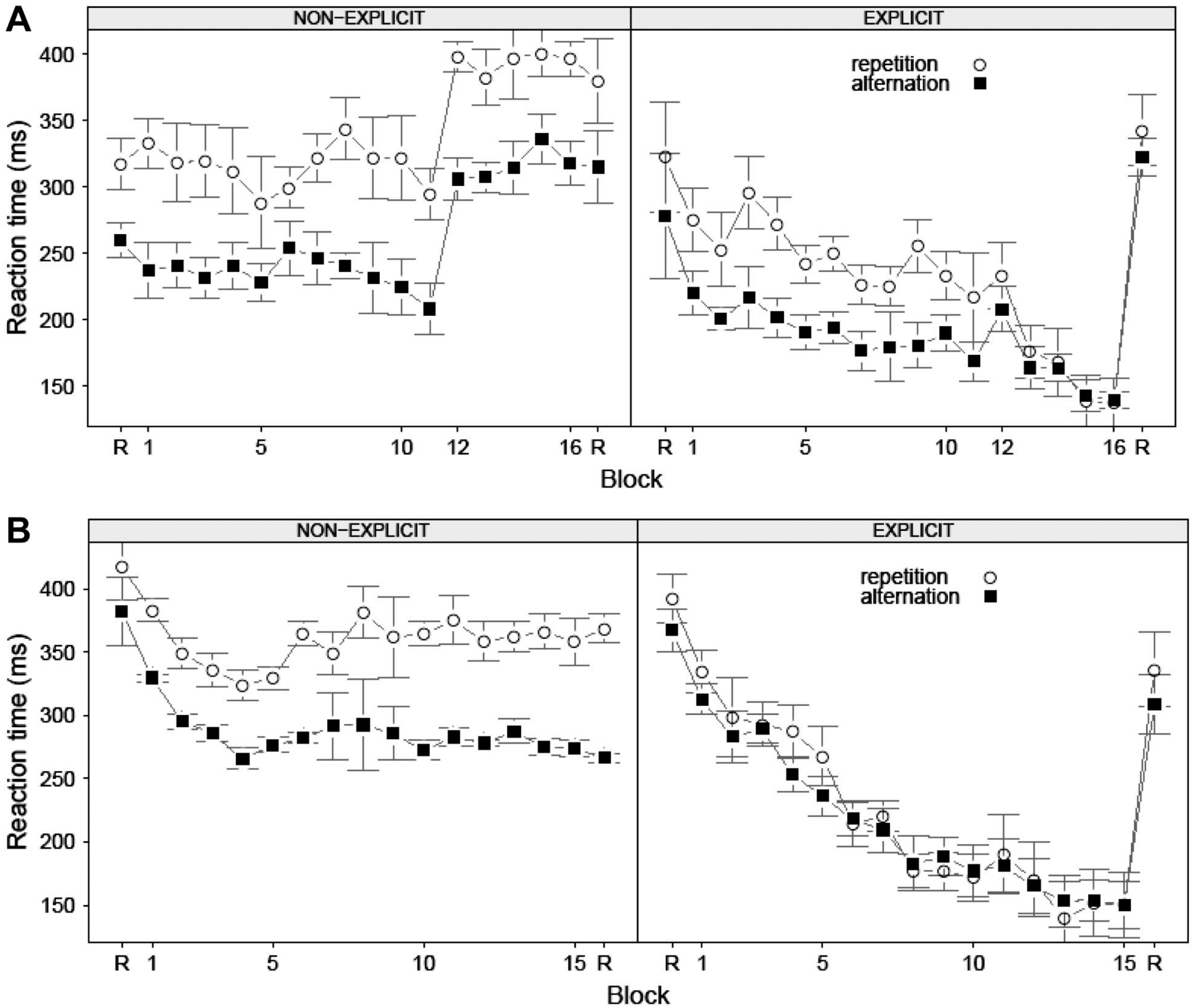

Figure 4. A: Results of Experiment 4A, showing reaction time as a function of learning block split by transition frequency (low repetition and high alternation) and knowledge. Blocks 1–11 included random tones with the stimuli. In Blocks 12–16, no tones were presented. B: Results of Experiment 4B, showing reaction time as a function of learning block split by transition frequency (low repetition and high alternation) and knowledge. Blocks 1–15 included random tones after the response. R = random blocks. Error bars denote 1 standard error.

Experiment 4B

In Experiment 4A, we presented the tone distractors together with the visual stimulus because they were likely to hamper response execution at this point in time (cf. Saito & Baddeley, 2004). The success of our manipulation may thus be taken to support the assumption that RT is indeed the crucial point, but, again, other interpretations are possible. In fact, it may be that any presentation of randomly varying tones is distracting, regardless of their presentation time and which temporal relationship they have with response execution. Experiment 4B was designed to test this possibility. Rather than presenting irrelevant and randomly varying tones simultaneously with the visual stimuli, we presented them after the response. According to our considerations, this should eliminate the impact of the tones on response execution and thus allow explicit learners to apply their knowledge. In other words, explicit learners should now show evidence of plan-based control.

Method

The same random tones used in Experiment 4A were triggered by the onset of the response in the current experiment. Participants were also given intentional instructions and were advised to try to ignore the tones and to attend only to the location of the letter *X*. A total of 15 students from the University of Barcelona participated for extra course credit.

Results and Discussion

The percentage of explicit learners was 73% (11 out of 15). RT data of 1 participant in the explicit group were dropped because her mean deviated by more than 2 standard deviations from the group mean. The rest of the data were treated as were data in the previous experiments. The ANOVA showed significant effects (all $ps < .001$) of block, $F(14, 168) = 8.34$, $\eta^2 = .41$; frequency, $F(1, 12) = 23.55$, $\eta^2 = .66$; and knowledge, $F(1, 12) = 17.02$, $\eta^2 = .59$. The Block × Knowledge and Frequency × Knowledge interactions were also significant, $F(14, 168) = 8.01$, $\eta^2 = .40$, and $F(1, 12) = 17.00$, $\eta^2 = .59$, respectively. Whereas block was only reliable for explicit learners ($\eta^2 = .71$), frequency was only significant in the case of nonexplicit learners ($\eta^2 = .95$). As Figure 4B reflects, the RTs of explicit learners decreased through blocks and were not affected by the transitional frequency.

Contrary to Experiment 4A, where the random tones apparently prevented the occurrence of plan-based control, presenting the tones after the response eliminated the transitional frequency effect in explicit learners, as plan-based control would predict. This result negates an interpretation of the impact of tones on performance in Experiment 4A as a general attentional distraction effect. Rather, random tones seem to prevent plan-based control only if they are presented sufficiently early to affect response execution.

Taken together, Experiments 4A and 4B strongly suggest that the time point at which sound distractors are presented matters. Therefore, one may conclude that phonetic codes have an important role in the control of the response timing. In Experiments 5A and 5B, we attempted to provide more direct evidence for the double function of phonetic coding: (a) facilitating the construction of sequential action plans in incidental learning conditions and (b) as a potential distractor because of its competition with the internal triggering signals related to response timing. Finally, in Experiment 5C, we attempted to demonstrate that replacing sound distractors with nonverbal visual distractors does not hamper performance.

Experiment 5A

In Experiment 4A, intentional instructions were given, which we assume promotes the acquisition of action plans. If action plans were indeed acquired, the fact that irrelevant tones were successful in preventing the standard plan-based-control pattern is informative with regard to the format of the action plan: It apparently is related to the way in which the tones were coded. In Experiment 5A, we attempted to tighten the link between action planning and phonetic coding by testing whether planning can be directly impacted by manipulating sounds. As discussed in the introduction, action plans can be considered ordered sequences of representations of action effects (Elsner & Hommel, 2001; Hommel, 1996). Any perceivable effect of an action element can be used to represent the element in the plan. For instance, one can code the same manual keypressing action in terms of the location of the key or of a visual effect that the keypress evokes (Hommel, 1993).

In the context of sequential learning experiments, Hoffmann, Stöcker, and colleagues demonstrated that sequences are acquired much more efficiently if each response is associated with a particular sound (i.e., a response-contingent auditory consequence; cf. Hazeltine, 2002), so that the response sequence produces a kind of melody (Hoffmann, Sebald, & Stöcker, 2001; Stöcker & Hoffmann, 2004; Stöcker et al., 2003). These findings suggest that response-contingent sounds facilitate the chunking of the response sequences into larger units as rudimentary action plans. This again has two important implications. First, it confirms earlier claims that responses are easily associated with—and can apparently be represented by—the sounds they produce (Elsner & Hommel, 2001; Hoffmann et al., 2001). Second, it suggests that sequential plans can be made, apparently more easily, out of auditory material, which fits nicely with our idea that plan-based control relies on phonetic (i.e., sound-related) codes.

These considerations suggest that we should be able to turn the largely ineffective auditory effects of Experiment 4B into highly effective action effects, which, accordingly, would be represented in the action plan, by introducing a systematic relationship between keypresses and tones. If we are correct in relating the findings of Hoffmann, Stöcker, and colleagues to what we call the plan-based control mode, we expect that evidence for plan-based control (the disappearance of the frequency bias) would be observed for explicit learners. More important, however, if presenting the response-contingent sounds facilitates plan-based control (because their format is more suitable for sequential planning), subsequent evidence of plan-based control even with incidental instructions should be observed. To replicate the dissociation between plan acquisition and plan execution found in Experiment 4A, we included transfer blocks at the end of the session, in which the response-sound contingencies were eliminated and randomly chosen sounds were presented with the visual stimuli. As shown in Experiment 4A, this manipulation should hamper the execution of the plan, which should reintroduce the frequency bias even after the participant has successfully learned and applied the plan.

Method

The high- and low-pitched tones used in Experiments 4A and 4B were presented after "right" and "left" responses, respectively. Participants (a total of 17 students) received incidental instructions. At Block 12, contingent tone effects were replaced by the random tones with the sequential stimuli. Block 17 was a fully random block (stimuli and tones). All participants were explicitly instructed to ignore the tones and to attend to only the locations of the letter *X*.

Results and Discussion

The percentage of explicit learners was 53% (9 out of 17). The mean percentage of errors was 2.5% (2% in the explicit group and 3% in the nonexplicit one). The RT data were treated as were the data in the previous experiments. The ANOVA with block (11 initial sequence blocks) and frequency as within-participant factors and knowledge (explicit and nonexplicit) as the between-participants factor showed that all main effects were significant (all $ps < .001$): for block, $F(10, 150) = 6.66$, $\eta^2 = .31$; for frequency, $F(1, 15) = 44.87$, $\eta^2 = .75$; and for knowledge, $F(1, 15) = 25.94$, $\eta^2 = .63$. Analyses also showed significant Block × Knowledge, $F(10, 150) = 6.5$, $\eta^2 = .30$, and Frequency × Knowledge, $F(1, 15) = 28.36$, $\eta^2 = .65$, interactions. Block was only reliable for explicit learners ($\eta^2 = .53$), and frequency was only significant for nonexplicit learners ($\eta^2 = .85$). As can be seen in Figure 5A, RTs of explicit learners decreased through blocks and, on average, were not affected by the transitional frequency bias.

In contrast, the ANOVA of the last blocks with the interfering tones (Blocks 12–16) showed that neither the main effect of knowledge nor the Frequency × Knowledge interaction was significant. Replicating the findings of Experiment 4A, the effect of frequency was highly significant, $F(1, 15) = 42.04$, $p < .001$, $\eta^2 = .73$, and it was reliable for both explicit ($\eta^2 = .69$) and nonexplicit ($\eta^2 = .77$) learners. As can be seen in Figure 5A, the transitional frequency bias reappeared in explicit learners. Furthermore, as shown in Experiment 4A, nonexplicit learners were faster in the last tone-interfering blocks than in the initial sequence blocks (on average, they were 53 ms faster). In contrast, RTs of explicit learners showed the opposite pattern (on average, they were 79 ms slower in the interfering blocks).

Consistent with previous studies (e.g., Hoffman et al., 2001), response-contingent auditory effects produced excellent performance in the case of explicit learners. It is important to note that comparisons between Experiment 1 (where no sounds were presented) and the present one (Experiment 5A, with response-contingent sounds) suggest that auditory action effects played a similar role in facilitating plan execution as intentional instructions did in Experiment 1. Assuming that intentional instructions promoted the association of each response with its corresponding phonetic code (see also the General Discussion), auditory action effects did the same job, which underscores the importance of action effects for action planning (Elsner & Hommel, 2001; Hoffman et al., 2001; Hommel, 1996). However, although auditory effects improved plan execution, they failed to enhance plan acquisition (the percentage of explicit learners with incidental instructions was the same with and without contingent tones, as a post hoc comparison of Experiments 1 and 5A showed; $p > .8$). This point is discussed further below.

Experiment 5B

The comparison of Experiments 5A and 1 (incidental condition) suggests that response-contingent sounds did not facilitate plan acquisition but did help plan execution. This conclusion is also supported when comparing Experiments 1 and 2 (incidental conditions): Verbal stimuli did not enhance the acquisition but did enhance the execution of the plan. Assuming that an optimal execution of the plan implies a precise control of the response timing, phonetic codes appear to be useful for this role. If it is true that timing issues are particularly important for plan execution, changing the timing demands of the task should be particularly harmful for plan-based control. In Experiment 5B, we tested this possibility by introducing blocks with randomly varying RSIs.[3] Random RSIs should render the availability of a temporally well-organized action plan useless and thus selectively impair performance in explicit learners.

Method

As in Experiment 5A, the high- and low-pitched tones appeared after "right" and "left" responses, respectively. Participants received the incidental instructions. After the 11th block, contingent tones were eliminated and the RSI was selected randomly among 50-, 250-, and 450-ms intervals. A total of 13 students participated in this experiment. All participants were explicitly instructed to ignore the tones and to attend to only the locations of the letter *X*. They were informed that the tones would disappear after several blocks, but they were not informed about the random RSI.

Results and Discussion

The percentage of explicit learners was 54% (7 out of 13). The RT data were treated as were the data in the previous experiments. An ANOVA with the 11 initial sequence blocks replicated the findings of Experiment 5A (see Figure 5B). An ANOVA of Blocks 12–16 (random RSI) showed that only frequency, $F(1, 11) = 187.00$, $\eta^2 = .94$, and the Frequency × Knowledge interaction, $F(1, 11) = 18.48$, $\eta^2 = .63$, were significant (all $ps < .001$). Although frequency had more impact on nonexplicit learners' RTs ($p < .001$, $\eta^2 = .99$), it was also reliable in the case of explicit learners' RTs ($p < .01$, $\eta^2 = .83$). As shown in Figure 5B, random RSI interfered with explicit learners' execution, but it had no effect on nonexplicit learners, who, if anything, showed the opposite pattern (nonexplicit learners were, on average, 40 ms faster in the random RSI than the contingent tone blocks; the difference between Blocks 11 and 12 was, however, nonsignificant). Explicit learners were, on average, 65 ms slower in the random RSI than the contingent tone blocks (the difference between Blocks 11 and 12 was significant, $p < .01$, $\eta^2 = .63$).

As predicted, random RSI interfered with performance in explicit learners, presumably because introducing temporal uncertainty prevents the execution of the acquired action plan. Possibly, the faster RTs observed in the nonexplicit learners could be explained as an increase in concentration due to an increase in temporal uncertainty, but we have no data to support this speculation. In any case, however, it seems clear that the RSI manipulation affected explicit and nonexplicit learners differently. This again supports our assumption that action plans integrate the temporal structure, at least in initial learning (see also Shin & Ivry, 2002), which would explain why phonetic coding is particularly suitable for representing sequential action plans.

Experiment 5C

Experiment 5A showed that irrelevant sounds are particularly effective in hampering the use of an acquired sequential plan. In Experiment 5C, we aimed to demonstrate, if possible, that such effects are unique to auditory distractors but do not occur with (nonverbal) visual stimuli. We used verbal symbols as in Experiment 2, except that now the location of the stimuli varied randomly. According to our account, this type of visual noise should not hamper plan-based control.

One technical problem was that sequential learning tasks using this particular design commonly fail to yield a high percentage of explicit learners (Tubau & Lopez-Moliner, 2004), which may have compromised our analyses. Given our observation in Experiments 5A and 5B that response-contingent tones strongly facilitate plan-based control, we used the same manipulation in the present experiment.

Method

As in Experiment 2, the sequence of responses was signaled by spatial symbols that are meaningful in Catalan (*E* indicating left

[3] We are grateful to Ulrich Mayr for this suggestion.

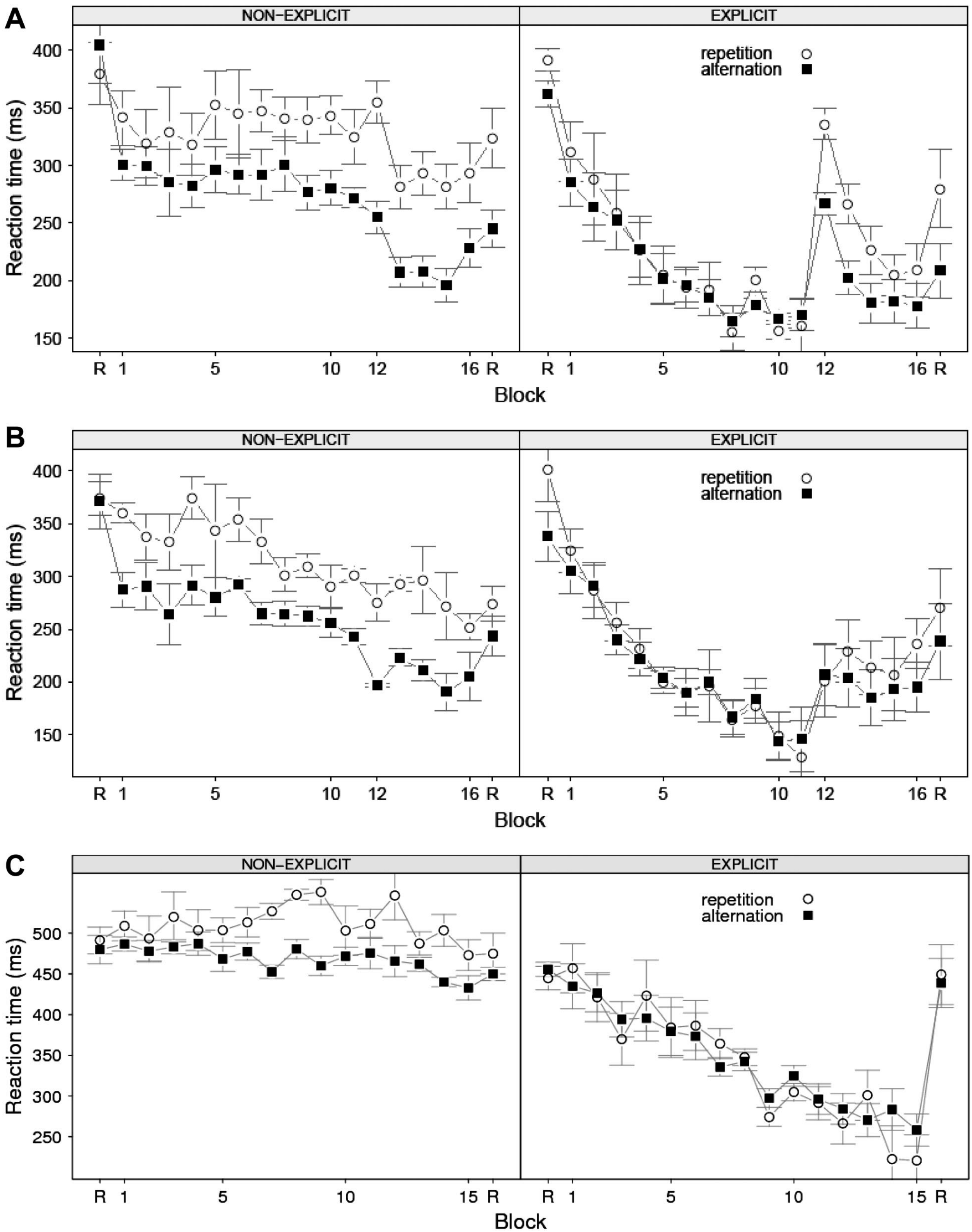

Figure 5. A: Results of Experiment 5A, showing reaction time as a function of learning block split by transition frequency (low repetition and high alternation) and knowledge. Blocks 1–11 included contingent tone effects. Blocks 12–16 included random tones with the stimuli. B: Results of Experiment 5B, showing reaction time as a function of learning block split by transition frequency (low repetition and high alternation) and knowledge. Blocks 1–11 included contingent tone effects. Blocks 12–16 included random response–stimulus intervals. C Results of Experiment 5C, showing reaction time as a function of learning block split by transition frequency (low repetition and high alternation) and knowledge. Blocks 1–15 included random spatial locations. R = random blocks. Error bars denote 1 standard error.

and *D* indicating right). However, the symbols appeared from the beginning at randomly varying left and right screen locations. Moreover, as in Experiments 5A and 5B, "left" and "right" responses triggered low- and high-pitched tones, respectively. In Block 17, both symbols and locations were presented randomly. A total of 14 students participated. They were instructed to attend to only the meaning of the symbol and to ignore its spatial position.

Results and Discussion

The percentage of explicit learners was 57% (8 out of 14). Data from 1 nonexplicit learner were discarded because of a very high error rate (more than 30%, as compared with the group mean of 5%). The remaining data were treated as were the data in the previous experiments. The ANOVA with block and frequency as within-participant factors and knowledge as a between-participant factor yielded significant effects of block, $F(14, 154) = 8.55$, $p < .001$, $\eta^2 = .44$; frequency, $F(1, 11) = 6.38$, $p < .05$, $\eta^2 = .37$; knowledge, $F(1, 11) = 16.11$, $p < .005$, $\eta^2 = .59$; Block × Knowledge, $F(14, 154) = 5.92$, $p < .001$, $\eta^2 = .35$; and Frequency × Knowledge, $F(1, 11) = 7.13$, $p < .05$, $\eta^2 = .39$. As shown in Figure 5C, explicit learners learned to respond much faster than did nonexplicit learners. Moreover, explicit learners, differing from nonexplicit learners, did not show the transitional frequency bias ($F < 1$).

A second ANOVA was performed with block and spatial congruency (the Simon effect) as within-participants factors and knowledge as a between-participant one. Results showed a significant congruency effect, $F(1, 11) = 89.43$, $p < .001$, $\eta^2 = .89$; a significant Block × Congruency interaction, $F(14, 154) = 2.38$, $p < .01$, $\eta^2 = .40$; and a significant Congruency × Knowledge interaction, $F(1, 11) = 6.97$, $p < .05$, $\eta^2 = .38$. Although explicit participants still showed the Simon effect at the end of the sequence blocks (the triple Block × Congruency × Knowledge interaction was not significant), its effect was much smaller for explicit (7 ms) than for nonexplicit learners (43 ms).

As previously observed by Tubau and López-Moliner (2004), explicit learners showed evidence of plan-based control; their responses were not affected by the transitional frequency bias and they were almost unaffected by the irrelevant stimulus location. Even more interesting was the fact that nonverbal visual distractors differed from irrelevant auditory stimuli in failing to hamper plan execution.

General Discussion

Our main goal in the present study was to characterize the two proposed control modes (stimulus based and plan based) by studying the impact of intention and the type of stimulus on learning a repeated sequence of left–right button presses. Overall, our findings showed a strong relationship between intention to learn, explicit knowledge, and plan-based action control. Plan-based control was induced by intentional instructions, by presenting target stimuli in a verbal (rather than spatial) format, and by providing learners with the opportunity to code their actions in terms of response-contingent sounds. Presenting visual–verbal target stimuli or auditory distractors had the opposite effect in inducing a stimulus-based control mode. Furthermore, interference with plan-based control was specific for sounds or verbal distractors; nonverbal visual noise (i.e., random stimuli locations) did not hamper performance.

Taken together, the experimental results support our main hypotheses. First, being able to communicate the order of the responses (i.e., to make it explicit) reflects the acquisition of an action plan, because only explicit learners showed a pattern that we attribute to plan-based control; in conditions of nonauditory interference (Experiments 1, 2, 4B, and 5C), explicit learners' RTs were independent from first-order transitional frequencies. In contrast, nonexplicit learners seemed to operate in a stimulus-based control mode, as their performance was always affected by the transitional frequency bias, in addition to being more sensitive to the physical dimension of the stimuli (auditory distractors facilitate performance; Experiments 4A and 5A).

Second, we found evidence that, at least with the sequences tested here, action plans are coded phonetically, that is, in a sound-related format. We saw that plan execution is facilitated by both visual symbols and response-contingent tone effects (Experiments 2, 5A, and 5B) but prevented by visual–verbal and auditory distractors (Experiments 3 and 4A). We speculated that the reason for phonetic coding is that it is particularly suited to create plans with a well-organized timing structure. This assumption is supported by our observation that introducing conditions that affect the timing of the response sequence (random RSI) specifically impairs performance in explicit learners, that is, in participants that are assumed to operate in plan-based control mode (Experiment 5B). In contrast, spatial distractors not linked to phonetic codes or to timing information did not hamper performance once the plan had been acquired (Experiment 5C). Sound distractors also failed to interfere with plan control when presented after the response (Experiment 4B), showing that auditory noise interferes only at a point in time when internal codes are triggering the next response to be performed. Finally, our results demonstrated a clear dissociation between plan acquisition, which goes hand in hand with the emergence of explicit knowledge, and plan execution, which depends on the availability and usefulness of internal signals for timing control.

It is interesting that intentional instructions produced the same level of explicit knowledge regardless of the physical features of the stimuli (on average, about 70%). The fact that this percentage was always higher than that achieved under incidental instructions (about 40%) suggests that instructions affected the encoding of the stimuli and/or the responses similarly in all experiments involving this type of instructions. Indeed, the intention to learn the sequence appeared to have an important impact on specific rehearsal strategies (see below).

An Event-Coding Account of Sequential Learning

In the following section, we attempt to make our assumptions more explicit and to embed them into a broader theoretical picture. We emphasize that this is only a first, preliminary step to model what we call stimulus- and plan-based control, but this step is helpful in appreciating the further implications of our approach and in applying it to other stimuli, response sets, and tasks. Given our assumption that switches between control modes imply changes in the way people translate stimulus events into actions, we choose to base our framework on the theory of event coding (TEC; Hommel, Müsseler, Aschersleben, & Prinz, 2001), which focuses on the role of action goals and their interactions with coding processes, that is, on the format in which perceptual events and action plans are cognitively represented.

According to TEC, both perceived stimulus events and planned actions are cognitively represented by distributed feature codes (see Hommel, 2003). For example, a left stimulus in Experiment 1 would activate codes referring to its leftward location (black color, X-like shape, etc.). Likewise, a "left" response would be coded in terms of its (anticipated) perceivable features, such as the expected kinesthetic feedback from the left index finger, the visible motion of the finger movement, the click sound the left button produces, the pitch of a possible tone effect, and so forth. The more features a stimulus and a response share, the more their cognitive representations will overlap and, hence, the more direct the stimulus–response translation would be (Hommel, 1997; Hommel et al., 2001).

First consider the spatial features. The relationship between codes of spatial stimulus locations and of left and right responses are sketched in Figure 6, where the spatial feature codes are indicated by arrowheads. Note that spatially varying *X*s are assumed to mainly activate spatial codes, which again are associated with the motor program operating the corresponding key—reflecting the fact that people are trained to produce left and right action effects by performing left and right finger movements, respectively (Hommel, 1997). Also note that the route from the stimulus location via spatial codes to spatially defined responses is assumed to be highly overlearned (Tagliabue, Zorzi, Umiltà, & Bassignani, 2000), as indicated by the thick arrows connecting asterisks to corresponding keys. Along this route, we assume, spatial stimuli are processed by default, that is, if instructions or other task characteristics do not suggest any other coding of the stimulus material. Given that feature codes are inert, that is, their activation curves outlive the physical presence of the stimuli and responses they code, sequential code activations will overlap. In other words, each given response will be selected and carried out in the presence of traces stemming from the previous response, which allows for the integration of local contingencies but not the development of higher order control structures. Such associations, which are necessarily directionally ambiguous in the sense of Münsterberg (1892), can produce local expectations. However, there is no reliable basis for representing a whole sequence of the sort we used. Accordingly, processing along this spatial route can be expected to be fast and reliable and to produce the transitional frequency bias, but it does not afford reliable chunking of sequence elements. This is exactly what we observed in the case of nonexplicit participants and in the incidental conditions of Experiments 1 and 3.

Symbolic stimuli with spatial meanings do not contain physical spatial features but only refer to them, so spatially processing them is not an option. Instead, visual words and letters are likely to activate phonetic structures according to overlearned grapheme–phoneme correspondence rules (cf. the dual-route assumption of word recognition; e.g., Ellis, 1982; Humphreys & Evett, 1985), which makes phonetic coding the most obvious choice for representing the stimuli used in Experiments 2 and 5C. Phonetic codes have three important characteristics for our purposes. First, they are also associated with "left" and "right" responses. That is, phonetic coding is a route to action control (Hommel, Daum, & Kluwe, 2004), although weak, as indicated by thin double arrows in Figure 6. Second, phonetic feature codes can also be used to code spatial events, such as left and right stimuli. This provides a coding alternative for the spatial conditions used in Experiments 1, 3, 4A, and 4B: People can choose between spatial and phonetic coding. Given that the spatial route is more efficient and less demanding, phonetic coding is certainly the less obvious choice under normal conditions, and yet it becomes an option if one is asked to verbally describe the sequence. That is, we assume that, all other things being equal, instructing people to try to learn the repeating pattern for further report increases the likelihood that stimuli and responses are phonetically coded, as indicated in Figure 6. The representation of a verbal action elicited by the stimuli or instructions entails that a set of distributed features denoting its perceived consequences (e.g., a phonetic code of "left" or "right") will be activated. Third, phonetic codes are commonly assumed to facilitate the creation of complex action plans and of hierarchical event representations in particular (Barkley, 1997; Luria, 1961; Meacham, 1984; Vygotsky, 1934/1986; Zelazo, 1999). Indeed, a number of studies have shown that vocal suppression and verbal noise hamper the development and/or application of task strategies, especially when one is learning to master a task (Baddeley, Chincotta, & Adlam, 2001; Emerson & Miyake, 2003; Goschke, 2000; Kray, Eber, & Lindenberger, 2004). Experiments 3 and 4A present evidence suggesting that visual–verbal and auditory noise can hamper the execution of the plan without necessarily preventing its acquisition, at least under the intention to acquire explicit knowledge.

Phonetic coding is probably the most flexible strategy, as it can be internally generated and applied to any stimulus or response that can be named or unequivocally described. Furthermore, considerable evidence shows that nonverbal phonetic coding can also be very effective, especially in coding actions and action plans. Numerous studies have shown that auditory action effects are quickly learned and associated with their accompanying responses so that representations of the auditory effect can serve as a retrieval cue for the action (e.g., Eenshuistra, Weidema, & Hommel, 2004; Elsner & Hommel, 2001; Hommel, 1996). Accordingly, it does not come as a surprise that introducing auditory action effects has been found to improve both sequence learning (Hazeltine, 2002; Hoffmann et al., 2001) and temporal chunking (Stöcker & Hoffmann, 2004). Indeed, it has been argued that phonetic coding is a natural and adaptive way for the brain to integrate timing information (Conway & Christiansen, 2002; Mahar, Mackenzie, & McNicol, 1994) and may allow sequential action control by a kind of internal humming (cf. Mehta, 2005)—analogous to control through internal speech afforded by verbal stimuli.

Plan Acquisition

We hesitate to conclude that any action planning has to be phonetic or even verbal to some degree, but our study strongly suggests that any experimental manipulation that facilitates phonetic coding increases the likelihood that people enter a plan-based control mode. One of these manipulations was an intentional instruction, which under all conditions tended to enhance the acquisition of explicit knowledge and, in some of them, to eliminate the impact of local contingencies (i.e., transitional frequency effects). A comparable pattern was produced by introducing verbal target stimuli or auditory action effects, which points to the con-

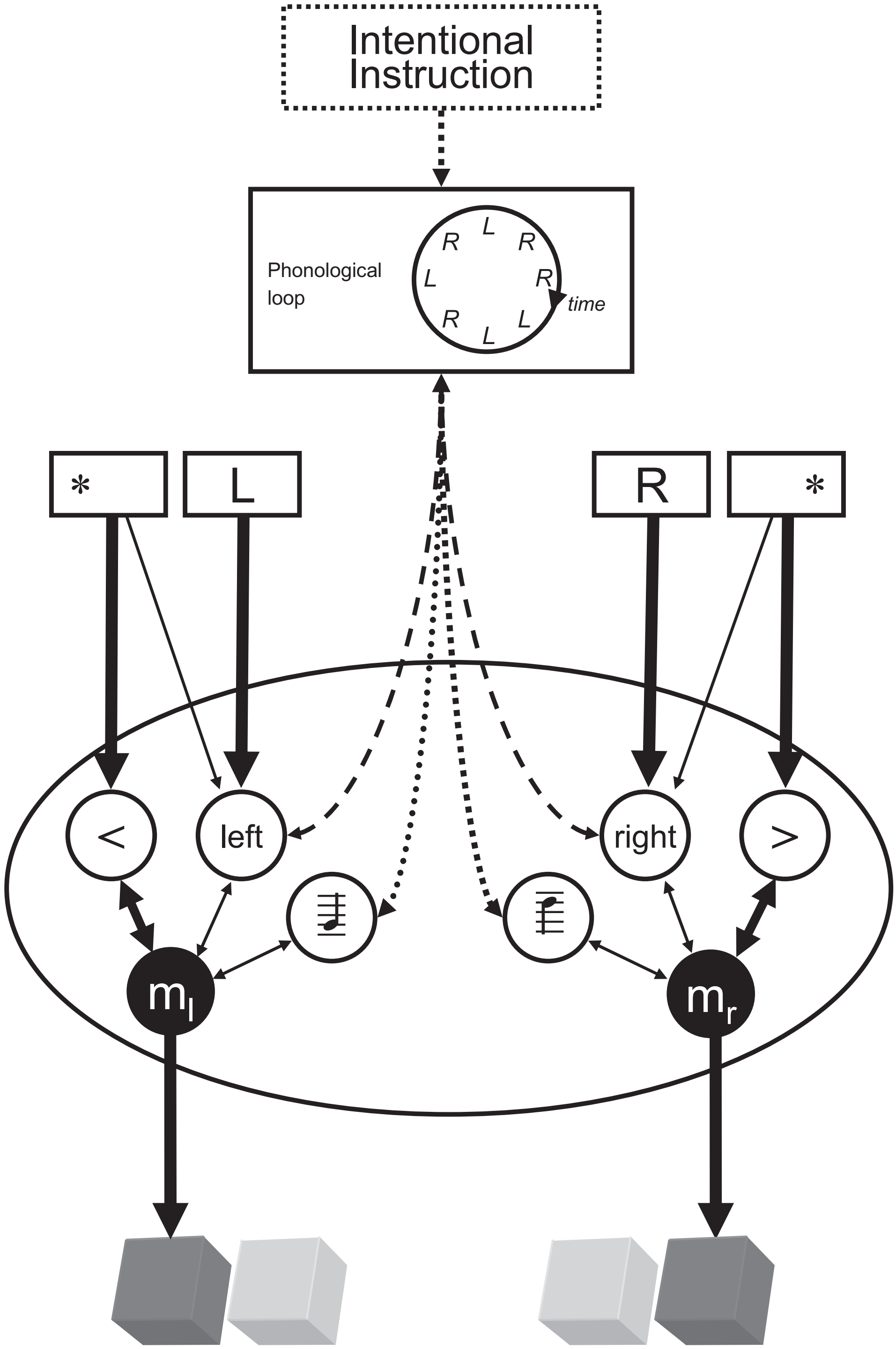

Figure 6. An event-coding account of stimulus-mode and instruction effects in sequential learning. Per default, spatial stimuli (asterisks) are coded spatially (indicated by arrowheads), whereas directional words and related verbal stimuli (letters) are coded phonetically (indicated by words). The intention to learn the order activates rehearsal processes, suggesting phonetic coding, which does not change matters for verbal stimuli but may induce a switch from spatial to phonetic coding with spatial stimuli. Actions (m_l and m_r = motor programs operating "left" and "right" responses, respectively) can be coded both spatially and phonetically—although spatial coding is more overlearned, so either code provides access to action control. The phonological loop thus facilitates the shift from stimulus-based to plan-based control, which integrates the relative response timing. R = pronouncing "right" or other sounds associated with the responses; L = pronouncing "left" or other sounds associated with the responses.

clusion that motivating or facilitating phonetic coding propagates plan-based control. However, intention appears to better enhance plan acquisition (as comparisons of Experiments 1, 2, and 5A suggest).

To account for the impact of intentional instruction, we assume that such a manipulation establishes the goal to acquire explicit knowledge about order, which again engages vocal rehearsal strategies (the phonological loop in Figure 6). With practice, rehearsal will lead to the consolidation of order information in the short term and at some point even in long-term memory. Vocal rehearsal consists of the repetitive sequential activation of phonetic codes, which, through consolidation, are likely to become part of a more complex action plan. To the degree that these codes are associated with the corresponding actions, either through overlearned long-term connections (as in Experiment 2) or by means of just-acquired action–effect links (as in Experiments 5A and 5B), the rehearsal-produced sequential (or hierarchical) structure will consist of direct action-retrieval cues and, hence, be an effective action plan. As a consequence, the emergence of an internal action plan and the ability to report the sequence explicitly go hand in hand.

Plan Execution

Once an action plan has been constructed and consolidated, it will take over the sequential triggering of the appropriate actions. For instance, the activation of the phonetic code of the upcoming action (say, "left") will spread to the action it is associated with (e.g., the left keypress), which drives the respective motor codes above threshold without requiring stimulus information. This does not mean that stimuli would be unable to provide additional input. For instance, note that explicit learners were faster in Experiment 1 than in Experiment 2. This suggests that the higher stimulus–response compatibility in Experiment 1 helped even those people who did not intend to and did not really need to consider stimulus information by adding external, stimulus-induced priming to the activation of motor codes provided by the action plan.

We have seen that evidence for plan-based control and for explicit knowledge is often correlated, but this correlation is far from perfect. For example, in Experiments 3 and 4A, the visual–verbal and auditory noise prevented plan-based control without affecting the acquisition of explicit knowledge. If we assume that the availability of explicit knowledge is a precondition for or perhaps an expression of the construction of an action plan, then even if the preconditions are met and perhaps even if action plans are constructed, auditory noise (i.e., distractor stimuli of a format that is similar or identical to that of the given plan) can prevent these plans from taking over action control (observed also in the last blocks of Experiment 5A). Noise may interfere by activating competing actions, which leads to a reduction of the activation of the correct action, and/or by inhibiting the correct action directly (cf. Pfordresher, 2003). This may either make the execution of an action plan impossible or at least render plan-based control unreliable. Note that this logic only applies to noise that appears while the next upcoming action is selected, which accurately fits with our observation that noise presented after a response has no effect (Experiment 4B).

As indicated by the outcome of Experiment 5B, plan execution is also impaired by randomizing the temporal structure of external mandatory stimuli. We take this to indicate that explicit learners developed action plans that kept the temporal structure provided by the stimuli. This does not mean that learners would be unable to carry out a plan at a different speed. In fact, there is evidence that sequential order and execution rate are independent, so an acquired sequence can be run at a lower or higher speed at will (Keele & Summers, 1976). However, in conditions where the triggering signal is externally presented, participants seem to be unable to escape the original rhythm, that is, the relative timing of the sequence components (Shin & Ivry, 2002; Summers, 1975). This suggests that the specific temporal pattern would be integrated with the internal phonological representation of the plan (see Figure 6). Therefore, acquired action plans are rendered useless by changes affecting the relative timing, which explains why performance in explicit learners (i.e., participants that we assume to operate in the plan-based control mode) was hampered by randomizing RSI.

Concluding Remarks

The proposed event-coding account allows us to disentangle a rather complex mosaic of observations resulting from manipulations of the intention to learn and the types and modalities of stimuli and action effects, as well as the dependence of these observations on the possession of explicit sequential knowledge. Our account borrows from both James (1890) and Münsterberg (1892) and integrates their views. The event-coding approach is Jamesian in assuming that actions are cognitively represented by perceptually derived codes of their effects. Action plans refer to and are actually made of those effect codes, so sequential learning actually is like perceptual learning. In contrast to James (1890), however, we suggest that local associations (i.e., associations linking the codes of directly succeeding sequence elements) are only part of the story and are often replaced or overruled by more complex cognitive structures, that is, true action plans. The event-coding approach is therefore Münsterbergian in assuming that advanced stages of sequential learning often include action programs (i.e., plans). In particular, we agree with Münsterberg (1892) that the acquisition of complex behavioral sequences with equivocal transitions between elements is unlikely to be based on simple, local associations.

Clearly, this is only a first step toward a more general theory of sequential action control, and whether and how our considerations can be generalized to other tasks, stimuli, and responses remains to be investigated. In particular, it would be very interesting to investigate whether alternative, nonphonetic (e.g., imagery based) planning strategies are as efficient as phonetic coding for sequential action control. Given the evidence that motor imagery plays a central role in several types of actions (Decety, 1996a, 1996b), it would be important to compare vision- and audition-based action planning.

This research also emphasizes the importance of taking into account individual differences. As revealed by previous studies of sequential learning (Hunt & Aslin, 2001), there seem to be at least two types of learners: those who preferentially process input at a more local level and those who process it at a more global level. Our results suggest that such individual preferences may reflect differences in encoding strategies, which may be codetermined by instruction and stimulus features. More specific research into this issue would improve psychologists' understanding of the specific mechanisms that realize the shift from stimulus-based to plan-based control.

References

Baddeley, A. D., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General, 130,* 641–657.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121,* 65–94.

Bertelson, P. (1965, April 10). Serial choice reaction-time as a function of response versus signal-and-response repetition. *Nature, 206,* 217–218.

Bryck, R. L., & Mayr, U. (2005). On the role of verbalization during task selection: Switching or serial order control? *Memory & Cognition, 33,* 611–623.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106,* 551–581.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120,* 235–253.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17–30.

Conway, C. M., & Christiansen, M. H. (2002). Sequential learning through touch, vision, and audition. In B. Bel & I. Marlien (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 220–225). Mahwah, NJ: Erlbaum.

Curran, T., & Keele, S. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 189–202.

Decety, J. (1996a). Do executed and imagined movements share the same central structures? *Cognitive Brain Research, 3,* 87–93.

Decety, J. (1996b). Neural representations for action. *Reviews in the Neurosciences, 7,* 285–297.

Destrebecqz, A., & Cleermans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review, 8,* 343–350.

Eenshuistra, R. M., Weidema, M. A., & Hommel, B. (2004). Development of the acquisition and control of action–effect associations. *Acta Psychologia, 115,* 185–209.

Eliassen, J. C., Souza, T., & Sanes, J. N. (2001). Human brain activation accompanying explicitly directed movement sequence learning. *Experimental Brain Research, 141,* 269–280.

Ellis, A. W. (1982). Spelling and writing (and reading and speaking). In A. W. Ellis (Ed.), *Normality and pathology in cognitive function* (pp. 113–146). London: Academic Press.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 229–240.

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language, 48,* 148–168.

Frensch, P. A., Lin, J., & Buchner, A. (1998). Learning versus behavioral expression of the learned: The effects of a secondary tone-counting task on implicit learning in the serial reaction task. *Psychological Research, 61,* 83–98.

Frensch, P. A., & Miner, C. S. (1994). Individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory & Cognition, 22,* 95–110.

Frensch, P. A., Wenke, D., & Rünger, D. (1999). A secondary tone-counting task suppresses expression of knowledge in the serial reaction task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 260–274.

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task set switching. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 331–355). Cambridge, MA: MIT Press.

Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience, 7,* 497–510.

Grafton, S. T., Hazeltine, E., & Ivry, R. (1998). Abstract and effector-specific representations of motor sequences identified with PET. *Journal of Neuroscience, 18,* 9420–9428.

Hazeltine, E. (2002). The representational nature of sequence learning: Evidence for goal-based codes. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance XIX* (pp. 673–689). Oxford, UK: Oxford University Press.

Hazeltine, E., Grafton, S. T., & Ivry, R. (1997). Attention and stimulus characteristics determine the locus of motor-sequence encoding: A PET study. *Brain, 120,* 123–140.

Helmuth, L. L., Mayr, U., & Daum, I. (2000). Sequence learning in Parkinson's disease: A comparison of spatial-attention and number-response sequences. *Neuropsychologia, 38,* 1443–1451.

Henson, R., Hartley, T., & Burgess, N. (2003). Selective interference with verbal short-term memory for serial order information: A new paradigm and tests of a timing-signal hypothesis. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 56*(A), 1307–1334.

Hoffmann, J., & Koch, I. (1997). Stimulus–response compatibility and sequential learning in the serial reaction time task. *Psychological Research, 60,* 87–97.

Hoffmann, J., Martin, C., & Schilling, A. (2003). Unique transitions between stimuli and responses in SRT tasks: Evidence for the primacy of response predictions. *Psychological Research, 67,* 160–173.

Hoffmann, J., Sebald, A., & Stöcker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 470–482.

Hommel, B. (1993). Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information. *Psychological Research, 55,* 270–279.

Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research, 59,* 176–186.

Hommel, B. (1997). Toward an action-concept model of stimulus–response compatibility. In B. Hommel & W. Prinz (Eds.), *Theoretical issues in stimulus–response compatibility* (pp. 281–320). Amsterdam: North-Holland.

Hommel, B. (2000). The prepared reflex: Automaticity and control in stimulus–response translation. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 247–273). Cambridge, MA: MIT Press.

Hommel, B. (2002). Responding to object files: Automatic integration of spatial information revealed by stimulus–response compatibility effects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55*(A), 567–580.

Hommel, B. (2003). Planning and representing intentional action. *TheScientificWorldJOURNAL, 3,* 593–608.

Hommel, B., Daum, I., & Kluwe, R. H. (2004). Exorcizing the homunculus, Phase two: Editors' introduction. *Acta Psychologica, 115,* 99–104.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24,* 849–878.

Hsiao, A. T., & Reber, A. S. (2001). The dual-task SRT procedure: Fine-tuning the timing. *Psychonomic Bulletin & Review, 8,* 336–342.

Humphreys, G. W., & Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *Behavioral and Brain Sciences, 8,* 689–740.

Hunt, R., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General, 130,* 658–680.

James, W. (1890). *The principles of psychology*. New York: Dover.

Jeannerod, M. (1997). *The cognitive neuroscience of action*. Cambridge, MA: Blackwell.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S. J., & Passingham, R. E. (1994). Motor sequence learning: A study with positron emission tomography. *Journal of Neuroscience, 14,* 3775–3790.

Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 236–259.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 369–381.

Jueptner, M., Frith, C. D., Brooks, D. J., Frackowiak, R. S. J., & Passingham, R. E. (1997). Anatomy of motor learning: II. Subcortical structures and learning by trial and error. *Journal of Neurophysiology, 77,* 1325–1337.

Jueptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S. J., & Passingham, R. E. (1997). Anatomy of motor learning: I. Frontal cortex and attention to action. *Journal of Neurophysiology, 77,* 1313–1324.

Keele, S. W., Cohen, A., & Ivry, R. (1990). Motor programs: Concepts and issues. In M. Jeannerod (Ed.), *Attention and performance XIII: Motor representation and control* (pp. 77–110). Hillsdale, NJ: Erlbaum.

Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review, 110,* 316–339.

Keele, S. W., & Summers, J. J. (1976). The structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109–142). New York: Academic Press.

Kerzel, D., Hommel, B., & Bekkering, H. (2001). A Simon effect induced by induced motion and location: Evidence for a direct linkage of cognitive and motor maps. *Perception & Psychophysics, 63,* 862–874.

Kirby, N. H. (1976). Sequential effects in two-choice reaction time: Automatic facilitation or subjective expectancy? *Journal of Experimental Psychology: Human Perception and Performance, 2,* 567–577.

Koch, I., & Hoffmann, J. (2000). The role of stimulus-based and response-based spatial information in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 863–882.

Kray, J., Eber, J., & Lindenberger, U. (2004). Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. *Acta Psychologica, 115,* 143–165.

Lotze, H. (1852). *Medicinische Psychologie oder Physiologie der Seele* [Medical psychology or the physiology of the soul]. Leipzig: Weidmann.

Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behavior.* New York: Liveright.

Mahar, D., Mackenzie, B., & McNicol, D. (1994). Modality-specific differences in the processing of spatially, temporally, and spatiotemporally distributed information. *Perception, 23,* 1369–1386.

Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., & Druhan, B. (1989). Role of explicit and implicit knowledge in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1083–1100.

Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 350–364.

Meacham, J. A. (1984). The social basis of intentional action. *Human Development, 27,* 119–123.

Mehta, M. R. (2005, January 6). Role of rhythms in facilitating short-term memory. *Neuron, 45,* 7–9.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the organization of behavior*. New York: Holt, Rinehart & Winston.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford, England: Oxford University Press.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115,* 123–142.

Müller, R.-A., Kleinhans, N., Pierce, K., Kemmotsu, N., & Courchesne, E. (2002). Functional MRI of motor sequence acquisition: Effects of learning stage and performance. *Cognitive Brain Research, 14,* 277–293.

Münsterberg, H. (1892). *Beiträge zur Experimentellen Psychologie* (Heft IV) [Contributions to experimental psychology (Issue IV)]. Freiberg: Mohr.

Nattkemper, D., & Prinz, W. (1997). Stimulus and response anticipation in a serial reaction task. *Psychological Research, 60,* 98–112.

Newport, E., & Aslin, R. (2004). Learning at distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48,* 127–162.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1–32.

Notebaert, W., Soetens, E., & Melis, A. (2001). Sequential analysis of a Simon task: Evidence for an attention-shift account. *Psychological Research, 65,* 170–184.

Perruchet, P., & Amorim, M. A. (1992). Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 785–800.

Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 949–964.

Rayner, K., & Springer, C. J. (1986). Graphemic and semantic similarity effects in the picture–word interference task. *British Journal of Psychology, 77,* 207–222.

Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning & Memory, 1,* 217–229.

Reber, P. J., & Squire, L. R. (1998). Encapsulation of implicit and explicit memory in sequence learning. *Journal of Cognitive Neuroscience, 10,* 248–263.

Saffran, J. (2002). Constraints on statistical learning. *Journal of Memory and Language, 47,* 172–196.

Saito, S., & Baddeley, A. (2004). Irrelevant sound disrupts speech production: Exploring the relationship between short-term memory and experimentally induced slips of the tongue. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 57*(A), 1309–1340.

Sanders, A. F. (1998). *Elements of human performance: Reaction processes and attention in human skill.* Mahwah, NJ: Erlbaum.

Schmidtke, V., & Heuer, H. (1997). Task integration as a factor in secondary-task effects in sequence learning. *Psychological Research, 60,* 53–71.

Shanks, D. R., & Johnstone, T. (1999). Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25,* 1435–1451.

Shin, J. C., & Ivry, R. B. (2002). Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 445–457.

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51,* 300–304.

Soetens, E., Melis, A., & Notebaert, W. (2004). Sequence learning and sequential effects. *Psychological Research, 69,* 124–137.

Stöcker, C., & Hoffmann, J. (2004). The ideomotor principle and motor sequence acquisition: Tone effects facilitate movement chunking. *Psychological Research, 68,* 126–137.

Stöcker, C., Hoffmann, J., & Sebald, A. (2003). The influence of response–effect compatibility in a serial reaction time task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 56*(A), 685–703.

Summers, J. J. (1975). The role of timing in motor program representation. *Journal of Motor Behavior, 7,* 229–241.

Tagliabue, M., Zorzi, M., Umiltà, C., & Bassignani, C. (2000). The role of long-term-memory links and short-term-memory links in the Simon effect. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 648–670.

Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during motor sequence learning: A whole-brain fMRI study. *NeuroImage, 8,* 50–61.

Toro, J. M., & Trobalón, J. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics, 67,* 867–875.

Tubau, E., & López-Moliner, J. (2004). Spatial interference and response control in sequence learning: The role of explicit knowledge. *Psychological Research, 68,* 55–63.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Engle, M. A. Goodale, & R. J. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.

Vygotsky, L. (1986). *Thought and language* (A. Kozulin, Trans.). Cambridge, MA: MIT Press. (Original work published in 1934)

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105,* 558–584.

Willingham, D. B. (1999). Implicit motor sequence learning is not purely perceptual. *Memory & Cognition, 27,* 561–572.

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1047–1060.

Willingham, D. B., Wells, L., Farrell, J., & Stemwedel, M. (2000). Implicit motor sequence learning is represented in response locations. *Memory & Cognition, 28,* 366–375.

Zelazo, P. D. (1999). Language, levels of consciousness, and the development of intentional action. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 95–117). Mahwah, NJ: Erlbaum.

Received May 19, 2004
Revision received July 20, 2006
Accepted August 3, 2006 ■