On the dynamic interplay between perception and action - a connectionist perspective

Haazebroek, P.

Author: Haazebroek, Pascal
Title: On the dynamic interplay between perception and action : a connectionist perspective
Issue Date: 2013-12-11

Chapter 4
Automaticity

This chapter is an integration of major parts of the following articles:

Haazebroek, P., Raffone, A., & Hommel, B. *HiTEC: A Connectionist Model of the Interaction between Perception and Action Planning.* Manuscript submitted for publication.

Haazebroek, P., van Dantzig, S., & Hommel, B. (2011a). A computational model of perception and action for cognitive robotics. *Cognitive Processing, 12*, 355-365.

Haazebroek, P., van Dantzig, S., & Hommel, B. (2011b). Interaction between Task Orient-ed and Affective Information Processing in Cognitive Robotics. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 59,* 34-41.

Haazebroek, P., van Dantzig, S., & Hommel, B. (2009). Towards a computational account of context mediated affective stimulus-response translation. P*roceedings of the 31st Annual Conference of the Cognitive Science Society (pp. 1012-1017)*. Austin, TX: Cognitive Science Society.

Traditional views on human information processing hold that responding to stimuli in our environment follows a sequence of separable stages of processing (e.g., Donders, 1868; Neisser, 1967; Sternberg, 1969) from stimulus perception, to decision making, up to response execution. Numerous empirical findings, however, have demonstrated that parts of human information processing do not seem to involve conscious cognitive decision making. Features of perceived objects (such as location, orientation, and size) can influence actions *directly* and beyond (tight) cognitive control, as illustrated by stimulus–response compatibility (SRC) phenomena (for general overviews, see Hommel & Prinz, 1997; Prinz & Hommel, 2002; Proctor & Vu, 2006), such as the Simon effect (Simon & Rudell, 1967) as simulated in Simulation 3.

To account for both controlled and automatic processing, various dual route process accounts have been proposed (e.g., Zorzi & Umilta, 1995; Kornblum, et al., 1990; but see Hasbroucq & Guiard, 1991, for a strictly perceptual account). These accounts propose that there is, next to the first cognitively controlled route, a second, direct route from perception to action that can bypass cognition, as explicitly modeled in various computational models of the Simon effect. Essentially, dual route accounts consider the observed direct stimulus-response interaction as an exception requiring an additional route. Moreover, they typically do not address the reason *why* some stimulus features directly influence action and others do not.

In this chapter we attempt to explain how and why automaticity occurs in the HiTEC connectionist model (see Chapter 2). We explicitly address how representational and processing characteristics of HiTEC *inevitably* lead to SRC effects. Here, common codes play a crucial role. Building upon this notion of common codes, HiTECs structure and processes allow stimulus features, both task relevant and task irrelevant, to be registered, processed and translated into responses. In this endeavor we focus on two key paradigms. In Simulation 3, a HiTEC instance is constructed to simulate the Simon task. In Simulation 4, we model the Stroop effect. As HiTEC treats stimulus and response representation in a similar way, it is to be expected that a model instance similar to the one used in Simulation 3 would be able to account for the Stroop effect as well. The empirical findings accounted for in this chapter have been modeled before by other (dedicated) computational models. We conclude this chapter with a comparison of some of these models with our approach.

Simulation 3: Simon effect

Original experiment

Simon and Rudell (1967) showed that people respond faster to stimuli if the location of the stimulus is compatible with (corresponds to) the response location, even when stimulus location is not task relevant. In the standard Simon task, stimuli with a non-spatial stimulus feature (e.g., auditory pitch) are presented at different locations (e.g., left or right). Participants are instructed to respond to the non-spatial feature by giving a spatially defined response (e.g., pressing a left or right key). Even though the location of the stimulus is not relevant

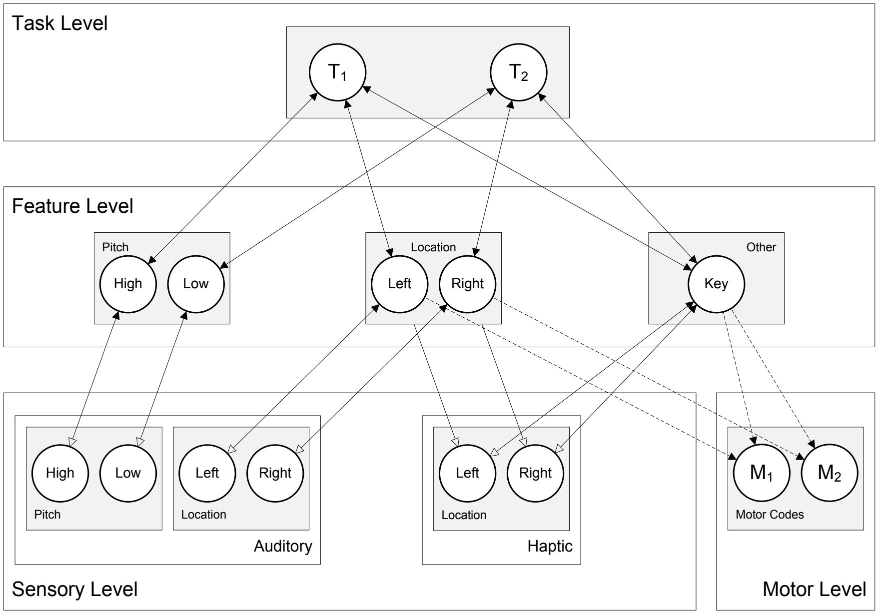


Figure 15. Specific HiTEC Model for Simulation 3. Feature codes are present for stimulus pitch and location. Note that location feature codes are used for encoding both stimulus location and response location. The task instruction is already internalized before presenting the learning trials. This biases the learning of connections between feature codes and motor codes. Note that in principle any feature code can be connected to any motor code. However, only some of them actually become (strongly) weighted reflecting the specific perceptual regularities.

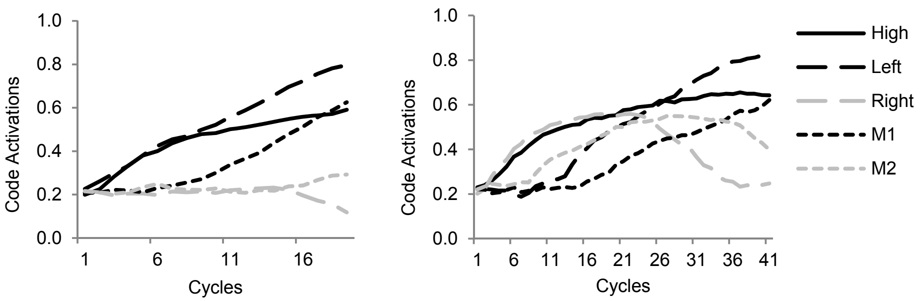
for this task, performance is facilitated when the chosen response corresponds spatially to the stimulus location.

HiTEC simulation

The Simon effect was modeled in HiTEC using sensory codes for auditory pitch[7], auditory locations and haptic locations. At the feature level there are feature codes for pitch, location and for 'Key'. The model, as shown in Figure 15, contains two motor codes, 'M1' and 'M2', representing pressing the left and the right key. During the learning phase, 'M1' and 'M2' are activated alternately and their respective action effects are presented to the model. As a result, associations are learned selectively between the motor codes and the 'Left' and 'Right' feature codes.

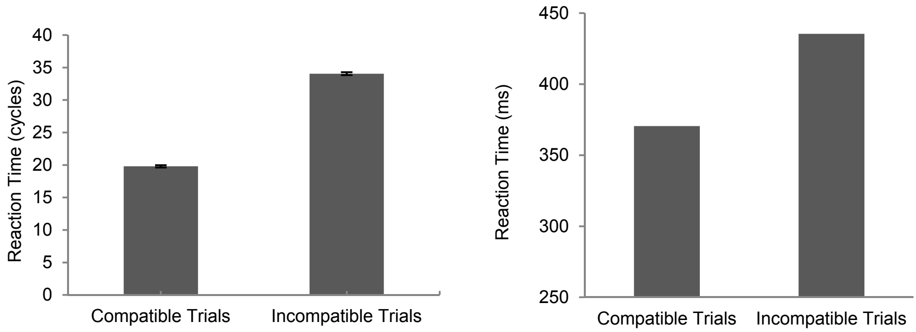
In the experimental trials, tones are presented and are responded to by anticipating and executing left or right keypresses (i.e., by activating 'Left' or 'Right' feature codes respectively). Crucially, the 'Left' and 'Right' feature codes are also activated when the tone stimulus is presented on the left or right, yielding a compatibility effect as demonstrated in Figure 16 and as reflected in the results. Because 'Left' and 'Right' are features that are relevant for

[7] We decided to simulate the auditory version of the Simon task, rather than the more common visual version, because that will make it easier for the reader to relate it to the auditory version of the Simon task that we modeled in Chapter 5. However, the logic of our modeling applies to visual versions just as well.



(a) (b)

Figure 16. Time courses of feature code and motor code activations in the experimental trials of Simulation 3. Panel A depicts the activations in the compatible condition. Here 'M1' reaches threshold in 19 cycles. Panel B depicts the dynamics in the non-compatible condition. Here 'M1' reaches threshold in 41 cycles. In the latter condition, activating 'Right '(as stimulus feature) biases the model into planning a 'right' action. This, however, is overcome due to the task connections so that 'Left' becomes stronger and eventually wins over 'Right'. Similarly, first the incorrect motor response, 'M2' becomes active, but eventually 'M1' reaches threshold. In effect, the model takes longer to respond in the non-compatible condition than in the compatible condition. Activations of the remaining feature codes, task codes and sensory codes are omitted for sake of clarity.



(a) (b)

Figure 17. Results of Simulation 3 compared with behavioral data (adopted from Simon & Rudell, 1967), showing average reaction time means and standard deviations. Human variance data was not available.

response coding, they are part of the task connections. As a consequence, stimulus location becomes of influence in the overall stimulus-response translation. As shown in Figure 16, in the compatible condition, the stimulus location already activates the correct spatial feature code and thereby speeds up response selection, on average.

Conversely, in the incompatible condition, stimulus location activates the wrong spatial feature code, which also already activates the wrong motor code. Meanwhile, however, the stimulus pitch is translated — through the task codes — into the correct spatial feature codes and the correct motor code. This latter pathway typically overcomes the head start due to the overlap-pathway, but the code overlap does slow down the overall translation as reflected in the results.

Simulation results

In the simulations (15 simulated subjects, each performing 20 trials in each condition), no errors were made and no subjects were excluded from analysis. Compatible trials yielded faster responses (M = 19.79 cycles, SD = 0.18) than neutral trials (M = 25.28 cycles, SD = 0.23), which again produced faster responses than incompatible trials (M = 34.04 cycles, SD = 0.73). The results are shown in Figure 17, where 17a shows the averaged simulated reaction times in cycles and 18b the empirical data from the study by Simon and Rudell (1967) in milliseconds. Overall, the simulation results fit well with the available behavioral data, demonstrating that and how code sharing between stimulus and response results in compatibility effects. Note that the processing logic according to which SRC effects are produced are identical to that responsible for action-effect compatibility effects as assessed in Simulation 1 (see Chapter 1).

Simulation 4: Stroop effect

As we do not differentiate between perceptual and action stages, one could argue that stimulus–response compatibility and stimulus–stimulus compatibility would need to work similarly in HiTEC.

Original experiment

Stroop (1935) showed that if people are instructed to name the ink color of color words, they are slower if the word (e.g., "blue") appears in an incompatible ink color (e.g., red). This compatibility effect is dramatically reduced if non-verbal responses are required (MacLeod, 1991), suggesting that the task-irrelevant words interfere (at least partly) with verbally naming the colors. Note that this interpretation of the Stroop effect bears a strong resemblance to the Simon effect as the effect is now attributed to incompatibility between a stimulus feature (ink color) and a response feature (verbal sound).

HiTEC simulation

In HiTEC the Stroop effect is simulated by having the model, as depicted in Figure 18, structured very similarly to the model used in Simulation 3 to simulate the Simon effect. The connections from visual shape to word feature codes have been made slightly stronger (weight of 0.45 instead of 0.4; see Appendix for further details) in order to take into account the richer experience of word reading as compared to color naming. During the learning trials, the model alternately executes 'M_1' and 'M_2', reflecting the 'physical' pronunciation of the respective words. The model is subsequently presented with the auditory feedback (i.e., reflecting the perception of this pronunciation) and associations are learned between motor codes and feature codes. During experimental trials, naming ink color of compatible color words benefits from facilitation whereas naming the color of incompatible color words suffers from interference.

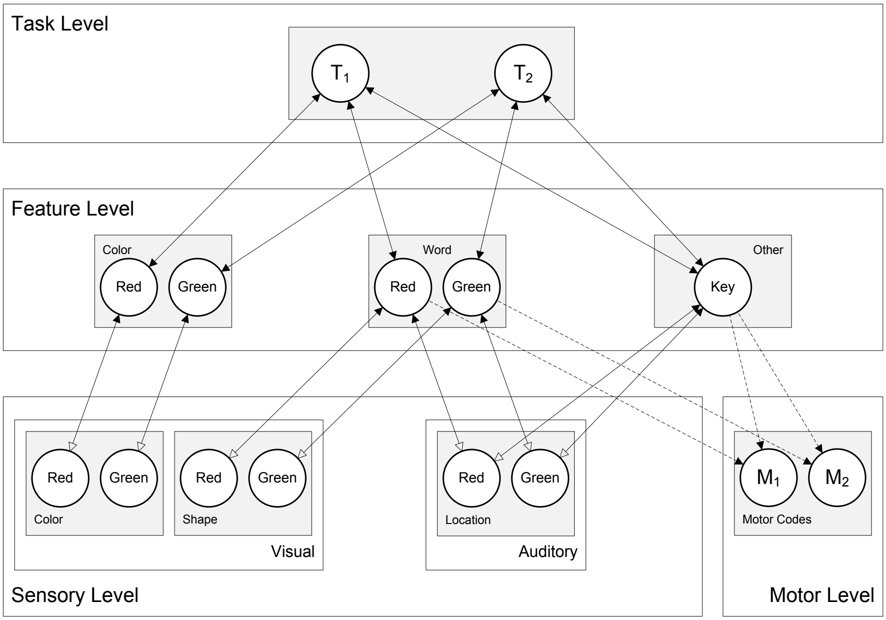
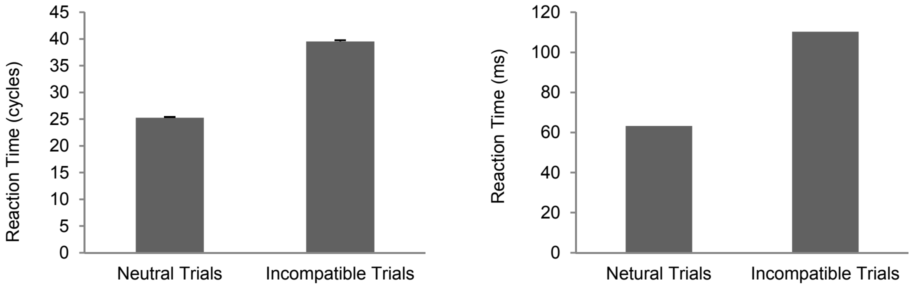


Figure 18. Specific HiTEC Model for Simulation 4. Feature codes are present for stimulus colors and words. Crucially, word feature codes are used for encoding both stimuli (i.e., the color words) and responses (i.e., the words to name the ink color). Note that this structure is in essence identical to the structure of the model used for Simulation 3. Connections between word feature codes and motor codes are learned during learning trials (i.e., pronouncing the words).



(a) (b)

Figure 19. Results of Simulation 4 compared with behavioral data (adopted from MacLeod, 1991), showing average reaction time means and standard deviations. Human variance data was not available.

Simulation results

In the simulation (15 simulated subjects, each performing 20 trials in each condition) 5% errors were made on average (all during incompatible trials) and one subject was excluded from analysis due to having more than 30% error trials. After removal of error trials, the results showed that responses were fastest with compatible trials (M = 19.22 cycles, SD = 0.14), intermediate with neutral trials (M = 25.27 cycles, SD = 0.24) and slowest with incompatible trials (M = 39.53 cycles, SD = 0.65). The global fit between simulation results and behavioral data is depicted Figure 19. Note that the Stroop simulation results point more strongly to an interference effect with non-compatible stimuli than to facilitation with compatible stimuli, a result that is also found in behavioral studies (MacLeod, 1991). In our simulation this is due to the stronger weights from visual shape sensory codes to word feature codes (see Appendix).

Discussion

This chapter attempts to address how and why compatibility effects arise in stimulus-response translation. These effects demonstrate that some aspects of stimulus-response translation occur automatically. As demonstrated in the simulations, HiTEC is able to account for these effects. In fact, SRC is an *inevitable* consequence of HiTECs structures and processing characteristics as we will now explain. First, in order to internalize task instructions into a task set, both stimuli and responses need to be represented on a distal level and associated through task codes (see Figure 20a). Secondly, actions are represented in terms of perceptual effects and therefore use the same distal codes as stimuli and, consequently, are grounded in the same perceptual world (Prinz, 1992; illustrated in Figure 20b).This means that code overlap is possible and – to the extent that stimuli and responses overlap in the external environment, such as spatial correspondence — very probable. Finally, HiTEC assumes integrated processing which means that stimulus coding and response coding also overlap in time. Thus, the task set results in a pathway mediated by task codes and defined in distal features, and in probable code overlap of these same distal features; as stimulus processing and response planning occur simultaneously, the cognitive system inevitably needs to combine task-driven and automatic feature code activation. As a result, code overlap between stimulus and response features results in either facilitation or interference effects (Hommel, 2004).

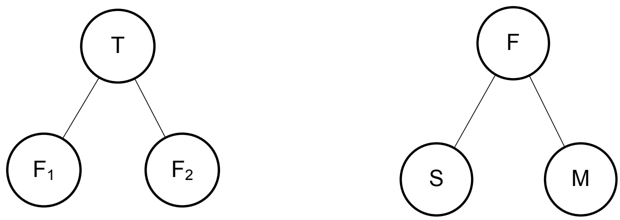


Figure 20. Schematic depiction of couplings between sensory codes, motor codes, feature codes and task codes. Panel (a) depicts a relation between task codes and feature codes that is part of the task set. F1 refers to a task relevant stimulus feature and F2 refers to an action effect feature of the required response. Panel (b) shows that feature codes are common codes, relating to both sensory and motor codes.

In Simulation 3, the simulation of the Simon effect, stimulus-response compatibility follows from the fact that responses are coded in terms of their spatial perceptual consequences (due to ideomotor learning, see Chapter 3). That is, left or right keypresses. In order to plan one of these keypress actions, the model needs to activate either the 'left' or 'right' feature code. Now, when a stimulus is presented left or right, the 'left' and 'right' feature codes will be activated both due to the exogenous excitation resulting from the presented stimulus and due to the endogenous excitation due to their roles as action effect features. When both stimulus perception and response anticipation activate the same 'left' or 'right' feature code, overall stimulus-response translation is faster, constituting a compatible trial. When they do not activate the same but competing codes, stimulus-response takes longer, constituting an incompatible trial.

In similar vein, in the simulation of the Stroop effect, the task irrelevant word feature only has influence because the response is coded using these features (which is a result from the action–effect learning). If the response is not verbally defined (e.g., in terms of key presses) the compatibility effect is dramatically reduced in behavioral studies (MacLeod, 1991). In HiTEC this would result in a different set of action effect features to be associated to the motor codes. Hence, code overlap with stimulus features would cease to occur, effectively eliminating the compatibility effect.

In typical computational models of SRC effects, such as the Simon effect, stimuli are represented in terms of non-spatial task-relevant codes (e.g., 'high tone' and 'low tone') and spatial task-irrelevant codes (e.g., 'left tone' and 'right tone'), and responses are also represented in terms of spatial codes (e.g., 'left key' and 'right key'). As depicted in Figure 21, stimulus codes and response codes are connected using two routes (e.g., Kornblum et al., 1990; Zorzi & Umilta, 1995; De Jong, Liang, & Lauber, 1994). A direct route connects the spatial stimulus codes to the corresponding spatial response codes, which is assumed to reflect the automatic process. The task instruction (e.g., "*when you hear a high tone, press the left key*") is implemented as a soft-wired connection from the non-spatial stimulus code (e.g., 'high tone') to a spatial response code (e.g., 'left key'), following the task instruction. This is assumed to reflect the controlled process. When a stimulus is presented, activation is propagated through the model towards the response codes. The response code that first reaches an activation threshold will be selected for execution. Now, when a compatible stimulus is presented (e.g., a high tone presented on the left), both the hard-wired spatial connections and the soft-wired task instruction-based connections contribute to a speedy activation of the correct response code. Conversely, when an incompatible stimulus is presented (e.g., a high tone presented on the right), the direct route activates the incorrect response. The controlled route, however, activates the response determined by the task instruction, which eventually is assumed to win this competition. As a result, processing incompatible stimuli results in longer reaction times than processing compatible stimuli. In sum, in dual route models, the stimulus–response compatibility effect arises from the interplay between the direct route, reflecting automatic comparison between spatial stimulus and response codes, and the controlled route, reflecting the task instructions. Thus, to account for SRC effects, these models drive on three main assumptions: (1) responses are represented by spatial codes, (2) attending to a stimulus automatically produces a spatial stimulus code, and (3) the outcome of a comparison between the spatial stimulus code and the spatial response code produces the compatibility effect. Here, this comparison is assumed to occur automatically and arise from the idea that stimuli and responses are similar (e.g., '*have dimensional overlap*', Kornblum et al., 1990; 1999).

Clearly, there are some strong similarities between these dual route models and HiTEC. First, the basic dynamic activation mechanisms of these models (i.e., codes, connections, activation levels) are very similar to HiTEC's connectionist implementation, and second, the general structure of the HiTEC model instance used to model the Simon (and Stroop) effect also shows some resemblance to 'two routes' (i.e., a route through the task codes and a route through the common codes). However, HiTEC does not share the main assumptions of the (strictly feedforward) dual route models and provides a different rationale for SRC. With respect to the main assumptions listed above, HiTEC assumes that (1) motor codes and representations of their perceptual effects are learned, allowing for the emergence of situation-specific meanings of actions (see Chapter 3), (2) task sets are implemented using common distal feature codes and recurrent connections with task codes. Including a feature code as response feature automatically makes it susceptible to stimulus based exogenous excitation and (3) compatibility between stimuli and responses (i.e., action effects) is due to the degree they are represented using the *same* common codes. These assumptions follow directly from key characteristics of the HiTEC model and do not require a notion of 'dimensional overlap' or 'similarity' that selectively applies to some combinations of stimuli and responses and not to others.

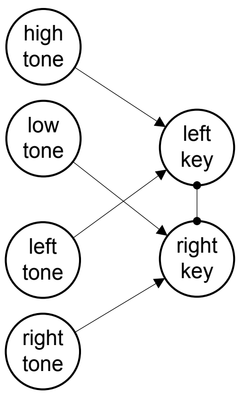


Figure 21. Dual route account of Simon effect (adapted from Zorzi & Umilta, 1995).

Also, the ideomotor learning of action-effect associations as employed in the simulations in this chapter allows for the flexibility and context dependence that is shown in a variety of SRC studies (see Chapter 3 for an elaborate overview). Moreover, in HiTEC task sets are implemented using recurrent connections only. These connections strictly follow the actual task instructions. In comparison, the Dimensional Overlap model (Kornblum et al., 1990; 1999), in addition to the controlled and automatic connections, assigns different activation dynamics to task relevant stimulus features than to task irrelevant stimulus features. Hence, we argue that HiTEC allows for a more parsimonious approach to controlled and automatic stimulus-response translation and provides a rationale – based on representations and processes – for *why* these SRC effects occur.

A related model of the Stroop effect (Cohen et al., 1990) also contains two routes. In this model, however, 'automatic' and 'controlled' is considered to depend on experience which they address explicitly. The model further allows for modeling multiple tasks (naming the ink color vs. naming the color word), showing somewhat of the task flexibility demonstrated by the HiTEC model. Task implementation in this model, however, is confined to injecting additional input to either one out of two task nodes thereby biasing the model to either one of the two implemented tasks.

The SLAM model (Phaf, van der Heijden, & Hudson, 1990) for attention in visual selection tasks is also used to model the Stroop effect. This connectionist model consists of multiple interacting levels of representation and employs two main processes, object selection and attribute selection, to perform a variety of filtering tasks. In order to account for the Stroop effect additional connections between stimulus features and response aspects are assumed ("privileged links") in similar vein as the automatic route in the dual process models described above.

Other models that include perception and action systems, such as the models by Ward (1999) and by Botvinick et al. (2009) do not address SRC; in these models stimulus features are simply connected to action features according to the task at hand; hence, stimulus features are just straightforwardly translated into action features. In contrast to the dual route models described above, however, connections in these models are recurrent. Hence, action activation can also influence stimulus perception, in similar spirit as HiTEC (see Chapters 2 and 3 for a more detailed comparison).

Another well-known SRC effect, which we did not explicitly model in HiTEC, is the Flanker effect (Eriksen & Eriksen, 1974). This effect is observed when participants are required to respond to a visual target with close-by distractors (flankers) which they are unable to ignore. For instance, if a discriminative response is required for a central target letter that is flanked by distractors, participants are faster if target and distractors are associated with the same response than with different responses. This result suggests that also for distractors the associated responses are activated and that this activation interacts with producing the response to the target. The Flanker effect is modeled by Cohen and Shoup (1997). In their model, displays of multiple stimuli are processed in terms of their individual features, which include location information. This process works separately for each feature dimension. At this stage, response competition is assumed to occur possibly yielding congruency effects. Finally, response activation from multiple dimensions is combined into a single actual response. Cohen and Shoup (1997) propose that the Flanker effect results from within-dimension competition. This set up somewhat resembles HiTECs architecture. Motor codes (responses) are associated to feature codes (features in dimensions). In contrast, however, HiTEC does not confine response competition within dimension, but rather assumes a model-wide integrated competition process. Crucially, to simulate the Flanker task, a model must be able to process a display of multiple objects and selectively treat one object as the 'target' and the others as 'distractors based on their location in the display. The HiTEC model currently does not provide for such differentiation but see (Cohen, Servan-Schreiber, & McClelland, 1992) for a PDP model of the Flanker effect.

To summarize, existing models of congruency in stimulus-response translation typically assume spatial response codes and special links between stimulus features and these response codes based on a certain 'similarity'. HiTEC does not need such assumptions as congruency effects follow naturally and inevitably from using common codes for both stimulus and response (i.e., action effect) representation.

Interestingly, dual route systems have also been proposed to account for fast and automatic responses to *affective stimuli* (LeDoux, 1996). In such a system, a 'low road', associated with the amygdala, automatically translates stimuli to responses. In parallel with this subcortical pathway there is a 'high road', associated with the cortical structures of the brain. This pathway analyzes the stimulus in a more fine-grained, but slower way. Together, these routes enable someone to respond quickly to affective stimuli and to process these stimuli in more detail in order to adjust behavior at a later point in time. Recent studies show that automatic processes may be affected by top-down influences (e.g., Beckers et al, 2002). The simulations in this chapter show that HiTEC is able to account for such influences. In Haazebroek et al. (2009b; 2011b) this is more explicitly applied to affective processing in a simulation of an affective version of the Simon effect (Beckers et al., 2002).

Although HiTEC accounts for some aspects of automatic processing, it must be noted that automaticity is a much broader field than these SRC effects alone suggest (see Moors & De Houwer, 2006 for an overview). Indeed, there is a long history of theorizing on the struggle between human will and habit (for a prototype, see Ach, 1910). With respect to the SRC effects discussed in this chapter, alternative explanations for automatic, uncontrolled or unconscious behavior include storing and retrieving action instances (Logan, 1988), integrating 'chunks' of behavior (Anderson, 1992) and over-learning of stimulus-response translation (Proctor and Lu, 1999; Tagliabue, Zorzi, Umiltà, & Bassignani, 2000). In this thesis, however, we have focused on aspects of automaticity that naturally follow from a set of key characteristics of our connectionist model of perception and action planning. Moreover, in HiTEC, important components of cognitive control are actually assumed to be exerted already before responding to any stimuli. This includes the prerequisites for code overlap, so that—somewhat paradoxically—automaticity is the result of control (Hommel, 2000a). In effect, we have eliminated the difference between automatic and controlled information processing in the model (i.e., everything is automatic). One could argue that this is there is more to cognitive control than modeled in current simulations. With respect to the simulated experimental paradigms, however, it seems that other types of (online) control are unnecessary.

Although we have explained how and why automaticity occurs in the HiTEC model by means of code overlap, one could still wonder *why* this would be beneficial for coordinating our behavior. Clearly, being slower or faster in a Simon task does not provide one immediate evolutionary advantages. However, even though the presence of such effects is convenient for the scientific study of perception-action relationships, their real benefit is prevalent in everyday life: object properties (e.g., location, shape) must often be translated into very similar action parameters (location, shape of hand) in order to efficiently interact with the environment. Perceiving an object and internally coding its features would therefore be likely to specify and literally prepare important components of the action plan that the given object affords (Hommel, 2009). Thus, rather than explicitly translating these stimulus features into response features (e.g., 'if big object, use large grasp action'), automaticity – in our framework using common codes (e.g., 'big') – allows for implicit, effortless translation of matching features.

To conclude, we have addressed how and why automaticity occurs in stimulus-response translation. In the HiTEC connectionist model stimuli and responses are represented using common codes. In typical SRC tasks, responses are defined in terms of features that are shared by the stimuli to be responded to. This means that a task set not only defines a controlled pathway but also an automatic translation path through the common codes used both for stimuli and responses (cf. Hommel, 2000b). In this chapter we focused on automaticity, in the next chapter we will discuss the role of task context more explicitly.