A Computational PDP Model for Explaining Automatic Imitation

Matthias Scheutz (mscheutz@cs.tufts.edu)
Department of Computer Science, 165 College Avenue
Medford, MA 02155 USA

Bennett I. Bertenthal (bbertent@indiana.edu)
Department of Psychology, 10th Street
Bloomington, IN 47405 USA

Abstract

Recent evidence suggests that automatic imitation is mediated by an observation-execution matching system that cannot be reduced to the same processes responsible for other stimulus-response (S-R) compatibilities. A computational model is developed with different patterns of connectivity for imitative and spatial compatibilities, and it is successful in simulating the results from three different S-R tasks. Variations of the model with identical connections for mediating the two compatibilities reveal a significantly poorer fit. These results provide converging evidence that imitative and spatial compatibilities are mediated by different processes.

Keywords: automatic imitation, S-R compatibility, connectionist modeling, direct matching hypothesis

Introduction and Background

The tendency to unintentionally and unconsciously mimic actions performed by others has long been noted. Charles Darwin, for example, commented that at leaping matches spectators would move their own feet as if imitating the athletes. More recently, Dijksterhuis and Bargh (2001) noted that we tend to whisper or speak louder when others do, scratch our heads upon seeing someone else scratch, or cycle faster after seeing a cycling race on TV. During social interactions, mimicry translates into a greater desire to want to cooperate and affiliate with those individuals imitating our gestures (Chartrand and Bargh, 1999). In spite of the frequency and significance of these behaviors, our understanding of the underlying mechanisms responsible for these automatic tendencies remains incomplete at best.

The prevailing hypothesis for explaining spontaneous mimicry or automatic imitation is that the perception of some actions automatically activates corresponding motor programs. There are by now more than 75 experimental studies investigating automatic imitation, and most of the evidence is based on stimulus response compatibility paradigms, in which both stimuli and responses involve human movements (Heyes, 2011). In this paradigm, faster responding is observed when stimuli and responses correspond along some perceptual, structural or conceptual dimension than when they do not (referred to as a "compatibility effect"). When both the stimuli and responses involve human movements, it is often assumed that automatic imitation is involved. One problem with this interpretation is that the pattern of results for automatic imitation and all other S-R compatibility effects is exactly the same (i.e., faster response times for compatible than incompatible responses to the stimuli). As a consequence, these results beg the question as to whether the processes mediating automatic imitation are specialized or instead are the same processes involved in other S-R compatibility tasks. In order to resolve this issue, it is necessary to find a paradigm where the results for automatic imitation and other S-R compatibility tasks are predicted to be different.

We recently provided such evidence by comparing imitative and spatial compatibilities in two experiments (Boyer, Longo, and Bertenthal, in press). The first tested for spatial compatibility with an imitative cue as the imperative stimulus, and the second tested for imitative compatibility with a spatial cue as the imperative stimulus. The stimulus consisted of a left or right hand with fingers spread apart and appeared on a computer screen from a third person perspective. Participants were instructed to respond to either the left-right spatial location or the anatomical identity of the index or middle finger tapping downward (Figure 1).

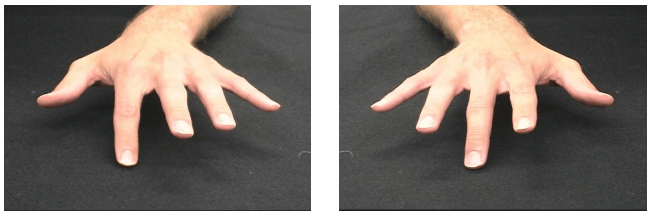

Figure 1. Left panel depicts the left hand stimulus with the index finger tapping down. Right panel depicts the right hand stimulus with the middle finger tapping down.

Responses consisted of pressing a key with the index or middle fingers on the right hand. In the standard S-R compatibility task (henceforth abbreviated "S-R task"), **the responses were compatible with a task-irrelevant spatial stimulus when the left hand was presented** (see Figure 1). For example, participants instructed to respond to the spatial cue would press a key with their index finger when responding to the left tapping finger. In this condition, both the stimulus and response are index fingers, and thus the response is facilitated via automatic imitation. Likewise, participants instructed to respond to the imitative cue would, for example, press a key with their middle finger when responding to the middle finger tapping. In this condition, both the stimulus and response correspond to the right side, and thus the response is facilitated via spatial compatibility. **When the stimulus corresponded to a right hand, the**

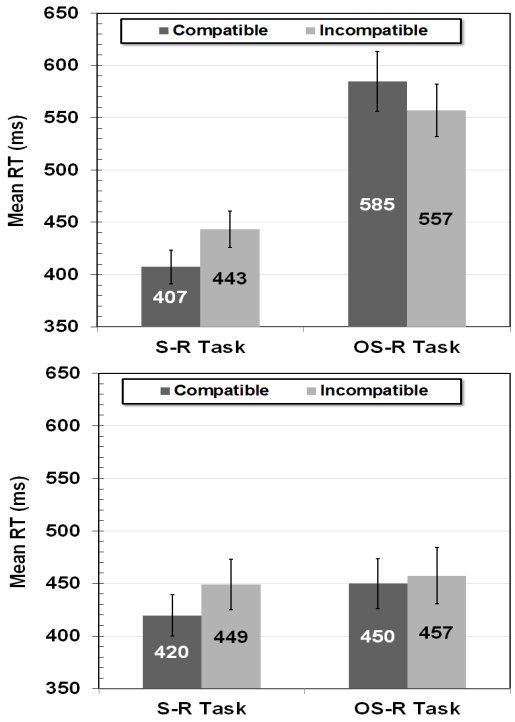

Figure 2. *Upper panel.* Mean response times (ms) to spatially compatible and incompatible stimuli as a function of task in Experiment 1. *Lower panel.* Mean response times (ms) to imitatively compatible and incompatible stimuli as a function of task in Experiment 2. (Error bars represent ± standard error of the mean.)

responses were not compatible with a task-irrelevant stimulus. In the Opposite S-R compatibility task (henceforth referred to as "OS-R task"), the task involved responding to the stimulus not presented on that trial (e.g., participants responded to a index finger with their middle finger). **Responses were compatible with a task-irrelevant stimulus that corresponded to the right hand, and incompatible with a stimulus corresponding to the left hand** (see Figure 2).

For the S-R task, Boyer et al. (in press) predicted similar results when the irrelevant stimulus was either spatially or imitatively compatible. By contrast, they predicted different results for the OS-R task. Switching instructions from a spatial S-R to a spatial OS-R task was first investigated by Hedge and Marsh (1975), who reported a "reverse compatibility effect". Although there is no consensus concerning the underlying mechanism, most hypotheses suggest that the recoding of the stimulus generalizes to the task irrelevant stimulus responsible for spatial compatibility and hence facilitates responding to the incompatible stimulus. By contrast, Boyer et al. (in press) hypothesized that imitative compatibility involves a direct and independent mapping between the task-irrelevant imitative stimulus and the response as suggested by recent neurophysiological evidence (Rizzolatti and Craighero, 2004). Thus, no reverse compatibility effect is predicted.

The results from the two experiments supported these predictions. In the S-R compatibility tasks, participants' response times were faster to the spatially and imitatively compatible stimuli than to the incompatible stimuli (see Figure 2). By contrast, in the OS-R tasks, participants' response times to the spatially incompatible stimuli were faster than to the spatially compatible (i.e., reverse compatibility effect), whereas response times to the imitatively compatible stimuli were faster than to the imitatively incompatible stimuli (although this difference was not significant). The results from this latter experiment thus suggested that imitative and spatial compatibilities are not mediated by the same domain-general process.

Although the preceding results suggest that spatial and imitative compatibilities are mediated by different processes, it remains an empirical question as to whether the response differences are a function of different neural architectures or more simply a function of differences in the parameterization of the same architecture. The standard processing model for explaining S-R compatibility effects is a dual route model whereby responses are activated by an intentional route as well as an automatic route (Zorzi and Umilta, 1995; Zhang, H., Zhang, J., and Kornblum, 1999). If the automatic route activates the same response as the intentional route, the response is facilitated. If, however, the automatic and intentional routes activate different responses, then the response is slowed down. In spite of sharing this general processing assumption, all dual route models are not the same. Some models are designed with some or all inputs mapped directly to the outputs, and other models are designed with a middle decision layer that selects the response (e.g., Sausser and Billard, 2002). In the current research, we hypothesized that the spatial compatibility task is mediated by the latter model, whereas the imitative compatibility task is mediated by a hybrid model (i.e., the automatic imitative route involves a direct mapping between input and output, but the controlled route involves a task-based mapping between the input and a middle decision layer). This latter model is consistent with recent theories suggesting that automatic imitation is due to a shared representation between the observation and execution of actions.

The purpose of the current investigation was to test this hypothesis with a computational model that was designed to simulate the empirical results from the previous study by Boyer et al. (in press).

A PDP Model of Spatial and Ideomotor Compatibility Effects

We started with our previous computational modeling efforts (Boyer et al. 2009) and develop a new three-layered (input-hidden-output) connectionist network, with nodes at each layer representing the stimulus input, the S-R translation, and the response, respectively. We use simplified interactive activation and competition connectionist units (Rumelhart and McClelland, 1986) with change in activation over time is given by $\Delta act/\Delta t = net - $

act(*netin*+*decay*), where *act*∈[0,1] is the *activation* of the unit, *netin*∈[0,1] is the *sum of the weighted inputs* to the unit and *decay*∈[0,1] is a constant *decay factor*. The model consists of eight units: two input units, called *finger units*, representing the perceived index ("I") and middle ("M") input fingers; two input units, called *location units*, representing the left ("L") and right ("R") location of the perceived input fingers (depending on the stimulus hand); two *output units* representing the index finger in the left location ("IL") and the middle finger in the right location ("MR") of the right hand; and two hidden units (or decision units), called *SR units*, affecting the S-R translation between inputs and outputs ("SR-IL" and "SR-MR"). As in (Boyer et al. 2009), we start with a base model which shows the participant's state before any task-based instructions and the task models which show the participant's condition after a task-based instruction. The base model consists of automatic connections between input and hidden, and hidden and output units. The input finger ("I" or "M") and spatial location ("L" or "R") are mapped onto the requisite hidden unit ("SR-MR" or "SR-IL") via the connections $I \xrightarrow[a]{}$ SR-IL, $M \xrightarrow[a]{}$ SR-MR, $L \xrightarrow[a]{}$ SR-IL, and $R \xrightarrow[a]{}$ SR-MR, respectively. These connections correspond to the compatible S-R translations between the imperative stimulus and the response, which we hypothesize are processed automatically by the hidden units presumably because they are overlearned and automatized. In addition to these automatic connections, we assume direct connections between input and output fingers that reflect the hypothesized direct matching pathways: $I \xrightarrow[i]{}$ IL and $M \xrightarrow[i]{}$ MR (note that there are no direct connections between input and output spatial locations). Lastly, hidden units are mapped onto corresponding output units via connections SR-IL $\xrightarrow[a]{}$ IL and SR-MR $\xrightarrow[a]{}$ MR.

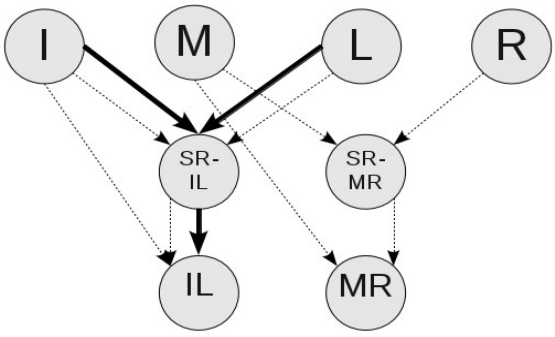

Figure 3. The proposed neural network model: the base model consists of only the dashed connections; the bold connections depict the S-R compatible task-based connections for the index finger.

From the base model, we construct the task models by adding additional connections that depend on the task instructions. Specifically, we add connections between input and hidden units: for the spatial compatibility models, we add task-based connections $I \xrightarrow[t]{}$ SR-IL and $M \xrightarrow[t]{}$ SR-MR reflecting that the imperative stimulus is the anatomical identity of the finger, and for the imitative compatibility models, we add task-based connections $L \xrightarrow[t]{}$ SR-IL, and $R \xrightarrow[t]{}$ SR-MR reflecting that the imperative stimulus is the left-right location of the finger. Furthermore, for the standard (S-R) mapping models we add task-based connections between hidden and output units SR-IL $\xrightarrow[t]{}$ IL and SR-MR $\xrightarrow[t]{}$ MR, while for the opposite (OS-R) mapping models we add task-based connections between hidden and output units SR-IL $\xrightarrow[t]{}$ MR and SR-MR $\xrightarrow[t]{}$ IL in which the controlled connection from the hidden unit crosses to the opposite output unit. This crossing is necessary given that the task instruction requires participants to select the opposite response to that selected in the standard mapping condition (i.e., either responding to the finger with the opposite identity or spatial location). Different from (Boyer et al. 2009), the current model has only excitatory connections (i.e., connections with positive values).

Inputs are applied to the model by fixing the netinput of the respective input units (e.g., "I" + "L") at a particular value to indicate, for example, the perceived index finger in the left position). The input is applied on each cycle of the trial because participants are able to observe the stimulus finger until they respond. The state of the model is updated in discrete time steps ("cycles") that correspond to 10 ms of real-time. Response selection is achieved whenever an output unit reaches the *action threshold* (i.e., the activation needed to perform a motor response by the finger). As such, the number of cycles computed from the introduction of the input (i.e., moving finger) until the action threshold is reached can be used to simulate the response times directly (e.g., 30 cycles correspond to 300 ms).

To minimize the number of free model parameters that can be used to fit models to the empirical data, we fix all base model parameters based on the study by Boyer et al., 2009. Specifically, we assume that all automatic connections $\xrightarrow[a]{}$ and direct mapping connections $\xrightarrow[i]{}$ have the same strength in all models, and they are set to a very low value of 0.001 (which is too low to generate any actions without task instructions, even if all input units are activated together). Moreover, we assume that the same decay factor of 0.05 for all computational units and also fix both the external input and the action threshold at 0.5. With those base model parameters fixed, we are left with the task-based connections as *free parameters* that can be used to fit the models to the empirical data.

Model Fitting and Simulation Results

We start with spatial and imitative compatibility in the S-R condition. There are four free parameters in each condition

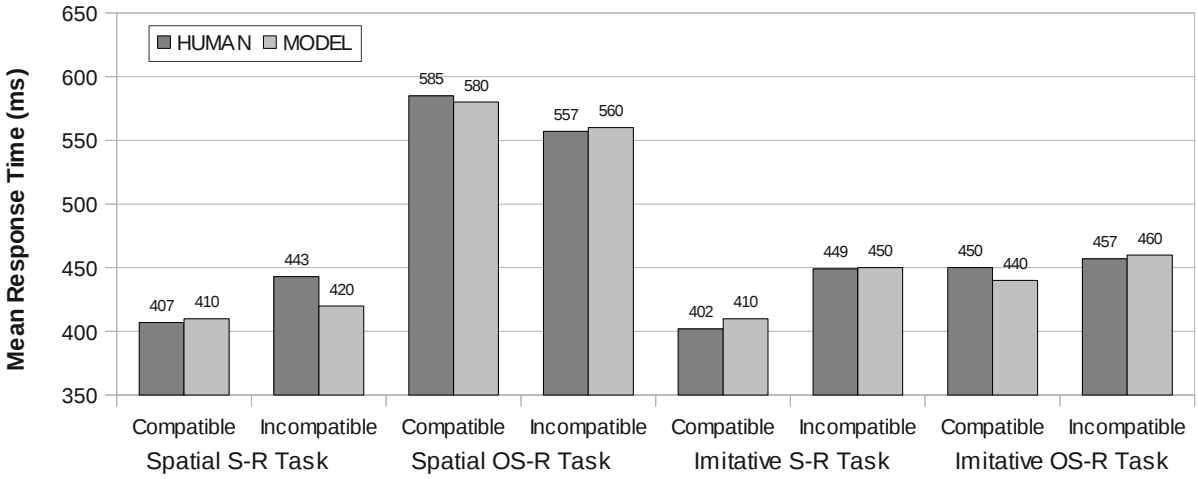

Figure 4. The simulation results of the proposed model for both compatibilities in the S-R and OS-R conditions.

corresponding to the input-to-hidden and hidden-to-output task connections, and we make the simplifying assumption that all task-based connections in the standard mapping models have the same strength. This assumption is plausible because both finger and location inputs require similar levels of encoding and integration followed by a similar S-R translation process regardless of whether the task instruction is to respond to the anatomical identity or location of the finger (Boyer et al., in press). Accordingly, we searched for a single positive value for all task-based connections that minimizes *the root mean squared error* for each model (*RMSE*). We found $\underset{t}{\rightarrow}$=0.086 to be such a value with *RMSE*=12.28 ms, which is just above the model resolution of 10 ms necessary for simulating a compatibility effect.

The next step was to use the model parameters in the S-R condition to predict the results in the OS-R condition. As previously described, the hidden-to-output connections for the task-based processes reverse between these two layers because the correct response is opposite to the one cued by the stimulus. We assume that, because only the hidden-to-output connections are different from the S-R models, the OS-R models should use the same connection strength (0.086) as the S-R models for the input-to-hidden connections. Moreover, we assume that both hidden-to-output connections have the same strength for both outputs; but based on the results of Boyer et al. (in press) showing longer response times for recoding the spatial cue, we introduce different strengths for imitative vs. spatial compatibility conditions. Hence, we are left with only one free parameter SR-IL$\underset{t}{\rightarrow}$MR=SR-MR$\underset{t}{\rightarrow}$IL in each of the two opposite mapping conditions.

In order to fit this free parameter, we begin with the OS-R condition for imitative compatibility. Unlike the results for the spatial compatibility condition which showed a reverse compatibility effect, the empirical results for this condition were similar to those for the S-R condition. Given this similarity, we predicted that the same value of 0.086 should work for SR-IL$\underset{t}{\rightarrow}$MR and SR-MR$\underset{t}{\rightarrow}$IL in the OS-R condition for imitative compatibility. Consistent with our hypothesis, this simulation was very successful with *RMSE*=7.38ms, which is less than the model resolution of 10ms.

The situation was somewhat different for fitting the spatial compatibility model in the OS-R condition. Recall that this condition differed from the other three conditions in two ways: (1) the results revealed a reverse compatibility effect, i.e., response times were faster to the incompatible than to the compatible stimulus; and (2) the response times in this condition were significantly higher than in the comparable imitative compatibility condition (i.e., 571 ms vs. 454 ms). Hence, in order to model these two results, we made two predictions: (1) the reverse compatibility effect is a function of the model architecture, and therefore it should be unnecessary to change the connection strength of 0.0856 for SR-IL$\underset{t}{\rightarrow}$MR=SR-MR$\underset{t}{\rightarrow}$IL from the connection strengths used in the other conditions; and (2) a lower connection strength (less than 0.086) is required for SR-IL$\underset{t}{\rightarrow}$MR=SR-MR$\underset{t}{\rightarrow}$IL to fit the model to the significantly longer response times with a small *RMSE*.

In the first simulation of this condition, we did not change the hidden-to-output connection weight (set SR-IL$\underset{t}{\rightarrow}$MR=SR-MR$\underset{t}{\rightarrow}$IL=0.086) which yielded a response time of 460 ms for the incompatible condition and a response time of 470 ms for the compatible condition. These results thus confirm our first prediction because they are consistent with a reverse compatibility effect. In the second simulation we lowered SR-IL$\underset{t}{\rightarrow}$MR=SR-MR$\underset{t}{\rightarrow}$IL=0.083 by about 3%, which yielded a response time of 560 ms for the incompatible condition and a response time of 580 ms for the compatible condition. In contrast to the previous simulation, these results revealed not only a reverse compatibility effect but also a very small *RMSE*=4.12, thus supporting our second prediction.

Several points are worth noting about the above modeling results. First and foremost, the current models are capable of capturing the reversed spatial compatibility effect in the

empirical data that Boyer et al. (2009) failed to capture. At the same time, the current model is simpler than the one presented in Boyer et al. (2009) because it has fewer nodes and fewer connections, does not use any inhibitory connections, and thus has overall fewer parameters. And

finally, the fits we obtained here are better than the fits in Boyer et al. (2009). The fact that spatial and imitative compatibilities in the OS-R condition required separate models is theoretically significant. This result thus provides an important source of evidence for concluding that the two compatibilities are not mediated by the same processes.

Testing the Generalizability of the Model

Given that the model succeeded at capturing the data from Boyer et al. (in press), we sought to test its generalizability as a means of providing further evidence that spatial and imitative compatibilities are not mediated by the same processes. Recently, Catmur and Heyes (2010) conducted a related study testing the effects of spatial and imitative compatibilities on response times. In this study, participants responded to a discriminative cue with an abduction of either the index or little finger of their right hand. On each trial, a left or right hand was displayed initially in a neurtral position with fingers spread apart and the outline of a small white circle appearing equidistant between the tips of the index and little fingers. Participants were instructed to respond as quickly as possible to the circle changing to orange or purple by abducting their index or little finger depending on the task instructions. Simultaneous with the appearance of the discriminative cue, the index or middle finger of the stimulus hand was abducted. By varying whether the stimulus corresponded to the left or right hand, it was possible to independently manipulate imitative and spatial compatibility, such that both compatibilities were present, only one, or neither.

To allow our current model to simulate the Catmur and Heyes (2010) task, we add two additional color nodes ("O" for "orange" and "P" for "purple") together with task-based connections $O \xrightarrow[t]{}$ SR-IL and $P \xrightarrow[t]{}$ SR-MR to the two hidden nodes "SR-IL" and "SR-MR", respectively, in the base model to reflect the requirement to abduct the little finger ("little finger") for the orange stimulus, and the index finger ("index finger") for the purple stimulus or vice versa (note that we can re-use the "middle finger" node for the "little finger"). We hypothesize that our model should also be able to simulate the results from this study with only minor adjustments. Keeping the task-based connections while inreasing the automatic connections (to 0.003) and lowering the direct mapping connections (to 0.0001), we get a result (with a small *RMSE* of 7.07ms) that very closely captures the Catmur and Heyes (2010) data (see Figure 5). This result confirms that our model is not limited to simulating results from only one specific experiment. It should be noted, however, that we do not simulate all results from the Catmur and Heyes (2010) experiment, because they also investigate the time course of the compatibility effect during the trial. In order to model these within trial timing effects, it would be necessary to design a stochastic model which is currently in development, but it is premature to report any results from this model.

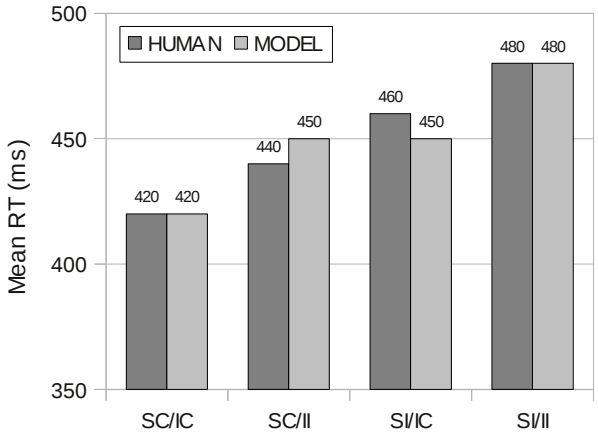

Figure 5. The simulation results of the model for the Catmur and Heyes (2010) experiment ("SC"="spatial compatible", "IC"="imitative compatible", "SI"="spatial incompatible", and "II"="imitative incompatible").

General Discussion

There are three principal conclusions from this investigation: (1) A computational model was presented that was successful in simulating the effects of spatial and imitative compatibilities in three separate conditions. (2) Unlike previous dual route models for explaining stimulus-response compatibilities, spatial and imitative compatibilities did not conform to the same architecture (i.e., imitative compatibility included a direct input-output connection, whereas spatial compatibility did not include this direct connection). (3) The finding that spatial and imitative compatibilities were modeled differentlty provides converging evidence that these two compatibilities are mediated by different processes.

A legitimate question is, therefore, whether the model would have been equally successful if we reversed which stimulus dimension was simulated with direct connections. Figure 6 shows the model with direct input-output connections for spatial, but not for imitative compatibilities. As can be seen, the model is just the mirror image of our previous model in the S-R conditions and thus fits the empirical results almost as well (with a slightly larger RMSE=15.52). In the OS-R condition, however, the model differs significantly from the empirical results, showing the exact opposite relations (as indicated by the two ovals in Figure 6). Not surprisingly, the RSME=38.87 for the OS-R condition is significantly higher than for the previous model with the direct imitative connections (RSME=7.12). It thus appears that these direct connections are necessary for predicting the imitative compatibilities in the OS-R condition. Given these results, we also investigated whether adding to our model direct connections for spatial compatibility improved the fit.

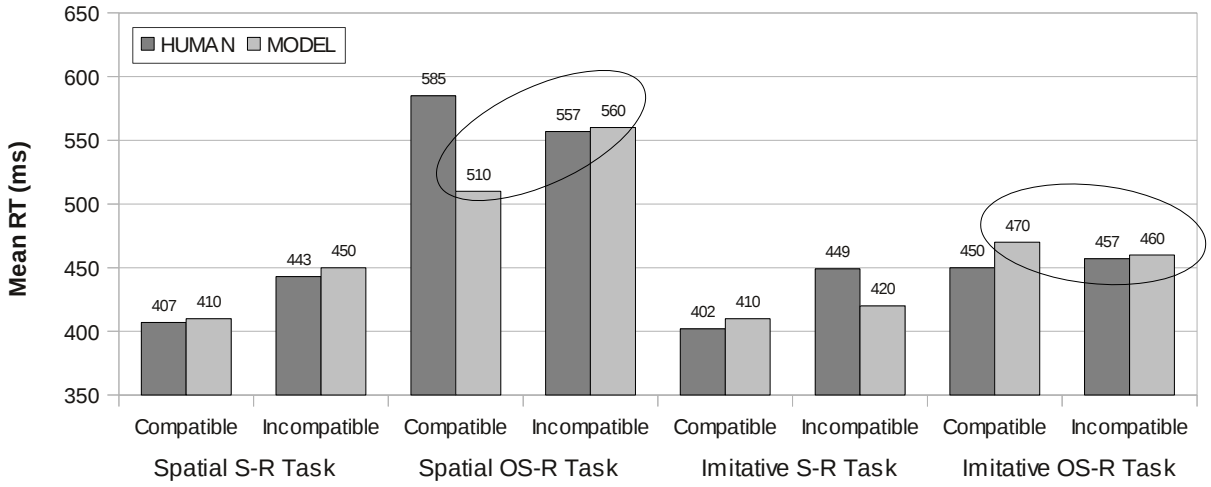

Figure 6. The simulation results of the model with no direct imitative, but direct spatial connections for both S-R and OS-R conditions. The ovals depict the two conditions in which the model predicts the opposite effects from the empirical data.

Preliminary experiments show that the addition of direct connections between the location and output nodes results in does even worse (without lowering of the task-based connections) in the S-R condition (with RSME=21.27 compared to RSME=13.99 for our previous model). However, as with the model including only direct spatial connections, this new model failed to match the empirical data in the reverse spatial compatibility condition, with an overall RSME=37.89 for both OS-R conditions. Hence, neither models with only direct spatial nor models with direct spatial and direct imitative connections are capable of matching the empirical data in the OS-R conditions as well as our proposed model under the given parameter assumptions.

Conclusions

In this paper, we introduced the first computational PDP model that can be fit to human data dissociating spatial from imitative compatibilities *and* can be used to make predictions about related tasks. While the results are constrained by the chosen constants and the modeling process, the comparison with alternative model architectures together with the model's ability to predict performance in a related, yet different task, are an encouraging step towards a full-fledged investigation of model parameters and model architectures that might be able to account for the empirical differences between imitative and spatial compatibility effects in a great variety of experimental paradigms.

References

Boyer, T. W., Scheutz, M., and Bertenthal, B. I. (2009). Dissociating ideomotor and spatial compatibility: Empirical evidence and connectionist models. In N. Taatgen, H. van Rijn, and L. Schomaker (Eds.) Proceedings of the 31st Annual Conference of the Cognitive Science Society (pp. 2280-2285). Austin, TX: Cognitive Science Society.

Boyer, T.W., Longo, M.R., and Bertenthal, B.I. (in press). Is automatic imitation a specialized form of stimulus-response compatibility? Dissociating imitative and spatial compatibilities. *Acta Psychologica*.

Catmur, C., and Heyes, C. (2011). Time course analyses confirm independence of imitative and spatial compatibility. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 409-421.

Chartrand, T. L., and Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893–910.

Dijksterhuis, A., and Bargh, J.A. (2001). The perception-behavior express-way: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology, 33*, 1-39.

Hedge, A., and Marsh, N. W. A. (1975). The effect of irrelevant spatial correspondence on two-choice response-time. *Acta Psychologica, 39*, 427-439.

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin, 137*, 463-483.

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169-192.

Rumelhart, D. E. and McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume I.* Cambridge, MA: MIT Press.

Sauser, E. L., and Billard, A. G. (2006). Parallel and distributed neural models of the ideomotor principle: An investigation of imitative cortical pathways. *Neural Networks, 19*, 285-298.

Zhang, H., Zhang, J., and Kornblum, S. (1999). A parallel distributed processing model of stimulus-stimulus and stimulus-response compatibility. *Cognitive Psychology, 38*, 386-432.

Zorzi, M., and Umilta, C. (1995). A computational model of the Simon effect. *Psychological Research, 58*, 193-205.