The microgenesis of action-effect binding

Ilona B. Dutzi · Bernhard Hommel

Abstract Ideomotor theories of human action control assume that performing a movement leads to the automatic integration of the underlying motor pattern with codes of its perceptual consequences. We studied the microgenesis of action-effect integration by varying the mapping of action effects upon actions from trial to trial. Experiments 1 and 2 showed that perceiving a tone repetition systematically affects one's tendency to carry out the response that produced that tone in the previous trial, suggesting that even the unintentional production of a stimulus creates a temporary binding of that stimulus with the action that brought it about. Experiments 3 and 4 extended this finding in suggesting that the integration and/or retrieval of action effects is modulated by attentional factors: Ongoing performance is more impacted by action effects if they are salient or match the current attentional set.

Introduction

People plan and perform voluntary actions in order to reach particular, intended goals, that is, to modify particular states of affairs or create particular events. Obviously, they can do so only if they have reliable knowledge at their disposal regarding which kind of action is likely to create the intended event. According to ideomotor theories of voluntary action (James, 1890; Lotze, 1852) and, to some degree, Piaget's (1946) sensorimotor approach to cognition, this knowledge is acquired "on the fly": Carrying out movements is assumed to be accompanied by a more or less automatic process of self-perception that integrates, without much ado, the motor patterns underlying the movement with the codes of that movement's perceptual consequences. In other words, actions become automatically associated with codes of their perceivable effects. This bilateral association provides the individual with a retrieval cue that allows creating that effect intentionally: One only needs to "think of" or "anticipate" (i.e., internally activate the codes of) particular action effects in order to prime and activate the action that has been experienced to produce that effect before.

Although this issue was neglected for quite some time, numerous recent studies provide increasing evidence that action effects are indeed picked up in an automatic fashion (for an overview, see Hommel & Elsner, 2009): People quickly acquire bilateral associations between actions and novel effects, such as keypress-contingent tones of a particular pitch or lights of a particular location, whether these effects are relevant to, or useful for the task at hand (Hoffmann, Sebald, & Stöcker, 2001; Hommel, 1993; Ziessler, 1998) or not (Beckers, De Houwer, & Eelen, 2002; Elsner & Hommel, 2001; Hommel, 1996). As studies using PET and fMRI have shown, once an action effect has been acquired its mere perception primes apparently associated motor structures (in the caudal supplementary motor area; Elsner et al., 2002; Melcher, Weidema, Eenshuistra, Hommel, & Gruber, 2008).

I. B. Dutzi
Bethanien Hospital, Geriatric Center, Heidelberg, Germany

B. Hommel
Leiden University Institute for Psychological Research and Leiden Institute for Brain and Cognition, Leiden, The Netherlands

B. Hommel (✉)
Department of Psychology, Cognitive Psychology Unit,
University of Leiden, Wassenaarseweg 52,
2333 AK Leiden, The Netherlands
e-mail: hommel@fsw.leidenuniv.nl

Even though the bulk of the evidence suggests that action-effect learning occurs spontaneously and without any intention to learn, Ziessler, Nattkemper, and Frensch (2004) have argued that effective action-effect acquisition depends on the active anticipation of the effects and is thus under attentional control. In their study, participants carried out pairs of manual responses signaled by visual letters ($S_1 \rightarrow R_1$, $S_2 \rightarrow R_2$). The second stimulus was systematically related to the preceding response so to allow for acquiring R_1–S_2 associations—which Ziessler et al. consider comparable to action-effect associations. As evidence for R_1–S_2 learning was obtained under undistracted conditions but not when participants were in addition to the sequential task to count tones presented in the R_1–S_2 interval, the authors conclude that R_1–S_2 acquisition cannot be automatic. But this conclusion is neither obvious nor necessary. First, ideomotor approaches claim that action-effect learning is automatic in the sense of not requiring an intention to learn, but they do not speak to the amount of cognitive resources involved. For instance, it is not unreasonable to assume that action-effect bindings need to be consolidated in order to affect subsequent behavior. As memory consolidation is known to be resource demanding and fragile (Jolicœur & Dell'Acqua, 1998), it may well be that it suffers from an overlapping task, such as tone counting. Second, even though it does not involve overt motor output, counting a tone may well be considered an intentional action. This means that Ziessler et al.'s tone-counting condition turned the original R_1–S_2 sequence into one where a third action intervened between R_1 and S_2, rendering it a R_1–R_2–S_2 sequence. If so, people might well have acquired R_2–S_2 associations, but that they failed to acquire R_1–S_2 associations is hardly surprising. Finally, the group that eventually showed the largest R_1–S_2 learning effects also showed by far the best performance on all measures from the very first trials on. For instance, their average reaction time for the first 12 performances of R_1 (the response that preceded and could thus not be affected by the tone) was already about 100 ms faster than the average of any of the other three groups. This strongly suggests major differences in motivation, which may also account for more efficient learning. In sum, we doubt that the available evidence provides strong support for a selective integration mechanism. On the contrary, numerous findings support the ideomotor expectation that carrying out a movement is indeed accompanied by the automatic (i.e., unintentional) integration of its perceptual consequences.

The present study focused on the microgenesis of this integration process, that is, the emergence of individual action-effect associations. According to the Theory of Event Coding (TEC) of Hommel, Müsseler, Aschersleben, and Prinz (2001) stimulus and action events are integrated in two phases. The first, *activation phase* consists of activating codes of a particular stimulus and/or action feature, be that internally driven, as in the case of action planning (accomplished by "anticipating" the intended action's attributes), or externally driven, such as when a stimulus event is perceived. The second, *integration phase* serves to bind the activated features together, hence, to integrate them into a sort of event file (Hommel, 1998). These event bindings may be actively maintained, such as when an action plan is held in preparation (Stoet & Hommel, 1999), or decay over time. In any case, however, event bindings seem to survive 1 s or longer (Hommel, 1998; Hommel & Colzato, 2004).

Here we applied TEC integration logic to action-effect integration. TEC claims that if the activations of codes (be they stimulus- or action-related) overlap in time, they get integrated. Hence, if the codes of an action plan are still activated to some degree when the effects of that action are coded, action and effects should become part of the same representational structure. Given that the codes of action plans commonly show activation 250 ms or longer after the corresponding action is carried out (Stoet & Hommel, 1999; Hommel, 1994), there are reasons to believe that the overlap is sufficient at least for immediate effects triggered by the action's onset. Indeed, studies of long-term action-effect acquisition have shown that actions and effects are spontaneously associated if the effects follow the action onset by up to 1 s but not longer (Elsner & Hommel, 2004)—at least if the action-effect interval is not "bridged" by intervening events (cf., Reed, 1999). Likewise, if participants estimate the extent their actions have caused a particular event, the accuracy of their judgments decreases considerably if actions and effects are separated by more than about 2 s (Shanks, Pearson & Dickinson, 1989). With respect to the short-term binding of stimuli and responses, it has been shown that stimuli are integrated with responses if they appear in a temporal neighborhood of about half a second but not if they are separated from the response by about 2.5 s (Hommel, 2005).

If actions and effects are spontaneously (i.e., non-intentionally) integrated into action-effect bindings and if these bindings have a lifetime beyond the presentation of the effect, the way they are bound together should affect subsequent performance. Assume, for instance, a left-hand keypress is heard to produce a low-pitched tone, in a task where high and low tones can appear and left and right keypresses are carried out. If the co-occurrence of left-hand keypress and low-pitched tone creates a binding between the codes LOW and LEFT, presenting a high or low tone shortly thereafter (i.e., while the binding is still intact) should systematically bias response selection. Figure 1 shows how. Given the high and low tones are the perceptual alternatives in the present context, the participant is likely to represent these two possibilities as
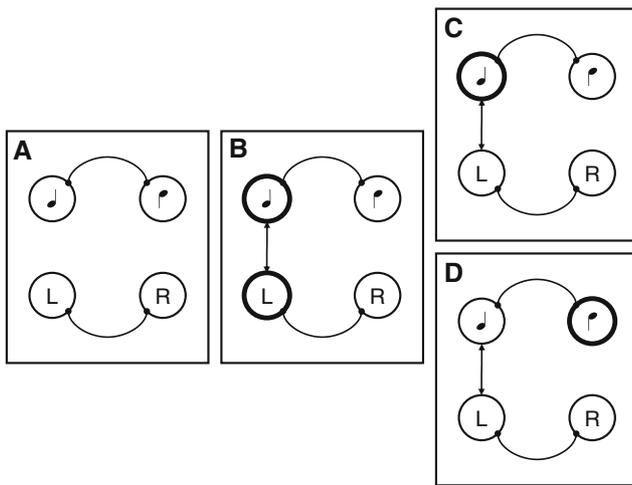
Fig. 1 Illustration of the creation and retrieval of action-effect bindings. **a** Being exposed to high- and low-pitched tones leads to the cognitive representation of these tones (low and high note for low and high tones, respectively), which given that the two tones are alternatives in the present context are connected by a mutually inhibitory link. Likewise, carrying out left and right responses leads to the representation of these (again mutually exclusive) alternatives (*L* and *R* for left and right responses, respectively). **b** Carrying out a left response followed by a low tone leads to the activation of the corresponding codes, which again leads to their integration (indicated by the *double arrow* between them). For the lifetime of the binding, the two codes to act as an unit. **c** Subsequently perceiving another low tone reactivates the corresponding code, which spreads activation to the left response code it is still integrated with. That is, a stimulus repetition primes a response repetition by biasing the competition between response codes toward the left code. **d** Subsequently perceiving the stimulus alternative (a high tone) activates the corresponding code, which will inhibit the code of the stimulus alternative (the low tone) via the inhibitory link. Given that the low tone is still integrated with the left response, this inhibition will spread to the left response code. Consequently, the competition between response codes is biased against the left code, so that stimulus alternation facilitates response alternation

sketched in Fig. 1a, where the codes for high and low tones are connected by an inhibitory link (see Bogacz, 2007). The same logic applies to the two alternative responses (the left and right keypress or L and R), which are also shown in this figure. If we assume that tones and responses vary independently and are thus uncorrelated, there are no long-term associations between tones and responses. However, according to our reasoning, a single co-occurrence of low tone and left response should induce a binding between their representations, as indicated in Fig. 1b.

What would happen if the tone repeats? As shown in Fig. 1c, activating the code of the low tone should prime the still bound response, the left keypress that is. This means that a stimulus repetition should induce a tendency to repeat the response as well. Now consider what a tone alternation would imply. As shown in Fig. 1d, presenting a high tone would activate the corresponding code, which is

not bound to any response (if we ignore previous trials for a moment). However, given the inhibitory link between the two tone representations, activating the code of the high tone should lead to the inhibition of the low-tone code. Given that this code is still bound with the left response code, inhibition will spread to that code as well. This follows from the integrated competition hypothesis suggested by Duncan and colleagues (Duncan, 1996; Duncan, Humphreys & Ward, 1997). They pointed out that the distributed cortical representation of perceptual and action codes calls for integration mechanisms that create coherent object-action compounds. Members of such a compound benefit from competitive gains achieved by other members of the same compound, so that, say, integrating RED with ROUND when processing the image of a cherry has the consequence that increasing the activation of the RED code also supports the ROUND code in its competition with other shape-related codes. The flipside of integrated competition is that losses in the competition also spread among members, so that outcompeting the RED code when seeing a banana somewhat later will also weaken the ROUND code associated with it. In other words, integrated elements win together and lose together. Applied to our example, this means that binding LOW and LEFT weakens LEFT if LOW loses against HIGH. Given that left and right responses are the only alternatives, this again implies that perceiving a high tone would bias response selection toward the right response, which would benefit from the indirect inhibition of the left response code.

Available evidence from stimulus-response integration studies provides support for both implications. For one, repeating stimulus features have been shown to speed up response repetitions as compared to response alternations (Hommel, 1998; Hommel & Colzato, 2004), suggesting that stimulus repetition indeed induces a response-repetition tendency. For another, alternations of stimulus features have been observed to speed up response alternations, sometimes even more than stimulus repetitions speed up response repetitions (e.g., Hommel & Colzato, 2004). Along the same lines, with multidimensional stimuli, response repetitions are particularly (i.e., over-additively) fast if signaled by a stimulus that repeats all the features of the previous stimulus (Bertelson, 1963), whereas response alternations are particularly slow under these circumstances (Hommel, Memelink, Colzato, & Zmigrod, 2008). Hence, stimulus alternations indeed seem to induce a response-alternation tendency.

According to these considerations perceiving a tone that does or does not match a just-experienced response-produced tone should systematically bias the decision to perform a left or right keypress. Importantly, this should be the case independently of previous experiences, hence,

Table 1 Conditions in Experiments 1 and 2

Condition	Induction $S \rightarrow R_1 \rightarrow E_A$		Test $E'_A \rightarrow R_2$
congruent	*****	$R_1 \rightarrow$ ♩ $R_1 \rightarrow$ ♩	♩ $\rightarrow$? ♩ $\rightarrow$?
incongruent	*****	$R_1 \rightarrow$ ♩ $R_1 \rightarrow$ ♩	♩ $\rightarrow$? ♩ $\rightarrow$?
no-go	*****	$R_1 \rightarrow$ ♩ $R_1 \rightarrow$ ♩	

Each trial consisted of two parts. In the induction part, a row of asterisks (Stimulus) triggered a free-choice response (R_1), which was followed by a randomly chosen high or low pitched tone (auditory action effect E_A). In congruent or incongruent go trials of the test part, a high or low pitched tone (which was congruent or incongruent with the preceding action effect) triggered another free-choice response (R_2)—the dependent measure (?) being the type of response (same as vs. different from R_1)

even if the overall probabilities for a high and low tone to follow a left or right response are equal. We tested this prediction as sketched in Table 1. Participants carried out free-choice responses by pressing a left or right key (for a discussion and validation of this technique, see Elsner & Hommel, 2001; Hommel, 2007). Each trial consisted of two parts. In the first, *induction* part participants made a freely chosen response (R_1) to a non-discriminative visual trigger stimulus (S). This response produced one of two auditory effects (E_A), a low- or a high-pitched tone. Importantly, the mapping of response keys to pitch varied randomly from trial to trial, so to prevent any incremental response-effect learning across the experimental session. One-second later, in the *test* part of each trial, participants encountered one of the two effect stimuli (E'_A), which now served as go signal (in 75% of the trials) to perform another freely chosen response (R_2). The measure of interest was the response choice in the test part (i.e., R_2). In particular, we analyzed the tendency to repeat the previous response ($R_2 = R_1$) as a function of the relationship between the effect tone E_A and the go-signal tone E'_A. According to our hypothesis, participants should be more likely to repeat a response if the two tones match ($E'_A = E_A$) than if the tones do not match ($E'_A \neq E_A$), because the tone's code should still be bound with the response that just had produced it.

Experiment 1

Experiment 1 was conducted as a first test whether action-related codes are spontaneously integrated with codes of their effects, as suggested by TEC. If so, we would expect response-repetition rates (%RR) to be higher if the R_2-go signal (E'_A) matches the preceding action effect (E_A) in pitch than if it does not.

Method

Participants

Twenty students served as paid participants. As was the case for all participants of this study, they reported having normal or corrected-to-normal vision and audition and were not familiar with the purpose of the experiment.

Apparatus and stimuli

Visual stimuli (a row of 13 white-on-black asterisks) were presented on a computer monitor and auditory stimuli (sinusoidal tones of 400 and 800 Hz) through external loudspeakers to the monitor's left and right. Responses were made by pressing the left or right of two external microswitches with the corresponding index finger. The experiment was controlled by a standard PC running under ERTS (Beringer, 1994).

Procedure

Each trial consisted of an *induction* part, to induce a particular action-effect binding, and a *test* part, to diagnose the presence of such bindings. Table 1 shows the sequence of events. After an intertrial interval of 3,000 ms, the asterisk string (S) appeared for 300 ms, requesting a speeded left or right keypress (R_1). Participants were instructed to choose the key randomly and to avoid any strategy apart from using the keys about equally often. If a response was made a randomly selected effect tone (E_A) was presented for 100 ms, its onset being synchronized with the keypress. Due to the random selection procedure, keypresses and tone pitches were uncorrelated, that is, in a given trial each keypress had the same probability to trigger either a low or high tone. Participants were told that these tones were completely irrelevant for the task and that there would be no systematic relationship between keypress and pitch.

In the second, test part of each trial one of the two effect tones was used as go signal (E'_A) to signal a second free-choice reaction (R_2) in 75% of the trials; in the remaining 25% no tone appeared and no second response

was to be performed (no-go trials). No-go trials were used to work against some of the most obvious strategies in free-choice tasks, such as choosing responses according to a standard predetermined pattern. In go trials one of the two tones sounded for 100, 1,000 ms after the previous effect tone had been presented. In 50% of these go trials the tone was the same as the previous effect tone (congruent trial); in the other 50% the signal tone was the alternative tone (incongruent trial). Participants were instructed to respond to the tone as quickly and as spontaneously as possible by pressing a randomly chosen response key and to refrain from responding if no second tone would occur. It was emphasized that only the *presence* of a tone mattered for the execution of R_2 while its *pitch* would be neither relevant nor informative. Participants were also urged to use both keys and not to apply any strategy. The program waited up to 1,500 ms for a response. Responses with reaction times exceeding 1,500 ms were counted as missing, those faster than 100 ms as anticipation, and responses in no-go trials as false alarms. All these errors were fed back to the participants. Following ten randomly drawn practice trials three blocks of 64 randomly ordered trials each were administered. After the session participants were asked whether they had obeyed to the instruction and had guessed the purpose of the experiment.

Results and discussion

Our dependent measure of choice is very sensitive to individual strategies, which may conceal or even prevent the possible impact of go stimuli on response choices. Particularly damaging would be strategies that determine response choices long before the go stimulus is presented, so that the selection process we intended to bias is already completed. Accordingly, we not only took measures to work against some of the strategies by speeding response selection and including no-go trials, but we also excluded participants that were likely to apply a particular "pre-selection" strategy. For this reason, we only considered participants who produced less than 20% false alarms and at least 90% correct trials altogether, and who did not report having used a response rule. All participants passed these criteria and no-one reported having paid any attention to pitch or having guessed the purpose of the experiment. In fact, most of them believed that reaction time was the important dependent variable. We also excluded participants if their mean %RR was lower than 10% or higher than 90%, which we consider strong evidence of an alternation or repetition strategy, respectively. This applied to two participants. After excluding trials with response omission (0.3%) or anticipation (0.4%) individual %RRs

were calculated as a function of congruency (see Table 1 for the coding scheme).[1]

In the induction part of the trials the two keys were pressed equally often and their frequencies (48.5 vs. 51.5%) did not differ from chance. This observation, which we also made in the following experiments, confirms that participants experienced all possible response-effect couplings about equally often. The mean %RR in the test part was 39.1%, but the repetition rate was modulated by $E_A - E'_A$ congruency: As shown in Table 2, congruent trials produced more response repetitions than incongruent trials, $t(17) = 4.86$, $p < 0.01$. That is, as expected, stimulus repetitions were associated with more response repetitions, suggesting that the present response choice was affected by the relationship between the previous response and its auditory effect.

It is interesting to note that the response-repetition frequencies for congruent and incongruent conditions were not distributed evenly around 50% but shifted toward response alternations (i.e., around 39.1%). There are at least two possible accounts for this observation, which we will also make in the following experiments. The first account considers that people are often biased toward response alternations, as can be seen in faster reaction times with response alternations that repetitions, presumably reflecting a general misconception about statistical probability (Bertelson, 1961; Soetens, Boer, & Hueting, 1985)—also known as gambler's fallacy. Interestingly, response alternations were faster than repetitions (335 vs. 357 ms) in the present experiment as well, $t(17) = 3.01$, $p < 0.01$. Hence, even though our study does not provide a "pure" measure of the alternation bias, the fact that it has been so often observed in other studies may be taken to suggest that our participants also showed such a bias. This again might suggest that our congruent and incongruent conditions were indeed symmetrically distributed around a mean that would

---

[1] We did not consider reaction times (with one exception in another context below) because, as demonstrated and discussed by Elsner and Hommel (2001), it is impossible to predict and interpret their pattern in the present free-choice task. For instance, fast responses in the congruent condition may indicate that (a) the response was particularly spontaneous, and thus free from strategic considerations, suggesting that conditions were particularly good for tones to affect response selection; or that (b) the response was preplanned already and thus under full strategic control, suggesting that conditions were particularly bad for tones to affect response selection. Slow responses in the congruent condition may be interpreted to reflect (a) particularly strong contributions from strategic considerations (a lot of thinking before responding), which one would expect to minimize the impact of the tone; but they may just as well reflect (b) a delay due to extended competition between strategic top-down factors and tone-driven bottom up factors, which again would point to a particularly strong contribution from the tone. Multiple interpretations also exist for fast or slow responses in the incongruent condition, all suggesting that reaction times are a misleading measure in the present context.

Table 2 Mean response repetition frequencies (in %) and response-repetition biases (congruent–incongruent) for Experiments 1–4 as a function of $E - E'$ congruency (match between effect of R_1 and go signal for R_2), action-effect modality, and modality of the task-relevant go signal for R_2

	Auditory action effect (pitch)			Visual action effect (color)		
	Congruent	Incongruent	Bias	Congruent	Incongruent	Bias
Experiment 1						
Auditory go signal						
	43.5 (12.1)	34.6 (12.1)	8.9**			
Experiment 2						
Auditory go signal						
	49.1 (12.3)	41.9 (15.0)	7.1**			
Experiment 3						
Auditory go signal						
	47.6 (22.1)	40.5 (19.9)	7.1**			
Visual go signal						
	50.1 (21.6)	46.9 (18.4)	3.2**			
Experiment 4						
Auditory go signal						
	46.4 (14.3)	41.4 (13.6)	5.0**	42.9 (12.7)	44.9 (14.4)	−2.0 NS
Visual go signal						
	47.1 (19.9)	41.9 (17.6)	5.2**	46.3 (19.1)	42.7 (18.3)	3.6**

Standard deviations of means are given in parentheses

NS non-significant bias

** Significant bias, $p < 0.01$

ideally approach 50%, but the whole distribution was torn to the lower half because of the gambler's fallacy.

The second account holds that our participants were not biased toward repetition or alternation in principle, and that the outcome for the congruent condition represents something like a neutral baseline. Indeed, it will turn out that the 43.5% we observed in the congruent condition of Experiment 1 is the lowest estimate of the present study, and that the other experiments will produce estimates very close to 50%. If so, the main impact of stimulus repetitions and alternations would consist in stimulus alternations biasing people toward response alternations. In other words, the effect sketched in Fig. 1d would be much stronger than the one in Fig. 1c. As mentioned earlier, this would fit with occasional observations that stimulus-response alternations produce faster and more accurate responses than complete stimulus-response repetitions, at least numerically (e.g., Colzato, Fagioli, Erasmus, & Hommel, 2005; Colzato, van Wouwe, & Hommel, 2007; Hommel & Colzato, 2004). Indeed, given that repetition-induced priming of previous bindings and alternation-induced integrated competition are different types of processes, there is no reason to believe that the reaction-time benefits they produce should be of exactly the same size.

As we neither have a pure measure of possible general alternation biases nor a noise-free measure of binding reactivation and integrated competition, it is premature to try deciding between these two interpretations. Importantly, however, they both rest on the same assumption, namely, that perceiving a self-produced stimulus event creates a temporary binding of the codes underlying the action and the codes representing the perceived event. As a consequence, perceiving the same event or its alternative systematically biases response selection. Taken altogether, Experiment 1 provides first evidence for our hypothesis that a single pairing of an action and an effect is sufficient to integrate their cognitive representations, and that this integration has a systematic effect on subsequent response selection.

Experiment 2

Even though the outcome of Experiment 1 is consistent with our expectation that action-effect binding affects subsequent response-selection processes, there is an alternative interpretation. Our participants had the task of producing random responses and response sequences, which is known to be very hard to do. One way to make this task easier and to still meet the task requirement of getting close to a 50:50 distribution of response repetitions and alternations would be to strategically repeat the response whenever the stimulus repeats. Note that this strategy only works well if the probability of stimulus repetition versus alternation in go trials is also 50:50. If this ratio would be drastically changed, such as if stimulus alternations would be much more frequent than stimulus repetitions, such a matching strategy would be bound to fail: either response alternations would now also become much more frequent than response repetitions or participants would notice that a matching strategy makes little sense and simply no longer apply it. This was the logic underlying Experiment 2, which repli-

cated Experiment 1 with a 25:75 probability of stimulus repetitions and alternations. According to a strategic interpretation of the congruency effect, this manipulation should eliminate the effect, whereas an interpretation in terms of action-effect binding predicts the same outcome as in Experiment 1.

Method

Twenty-one students served as paid participants. The method was exactly as in Experiment 1 with only one exception: the go trials of the test phase did not consist of 50% congruent and 50% incongruent conditions but of 25% congruent and 75% incongruent conditions; i.e., the trigger tone matched the previous action effect tone in only one quarter of the trials.

Results and discussion

Applying the same criteria as in Experiment 1 led to the exclusion of one participant. Again, trials with response omissions (<0.9%) and anticipations (<0.3%) were excluded. The overall response-repetition rate in the test part was 43.7%, which is higher than in Experiment 1 and statistically no longer different from chance. Clearly, this observation does not support the idea that participants might have strategically matched the response repetition rate to the stimulus repetition rate. The response repetition rate was again modulated by $E_A - E'_A$ congruency, $t(19) = 3.07$, $p < 0.01$, due to that congruent trials produced more response repetitions than incongruent trials did (see Table 2). An ANOVA on the combined data from Experiments 1 and 2 did not yield any hint to an interaction between experiment and congruency effect, $p > 0.5$, confirming that the congruency effect was equivalent in the two experiments. Taken together, these findings suggest that the congruency effect does not reflect a deliberate response selection strategy but rather represents an automatic by-product of action-effect binding.

Experiment 3

Our next experiment was conducted to see how automatic the impact of action-effect bindings really is and whether, or to what degree it is sensitive to attentional effects, that is, to the task relevance of the effect's perceptual characteristics. In Experiments 1 and 2, tones were used as both action effects (E_A) and R_2-go signals (E'_A). Accordingly, although neither the pitch nor the presence of the action effect was of any relevance, tones did play an important role and could not be ignored entirely. It may have been this, somewhat indirect type of task relevance that drew

sufficient attention to the effects to integrate and bind them with the responses and/or to retrieve the just-bound response when the effect stimulus was encountered again.[2] If so, it should be possible to reduce or eliminate the impact of action-effect bindings on ongoing response selection by defining the R_2-go signal in another than the auditory modality, so that tones are no longer of any relevance. This is what we did in Experiment 3. In an auditory-go condition we replicated Experiment 1 by using again an auditory E'. But we also ran a visual-go condition, where the R_2-go signal was a visual stimulus (E'_V). Although no longer of any relevance for the task, the tone was still presented as E'_A, thus accompanying the visual go signal in go trials and as the only stimulus in the test part of no-go trials. If task relevance affected the creation and/or retrieval of action-effect bindings we would expect the response-rate effect—that is, higher response-repetition rates if the R_2-go signal matches the preceding action effect—in the auditory-go condition but not (or less so) in the visual-go condition.

Method

Another 26 female and 19 male students were randomly assigned to two groups of 23 and 22 participants, respectively. For the first, auditory-go group the method was exactly as in Experiment 1. For the second, visual-go group several modifications were introduced. The relevant go signal in the test part of each trial was not a tone but a red 3x3 cm square (E'_V)[3] appearing for 300 ms at screen center. Just like the tone in the auditory-go group, the square was presented in 75% of the trials to signal R_2 (go trials) and participants were to withhold R_2 in the remaining no-go trials. The pitch of the tone matched the previous action-effect tone in 50% of the go trials and the alternative tone in the other 50%. It was pointed out to the participants that both presence and pitch of the tone would be completely irrelevant for the task.

---

[2] Note that the design of our study (or of any other analysis of sequential effects) does not allow disentangling possible effects on the creation of action-effect bindings (i.e., integration) and on their retrieval. Obviously, bindings can only be retrieved if they have been created earlier, so that changes in the impact of action-effect bindings on performance may be due to changes in the binding process, changes in the retrieval process, or both. We will get back to this issue in "General discussion".

[3] In keeping with the terminology introduced in Experiment 1 we refer to the visual go signal as E'_V. However, note that in the induction part of Experiment 3 there were only auditory action effects (E_A) but no visual action effects (E_V).

Results and discussion

Applying the same criteria as in Experiment 1 led to the exclusion of three members of the auditory and of two of the visual group. Again, trials with response omissions (<0.6%) and anticipations (<0.1%) were excluded. The overall response-repetition rate in the test part was 46.3%, which is almost the same as in Experiment 2 and statistically not different from chance. Mean %RR were analyzed as a function of $E_A - E'_A$ congruency (i.e., whether the two tones in each trial matched or not) and R_2-go-signal modality (i.e., whether R_2 was carried out in response to the second tone or a color square; see Table 2). The only reliable finding was a main effect of tone congruency, $F(1,38) = 8.69$, $p < 0.005$, while the interaction with go-signal modality was far from significance, $F(1,38) < 1$. That is, irrespective of the tone's task relevance, R_1 is repeated more often if the tone in the test phase matches the action-effect tone.

All in all, the outcome of Experiment 3 is somewhat mixed. Statistically speaking, task relevance had no impact on the congruency effect, suggesting that the auditory action effects were integrated and retrieved in either go-signal condition. However, numerically the induced bias in the visual condition was not even half as big as that obtained in the auditory condition. Moreover, auditory stimuli and their impact on perceptual processing have been demonstrated to be more salient, hence, to rely much less on attention than visual stimuli (Posner, Nissen, & Klein, 1976). One therefore may argue that Experiment 3 provides a rather conservative test of the impact of attention.

Experiment 4

To provide a more sensitive test we ran Experiment 4, where responses produced both auditory and visual effects. We also presented stimuli of both modalities in the test part of the go trials and varied their relevance. In one block, R_2-go signals were auditory, just like in Experiments 1–3, which rendered the visual stimuli in either part of the trial irrelevant. The saliency hypothesis suggests that auditory action effects should be integrated under such conditions while visual effect may not. If so, repeating tone pitch ($E'_A = E_A$) should lead to higher response-repetition rates than alternating pitch, whereas repeating color ($E'_V = E_V$) should yield the same rates as alternating color. In another block, R_2-go signals were visually defined, which rendered the auditory stimuli in either part of the trial irrelevant. According to the saliency hypothesis and in view of Experiment 3, we would expect that both auditory and visual action effects are integrated and retrieved, so that response-repetition rates should depend on whether pitch or color is repeated.

Method

Another 27 female and 23 male students served as paid volunteers. The method was as in Experiment 3 (visual-go group) with the following exceptions. With regard to the induction part, performing R_1 now caused the simultaneous presentation of a low- or high-pitched tone (for 100 ms) *and* a red or green square at screen center (for 200 ms); hence, each R_1 had both an auditory and a visual effect.

In the test part of the trials, three independent variables were manipulated: the modality of the R_2-go signal (tone or square), the congruency between the pitch of the action-effect tone from the induction part (E_A) and the pitch of the tone presented in the test part (E'_A), and the congruency between the color of the action-effect square from the induction part (E_V) and the color of the square presented in the test part (E'_V). Like in the visual-go group of Experiment 3, there were two stimuli in the test part of go trials, a low- or high-pitched tone and a red or green square. However, in a given block only one of them was task-relevant by virtue of signaling a go trial, whereas the other was entirely irrelevant.

The experimental session consisted of two blocks, an auditory-go block, where no-go trials were defined by the absence of a tone in the test part of the trial, and a visual-go block, where no-go trials were defined by the absence of a square in the test part of the trial. Block order was balanced across participants. Each block was composed of 10 randomly drawn practice trials and 192 randomly ordered experimental trials. The 192 experimental trials comprised 144 go and 48 no-go trials, so that the go probability was again 75%. The 144 go trials were composed of 36 trials in which both the tone and the color presented in the test part matched the action effects of the preceding induction part (i.e., $E'_A = E_A$ and $E'_V = E_V$), 36 trials in which only the auditory stimuli matched (i.e., $E'_A = E_A$ and $E'_V \neq E_V$), 36 trials in which only the visual stimuli matched (i.e., $E'_A \neq E_A$ and $E'_V = E_V$), and 36 trials in which neither the auditory nor the visual stimuli matched (i.e., $E'_A \neq E_A$ and $E'_V \neq E_V$). To ensure that participants registered the visual action effects even in the auditory-go block, the instruction emphasized that they should always fixate the center of the computer screen.

Results and discussion

Applying the same criteria as in Experiment 1 led to the exclusion of 10 participants. Trials with response omissions (<0.2%) and anticipations (<0.1%) were excluded. Mean %RR were calculated for each participant as a function of auditory ($E'_A - E_A$) congruency, visual ($E'_V - E_V$) congruency, and modality of the relevant go signal (see Table 2). A corresponding $2 \times 2 \times 2$ ANOVA produced

two significant results: a main effect for auditory congruency, $F(1,39) = 17.98$, $p < 0.001$, and an interaction of visual congruency and go-signal modality, $F(1,39) = 9.26$, $p < 0.005$. As Table 2 shows, congruent pitch yielded a higher rate of response repetitions independently of go-signal modality whereas congruent color affected the repetition rate only if go trials were defined by the presence or absence of visual stimuli. Indeed, separate t-tests revealed a highly significant effect of color congruency in the visual-go block, $t(39) = 3.27$, $p < 0.005$, but not in the auditory-go block, $t(39) = 1.39$, $p > 0.05$. This pattern supports an account in terms of stimulus saliency: Action effects are integrated and retrieved if they are either relevant to the task or salient enough to attract attention in a bottom-up fashion.

Another important finding of Experiment 4 is that it for the first time demonstrates the integration of multiple action effects. Although previous studies on action-effect acquisition employed a variety of to-be-learned effect stimuli they were always restricted to one type of stimulus at a time. Yet, both auditory and visual action effects influenced performance in the visual-go condition of Experiment 4, which suggests that participants had integrated their actions with both pitch and color.[4]

General discussion

Our study tested a TEC approach to the integration of actions and their effects. We hypothesized that the likely temporal overlap of activation of action- and effect-related codes induces the temporary binding of those codes. This binding may affect subsequent behavior by biasing it toward response repetition in case of a stimulus repetition, biasing it toward response alternation in case of stimulus alternation, or both. Consistent with this expectation, Experiment 1 showed that varying the pitch of a go signal systematically affects the tendency to repeat or alternate the response that was just experienced to produce a tone of that pitch. Together with the outcome of Experiment 2, which rules out a strategic interpretation of the response repetition bias, this suggests that codes of that action are still bound with codes of the tone it produced. As a consequence, re-activating the tone-related code spread activation to the corresponding action-related code, thus priming the previous action as indicated in Fig. 1c, while activating the alterna-

tive tone code led to the inhibition of the codes of both the previous tone and the previous response, resulting in a preference for response alternation (Fig. 1d). Interestingly, the distribution of response-repetition frequencies was shifted toward response alternation in all experiments. This might reflect a general impact of the gamblers fallacy and represent the same bias that has been shown in studies of sequential stimulus and response effects (Bertelson, 1961; Soetens et al., 1985). Alternatively, it might indicate that alternations of action-effect stimuli bias subsequent response selection more toward response alternations than effect repetitions bias selection toward response repetitions. In other words, the priming of response repetitions as sketched in Fig. 1c may be less efficient than the inhibition of response repetitions as sketched in Fig. 1d. The present study does not allow disentangling these two possibilities, which calls for a more detailed experimental analysis. However, both possibilities imply that actions and effects are spontaneously integrated into temporary bindings, which supports our main hypothesis.

As the action effect in Experiment 1 was not relevant or informative, effect integration seems to be spontaneous in the sense that it does not require the explicit intention to learn about those effects. This supports Elsner and Hommel's (2001) assumption that effect integration is an automatic by-product of moving and acting. However, this does not mean that goals and intentions, and the attentional set they bring about, have no impact on effect integration and/or retrieval (Hommel et al., 2001). To the contrary, Experiments 3 and 4 provide evidence that the likelihood with which action-effect bindings affect performance depends on both bottom-up and top-down attentional factors. If an effect is salient enough to attract attention in a bottom-up fashion, as can be assumed for tones (Posner et al., 1976), action effects impact behavior even if they are neither directly nor indirectly related to the task at hand. This fits well with observations from learning studies, where auditory (e.g., Hoffmann et al., 2001; Hommel, 1996) and electrocutaneous (Beckers et al., 2002) action effects were spontaneously acquired and retrieved in otherwise purely visual-manual tasks. Less salient effects, however, such as visual effects in an otherwise auditory-manual task, seem to depend more on the fit of their attributes with the current attentional set. Even though we manipulated saliency by contrasting auditory and visual effects, we consider saliency to be a matter of degree and of the particular stimulus-context relations rather than an absolute characteristic of a particular sense modality; but more systematic studies are necessary to elucidate that issue.

Another question is which process exactly is affected by saliency. One possibility is that integration proper depends on some minimal activation of effect codes, which they may reach only if they are either top-down primed because of

---

[4] Alternatively, participants might have alternated between the integration of auditory and visual effects. However, this should have decreased the response bias for both types of effects, which does not fit the finding that the bias for auditory action effects in the visual-go condition was numerically stronger in Experiment 4 (where alternation might have taken place) than in Experiment 3 (where action effects were all auditory).

their task relevance or particularly salient (see Hommel, 2004). However, a major disadvantage of a selective integration mechanism would be that infants, children, and adults facing a novel task would no longer be able to pick up unpredicted but consistent action-effect relations on the fly—a characteristic of action-effect acquisition that one may consider essential for the development of voluntary action and action skills (Elsner & Hommel, 2001; James, 1890; Piaget, 1946). Another possibility is that action-effect binding is truly automatic and may not even be sensitive to the availability of attentional resources, which would leave binding retrieval as a possible target of our saliency manipulations. Indeed, we cannot exclude that saliency affected the retrieval of just-created action-effect bindings rather than the creation of bindings. That is, stimuli may be more effective to trigger the retrieval of previously created bindings if they are task-relevant or salient. Again, studies on stimulus-response integration suggest that the retrieval of bindings is more sensitive to attentional manipulations than the creation of bindings is (Hommel et al., 2008), which would fit better with a retrieval-based interpretation of saliency effects. Nevertheless, the final word on this matter presupposes a better understanding of how action-effect binding and retrieval processes work, and how they are controlled.

In view of the previous demonstrations of the acquisition of stable action-effect associations on the one side and of the present evidence for transient bindings between actions and effects on the other, it would be tempting to assume that the latter are functional predecessors of the former: The transient coupling of action and effect codes may reflect the presence of reverberatory loops in the sense of Hebb (1949), which again may serve to establish and consolidate more enduring cell assemblies. In other words, binding may represent the first step to long-term memory (cf., Raffone & Wolters, 2001, but see Colzato, Raffone, & Hommel, 2006). However, in the absence of clear-cut evidence that action-effect learning is impossible without binding (and in view of the major methodological challenges demonstrations of such evidence would need to overcome) this is no more than an interesting speculation.

References

Beckers, T., De Houwer, J., & Eelen, P. (2002). Automatic integration of non-perceptual action effect features: The case of the associative affective Simon effect. *Psychological Research, 66*, 166–173.

Beringer, J. (1994). ERTS: A flexible software tool for developing and running psychological reaction time experiments on IBM PCs. *Behavior Research Methods. Instruments & Computers, 26*, 368–369.

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology, 13*, 90–102.

Bertelson, P. (1963). S-R relationships and reaction times to new versus repeated signals in a serial task. *Journal of Experimental Psychology, 65*, 478–484.

Bogacz, R. (2007). Optimal decision-making theories: Linking neurobiology with behavior. *Trends in Cognitive Sciences, 11*, 118–125.

Colzato, L. S., Fagioli, S., Erasmus, V., & Hommel, B. (2005). Caffeine, but not nicotine enhances visual feature binding. *European Journal of Neuroscience, 21*, 591–595.

Colzato, L. S., Raffone, A., & Hommel, B. (2006). What do we learn from binding features? Evidence for multilevel feature integration. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 705–716.

Colzato, L. S., van Wouwe, N. C., & Hommel, B. (2007). Feature binding and affect: Emotional modulation of visuo-motor integration. *Neuropsychologia, 45*, 440–446.

Duncan, J. (1996). Cooperating brain systems in selective perception and action. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI* (pp. 549–576). Cambridge, MA: MIT Press.

Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology, 7*, 255–261.

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

Elsner, B., & Hommel, B. (2004). Contiguity and contingency in the acquisition of action effects. *Psychological Research, 68*, 138–154.

Elsner, B., Hommel, B., Mentschel, C., Drzezga, A., Prinz, W., Conrad, B., et al. (2002). Linking actions and their perceivable consequences in the human brain. *NeuroImage, 17*, 364–372.

Hebb, D. O. (1949). *The organization of behavior: A neurophysiological theory.* New York: Wiley.

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. (1994). Spontaneous decay of response code activation. *Psychological Research, 56*, 261–268.

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

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition, 5*, 183–216.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*, 494–500.

Hommel, B. (2005). How much attention does an event file need? *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1067–1082.

Hommel, B. (2007). Feature integration across perception and action: Event files affect response choice. *Psychological Research, 71*, 42–63.

Hommel, B., & Elsner, B. (2009). Acquisition, representation, and control of action. In E. Morsella, J. A. Bargh & P. M. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 371–398). New York: Oxford University Press.

Hommel, B., & Colzato, L. S. (2004). Visual attention and the temporal dynamics of feature integration. *Visual Cognition, 11*, 483–521.

Hommel, B., Memelink, J., Colzato, L. S., & Zmigrod, S. (2008). Attentional control of the creation and retrieval of stimulus-response bindings (submitted).

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.

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

Jolicœur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36*, 138–202.

Lotze, R. H. (1852). *Medicinische Psychologie oder die Physiologie der Seele*. Leipzig: Weidmann'sche Buchhandlung.

Melcher, T., Weidema, M., Eenshuistra, R. M., Hommel, B., & Gruber, O. (2008). The neural substrate of the ideomotor principle: An event-related fMRI analysis. *NeuroImage, 39*, 1274–1288.

Piaget, J. (1946) *La formation du symbole chez l'enfant*. Delachaux & Niestlé.

Posner, M. I., Nissen, J. J., & Klein, R. M. (1976). Visual dominance: An information processing account of its origins and significance. *Psychological Review, 83*, 157–171.

Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience, 13*, 766–785.

Reed, P. (1999). Role of a stimulus filling an action-outcome delay in human judgments of causal effectiveness. *Journal of Experimental Psychology: Animal Behavior Processes, 25*, 92–102.

Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989). Temporal contiguity and the judgment of causality by human subjects. *Quarterly Journal of Experimental Psychology, 41B*, 139–159.

Soetens, E., Boer, L. C., & Hueting, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception & Performance, 11*, 598–616.

Stoet, G., & Hommel, B. (1999). Action planning and the temporal binding of response codes. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1625–1640.

Ziessler, M. (1998). Response-effect learning as a major component of implicit serial learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 962–978.

Ziessler, M., Nattkemper, D., & Frensch, P. A. (2004). The role of anticipation and intention for the learning of effects of self-performed actions. *Psychological Research, 68*, 163–175.