Cognitive Control in Number Processing – A Computational Model

Stefan Huber (s.huber@iwm-kmrc.de) and Korbinian Moeller (k.moeller@iwm-kmrc.de)
Knowledge Media Research Center, Schleichstrasse 6, 72076 Tuebingen, Germany

Hans-Christoph Nuerk (hc.nuerk@uni-tuebingen.de)
Department of Psychology, Eberhard Karls University, Schleichstrasse 4, 72076 Tuebingen, Germany and
Knowledge Media Research Center, Schleichstrasse 6, 72076 Tuebingen, Germany

Pedro Macizo (pmacizo@ugr.es)
Facultad de Psicología, University of Granada, Campus de Cartuja S/N
18071 Granada, Spain

Amparo Herrera (aherrera@um.es)
Departamento de Psicología Básica y Metodología, Universidad de Murcia,
30003 Murcia, Spain

Klaus Willmes (willmes@neuropsych.rwth-aachen.de)
Department of Neurology, Section Neuropsychology, University Hospital, RWTH Aachen University, Pauwelsstraße 30
52074 Aachen, Germany

Abstract

Recent research indicated that processes as easy as two-digit number comparison are under cognitive control. In the present study we evaluated the validity of the cognitive control architecture proposed by Macizo and Herrera (2011, 2012) to account for their empirical findings regarding the unit-decade compatibility effect: longer responses for unit-decade incompatible (e.g., 37_52, 3<5, but 7>2) as compared to compatible number pairs (e.g., 42_57, 4<5 and 2<7). Artificial neural network models implementing a cognitive control circuit were trained for magnitude comparison on stimulus sets with different proportions of compatible and incompatible pairs (20%, 50% and 80% incompatible pairs). In line with the empirical observations of a proportion congruity effect, the simulated compatibility effect increased with the number of compatible pairs. Finally, with the same model specifications we were also able to simulate the Gratton effect, providing further evidence for cognitive control in number processing to be substantiated by future empirical studies.

Keywords: two-digit number comparison, cognitive control, computational modeling, numerical cognition

Introduction

Generally, magnitude information is the most important semantic information conveyed by numbers. One of the most common tasks to investigate the magnitude representation of numbers is the magnitude comparison task in which participants have to single out the larger one of two numbers. Systematic evaluation of reaction times (RT) and error rates (ER) is supposed to be informative with respect to the mental representation of number magnitude. For multi-digit numbers three different models have been proposed: (i) the holistic (e.g., Dehaene, Dupoux, & Mehler, 1990), (ii) the strictly decomposed (e.g., Verguts & De Moor, 2005) and (iii) the hybrid model (e.g., Nuerk & Willmes, 2005). In this context, holistic indicates that multi-digit numbers are represented as integrated entities not retaining place-value information. Contrarily, in the strictly decomposed model the place-value structure of the Arabic number system is considered specifically, since separate representations of units, tens, hundreds, etc. are assumed. Finally, the hybrid model combines both approaches by proposing that two-digit numbers are represented both holistically and decomposed.

All of these three models can account for the numerical distance effect and the problem size effect. The distance effect describes the finding that magnitude comparisons become increasingly difficult as the numerical distance between the to-be-compared numbers decreases (e.g., Moyer & Landauer, 1967). On the other hand, the problem size effect indicates that numerical tasks get more difficult the larger the numbers involved (e.g., Brysbaert, 1995). However, there is another effect which allows to differentiate between the three models: the unit-decade compatibility effect (Nuerk, Weger, & Willmes, 2001). When comparing two two-digit numbers the pair can be either unit-decade compatible, if separate comparisons of tens and units bias the same decision (e.g., 42_57, 4<5 and 2<7) or unit-decade incompatible, if comparing tens and units separately leads to opposing decision biases (e.g., 37_52, 3<5, but 7>2). Although overall distance was matched for compatible and incompatible number pairs, reliably longer RT and higher ER were observed for incompatible as compared to compatible pairs. This reflects an interfering influence of the irrelevant unit digits on the overall comparison process. Thus, the compatibility effect corroborates the notion of a decomposed representation of two-digit numbers, either without (i.e., strictly decomposed model) or with an additional holistic representation (i.e., hybrid model). Further support for strictly decomposed representations of tens and units comes from computational modeling. Moeller, Huber, Nuerk, and Willmes (2011) implemented computational models reflecting each of the three models and evaluated these models with respect to their ability to account for the numerical effects described above. Taking into account model parsimony the strictly decomposed model accounted best for the empirical data.

Unit-decade compatibility and cognitive control

In addition to just simulating two-digit number magnitude comparison, Moeller et al. (2011) also investigated the impact of stimulus properties on the compatibility effect. In particular, they evaluated whether the compatibility effect indeed depends on the number of within-decade filler items (e.g., 53_58), as previously suggested by Nuerk and Willmes (2005). It is assumed that a higher number of within-decade fillers should prevent participants from focusing on the tens only, because for within-decade fillers with an identical tens digit the units are decision relevant. Thus, it would be a beneficial strategy to focus more closely on the units under such conditions. As a consequence unit-based interference in incompatible pairs should be stronger, which in turn would strengthen the unit-decade compatibility effect. Indeed, the simulations by Moeller et al. (2011) corroborated this rationale. When manipulating the number of within-decade fillers (i.e., 25%, 50%, and 75%) in an additional learning phase, they found that the strength of the unit-decade compatibility effect increased linearly with the number of within-decade fillers.

Only recently, this prediction was also corroborated empirically in a study by Macizo and Herrera (2011), however, so far for number word stimuli only. In their study, Spanish speaking students had to compare number words denoting pairs of either compatible or incompatible two-digit numbers. Comparable to the simulations by Moeller et al. (2011), Macizo and Herrera (2011) manipulated the number of within-decade fillers (i.e., 20%, 50% and 70%). Usually, for Spanish speaking participants a reverse unit-decade compatibility effect with longer RT for compatible pairs of number words can be observed (e.g., Macizo & Herrera, 2010). This is argued to be due to the fact that in Spanish number words the tens are spoken before the units (e.g., 37 = treinta y siete) and thus participants can focus more easily on the tens and more or less neglect the units. When overall distance is matched, the distance between the decade digits is necessarily larger for incompatible number pairs resulting in faster responses for incompatible pairs. Importantly, Macizo and Herrera (2011) observed that this reversed compatibility effect changed into a regular unit-decade compatibility effect (with faster responses for compatible pairs) in the condition with 70 % within-decade fillers. Thus, in line with the simulation results by Moeller et al. (2011), increasing the relevance of unit digits in the overall decision process via increasing the number of within-decade fillers led to an increase of the unit-decade compatibility effect.

Moreover, Macizo and Herrera (2011) interpreted their findings to reflect influences of cognitive control. Generally, cognitive control is relevant in conflict situations in which participants have to inhibit interfering but irrelevant information. Transferred to the case of two-digit number comparison, participants should basically ignore the units in favor of the tens in the critical between-decade pairs as the units are irrelevant for the overall decision. However, the presence of the compatibility effect indicates that focusing on the tens is not perfect. Additionally, increasing the saliency of the units by increasing the number of within-decade fillers even increased the influence of conflicting information. However, manipulating the number of within-decade fillers is not the only way to investigate influences of cognitive control with respect to the compatibility effect.

Only recently, Macizo and Herrera (2012) validated their account by showing that the unit-decade compatibility effect is also subject to a proportion congruity effect (Tzelgov, Henik, & Berger, 1992), this means its size depends on the proportion of compatible to incompatible items in the stimulus set. The unit-decade compatibility effect decreased as the number of incompatible items in the set increased. Thus, the effect was largest in a condition with only 20% incompatible items and smallest in a condition with 80% incompatible items. Macizo and Herrera (2012) interpreted their findings in terms of the adaption-by-binding account on cognitive control (Verguts & Notebaert, 2008, 2009) suggesting that connections between task relevant dimensions will be strengthened whenever conflict is detected. Transferred to the case of two-digit number comparison, this means that the more conflictive (i.e. incompatible,) items a stimulus set involves, the better the cognitive system adapts to unit-decade incompatibility which then reduces the compatibility as a consequence.

The present study

In the current study we aimed at further investigating the impact of cognitive control on two-digit number processing, as indicated by modulations of the unit-decade compatibility effect by means of computational modeling. In a two-stage approach we first attempted to simulate the findings of Macizo and Herrera (2012) about the proportion congruity effect, using a modified and extended version of the strictly decomposed model of Moeller et al. (2011). We focused on the strictly decomposed model, as it simulated the existing empirical data on the compatibility effect best. Additionally, as found by Macizo and Herrera (2011), the strictly decomposed model already provided valid predictions with respect to the influence of differing numbers of within-decade fillers. Against this background we expected (i) the strictly decomposed model to be flexible enough to account for changes in the proportion of compatible and incompatible items, with the simulated compatibility effect assumed to be smallest in the condition with the highest proportion of incompatible number pairs.

In a second step we then aimed at making predictions for future empirical research. Macizo and Herrera (2011) also found first evidence for a Gratton effect (e.g., Gratton, Coles, & Donchin, 1992), which proposes that any congruity effect should be reduced when the preceding trial was already incongruent compared to the case of a congruent preceding trial. Accordingly, the authors found the compatibility effect to be smaller after a preceding incompatible item than after a preceding compatible item. However, the primary focus of their study was not to investigate the Gratton effect. As a consequence, item sequence was not manipulated systematically, which led to distances and problem sizes not being matched for different sequences of compatible and incompatible items. This might have biased the Gratton effect. Thus, the main purpose of this study was to investigate the Gratton effect using a more controlled item set systematically manipulating item sequence. We expected to observe a reliable Gratton effect for the compatibility effect, employing the same model specifications as for the replication of the proportion congruity effect

Method

Model description

The model is based on the decomposed model by Moeller et al. (2011), but slightly modified in some respects: Model components were restructured in a different way and a task demand layer as well as a conflict monitoring unit were added to the network (see Verguts & Notebaert, 2008 for a similiar architecture). Hence, the new feature of the current model is that it integrates single-digit comparison networks for tens and units with a cognitive control network (see Figure 1).

Digit comparison networks

The digit comparison networks each have an input layer and a comparison layer (see Figure 1A and 1B). In the input layer each digit is represented by 10 nodes using place coding characteristics (see Verguts, Fias, & Stevens, 2005, for a similar approach). The activation function of node *i* for digit *j* can be formalized by

$$f_i(j) = e^{-10*|i-(j+1)|}, 1 \leq i \leq 10; 0 \leq j \leq 9 \quad (1)$$

All input nodes have feed forward connections to the two comparison nodes which code “left digit larger” and “right digit larger”. Furthermore, there are inhibitory connections between comparison nodes with w = - 2. The activation function of the comparison nodes is a sigmoid function. To simulate reaction times the propagation function includes a constant τ, which affects how fast the maximum activation of a node is reached and reflects the average input over time for a comparison node (see Moeller et al., 2011; equation 2). Moreover, random Gaussian noise is added at each time step (M = 0, SD = 0.005). The number of time steps, by which a threshold activation θ of the output nodes is reached, determines the simulated reaction times (RT). However, since it is possible that this threshold will never be reached, we additionally defined a maximal value for t. The constant τ was set to 0.01, the threshold θ to 0.75 and the maximum value for t to 100.

Initial weights were pseudo-random values from a uniform distribution in the interval [-1; 1]. The computational model was trained using the delta rule with a learning rate of 0.01. The frequency of occurrence of each digit in a training set was determined by a Google survey providing an estimate of the frequency of occurrence in daily life. After 100,000 training trials the network compared all possible combinations of digits correctly. Only one digit comparison network was trained.

Trained weights were then reused for the second network. Thus, the specifications of the same single-digit comparison network were used for the comparison of tens and units with these two single-digit comparison networks being monitored by the cognitive control network. Moreover, this architectural design also allows extending the network to an arbitrarily number of digits.

Cognitive control network

A model architecture as suggested by Verguts and Notebaert (2008) served as the basis for our cognitive control network (see Figure 1C). Comparison nodes of the two single-digit comparison networks served as input nodes to the cognitive control network. These nodes were connected to a response layer with two nodes coding “left digit larger” and “right digit larger”. Moreover, there is a task demand layer for the two tasks of comparing either tens or units. Finally, our model also incorporated a conflict monitoring unit which gets input from the response nodes. The activation of a node in the comparison layer is calculated as in the study of Verguts and Notebaert (2008; equation A1), except that we do not use an indicator function, but the activation propagated from the input layer to the comparison layer. Moreover, the activation function for a response node and for the conflict monitoring unit are identical to the functions used by Verguts and Notebaert (2008; equations A2 and A3). However, we added random Gaussian noise at each time step (M = 0, SD = 0.1)[1]. Finally, our model incorporates the same Hebbian learning rule as in the model of Verguts and Notebaert (2008; equations A4 and A5).

Parameter values are mostly identical to the ones used by Verguts and Notebaert (2008): τ = 0.75, β_{in} = 0.2, w^{inh} = -0.5, C = 0.7, β_{con} = 1, λ_{con} = 0.7, λ_w = 0.7, α_w = 5 and β_w = 0.5. The maximum value of t was set to 200 and threshold θ to 0.8. Moreover, activations of nodes in the task demand layer were preset to 1 for the decade comparison task and to 0.1 for the unit comparison task. Initially, connection weights between task demand nodes and corresponding comparisons nodes were 0.5. For decade comparison nodes connection weights between comparison nodes and response

[1] SD = 0.1 was chosen to resemble variances found in human RT

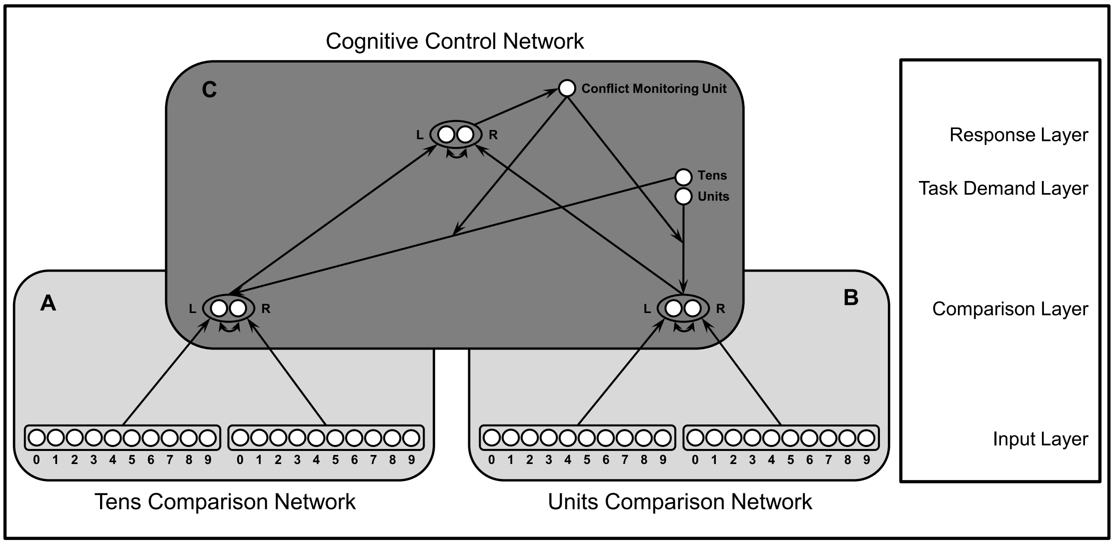

Figure 1: Schematic illustration of the model architecture: Networks A and B depict digit comparison networks for tens (A) and units (B) and network C the cognitive control network. L = "left digit larger" node, R = "right digit larger" node.

nodes were 0.91 and for unit comparison nodes 0.64.[2] Connection weights between task demand nodes and corresponding comparison nodes are adapted according to the Hebbian learning rule of Verguts and Notebaert (2008), whereas connection weights between comparison nodes and response nodes are kept fixed.

Analysis and Results

To obtain simulated RT data, 30 artificial neural networks were trained as described above and were then presented the two-digit number stimuli used by Macizo and Herrera (2012). The stimulus set contained three groups of two-digit number pairs with different proportions of compatible and incompatible number pairs (i.e., 20%, 50% and 80% incompatible number pairs). Trial order was randomized for each network. We simulated all three groups with the same 30 networks. However, after each proportion congruity condition we reset weights from the task demand layer to the comparison layer to their initial values to investigate their adaptation to the different proportion conditions. Thereby, our simulations aimed at reflecting the between-subject design employed by Macizo and Herrera (2012).

Model Validation

The validation of the model was threefold: First, we tested whether networks could replicate the most important effect for number magnitude comparison: the numerical distance effect. Hence, we calculated mean RT for each item over all 30 networks and correlated mean RT with absolute distance. Distance reliably predicted simulated RT in all three proportion congruity conditions [20% incompatible items: $r(98) = -.64$, 50% incompatible: $r(98) = -.63$ and 80% incompatible: $r(98) = -.79$; all $p < .001$]. The negative correlations indicate simulated RT to increase with decreasing numerical distance between the to-be-compared numbers as is also typical for human RT patterns.

Second, we were interested in how well our simulated RT corresponded to empirical RT observed by Macizo and Herrera (2012). Therefore, we calculated mean simulated and mean empirical item RT over all networks and participants for each proportion congruity condition separately. Simulated RT correlated reliably with empirical RT, further corroborating model validity [20% incompatible items: $r(98) = .68$, 50% incompatible: $r(98) = .57$ and 80% incompatible: $r(98) = .58$; all $p < .001$].

Third, we aimed at replicating the proportion congruity effect as found by Macizo and Herrera (2012). Simulated RT was analyzed by repeated measures ANOVA, discerning the factors proportion congruity condition (20% vs. 50% vs. 80% incompatible items) and unit-decade compatibility (compatible vs. incompatible). In line with Macizo and Herrera (2012), we observed a significant unit-decade compatibility effect [$F(1, 29) = 1331.94$, $p < .001$, $\eta_p^2 = .98$] with longer RT for incompatible trials [$M = 8.91$ vs. $M = 11.00$]. Moreover, proportion congruity conditions differed significantly [$F(2, 58) = 53.80$, $p < .001$, $\eta_p^2 = .65$]. Bonferroni-Holm corrected *t*-tests indicated reliable differences between the conditions with 20% and 80% as well as with 50% and 80% incompatible pairs [$p < .001$]. Mean simulated RT for conditions with 20%, 50% and 80% incompatible pairs was: $M = 10.21$, $M = 10.17$ and $M = 9.50$, respectively. Most importantly, the interaction between group and compatibility was significant [$F(2, 58) = 5.33$, $p < .01$, $\eta_p^2 = .16$]. As indicated by a linear trend contrast, the compatibility effect increased linearly with the number of compatible number pairs in the stimulus set [$F(1, 29) = 8.23$, $p < .01$, $\eta_p^2 = .22$].Compatibility effects for proportion conditions with 20%, 50% and 80% incompatible items were: $M = 2.35$, $M = 2.08$ and $M = 1.85$.

[2] The qualitative pattern of results was robust to changes of these parameters.

In line with the empirical data, the compatibility effect differed reliably between conditions with 20% and 80% incompatible pairs [$t(29) = 2.87$, $p < .05$], but did not differ between proportion congruity conditions with 20% and 50% incompatible items and between conditions with 50% and 80% incompatible items [both $t < 1.81$, $p > .09$].

Gratton effect

The Gratton effect proposes a reduced congruity effect after a preceding incongruent trial than after a congruent trial. Transferred to the case of two-digit number comparison, we should find a less pronounced compatibility effect when the preceding number pair was incompatible than when it was compatible. To investigate the Gratton effect, we created a stimulus set with matched distances (M = 36.65, SD = 0.52) and problem sizes (M = 59.30, SD = 0.49) for different sequences of three consecutive trials, which resulted in eight different groups of stimulus sequences: ccc, cci, cic, cii, icc, ici, iic, and iii (c = compatible item and i = incompatible trial). We created 30 number pairs per group resulting in a total of 240 pairs.

To evaluate the possible influence of a Gratton effect on the compatibility effect, we again simulated RT with 30 networks. Subsequently, we computed the compatibility effect for each case of two identical preceding number pairs (i.e., RT cci - RT ccc, RT cii - RT cic, RT ici - RT icc and RT iii - RT iic). Simulated RT were then analyzed with a repeated measure ANOVA with the factor preceding number pairs. The analysis revealed a significant influence of the preceding pairs [$F(3, 87) = 5.41$, $p < .01$, $\eta_p^2 = .16$; see Figure 2]. Bonferroni-Holm corrected t-tests indicated (marginally) significant differences between the compatibility effect following two compatible pairs and (i) one compatible and one incompatible pair [$t(29) = 3.85$, $p < .01$] as well as (ii) two incompatible pairs [$t(29) = 2.80$, $p < .05$] and (iii) one incompatible and one compatible pair [$t(29) = 2.51$, $p = .07$]. In line with the proposition of a Gratton effect, these results indicated that the unit-decade compatibility effect was less pronounced when at least one of the preceding two pair was incompatible.

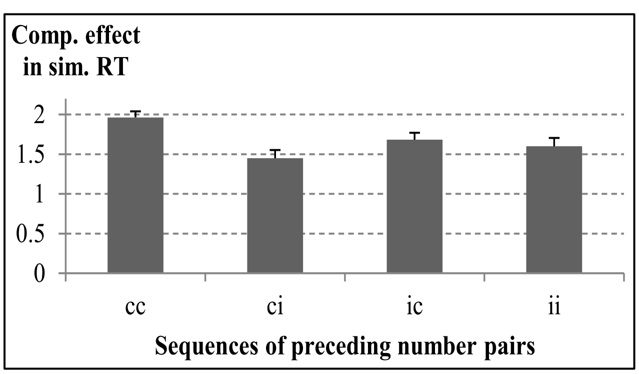

Figure 2: Simulated compatibility effects for sequences of preceding number pairs cc, ci, ic and ii (c = compatible, i = incompatible; Compatibility effect = simulated RT of incompatible – compatible trials)

Discussion

In the present study we employed computational modeling to (i) simulate recent empirical effects of cognitive control in two-digit number processing and (ii) to make predictions for future empirical studies. In particular, we were interested whether the cognitive control account, as formulated by Macizo and Herrera (2011, 2012) and implemented by the strictly decomposed computational model (Moeller et al., 2011) augmented by network components for cognitive control (adapted from Verguts & Notebaert, 2008) can (i) account for the proportion congruity effect on the unit-decade compatibility effect (Macizo & Herrera, 2012) and (ii) consider another effect of cognitive control: the Gratton effect. The current data were informative in both respects.

Cognitive control in number processing

Our study provides further support for the argument of Macizo and Herrera (2011, 2012), see also Moeller, Klein, and Nuerk (2013) that cognitive control is involved in two-digit number comparison. We implemented their suggested model of cognitive control in number processing which comprised (i) separate input layers for the processing of tens and units, (ii) a response layer, (iii) a task demand layer specifying the current task (i.e., whether the tens or the units are decision relevant), and (iv) a conflict monitoring unit (adapting to possible conflict from interfering unit information as in the case of unit-decade incompatible number pairs). In addition, we included an input layer for the representation of the single digits of tens and units. Thereby, the input layers in Macizo and Herrera (2011, 2012) correspond to our comparison layer. These authors also suggested that the adaption by binding theory of (Verguts & Notebaert, 2008, 2009) should account for the empirical findings.

To illustrate how adaption by binding works, the Stroop task is usually referred to. In the following we will reiterate this example with the relevant aspects being transferred to the case of two-digit number comparison. The input layer consists of nodes for colors and words (i.e., tens and units in the case of two-digit number comparison). The response layer contains all possible responses (e.g., red, vs. green, i.e., left larger vs. right larger in our case). In the task demand layer the tasks - naming either color or word - are represented (reflecting decisions based on either tens or units for two-digit number comparison). The most important node in this model is the conflict monitoring unit. It takes input from the response layer and modifies weights between task demand and comparison layer, such that the model can resolve conflicts more easily by focusing on the relevant dimension (color or word, i.e., tens or units in our case).

In the adaption by binding theory, the conflict monitoring unit triggers an arousal system which influences connections between the task demand layer and the input layer using Hebbian learning. Thus, conflicts between two tasks are not resolved by focusing on the relevant task, but by modifying task-relevant connections. Transferred to the case of two-digit number comparison, units are made less conflicting by changing weights between the task demand layer and the input layer, which was confirmed by the current simulation study: We successfully replicated the proportion congruity effect on the compatibility effect. The simulated compatibility effect was least pronounced in the case of 80% of incompatible number pairs in the stimulus set. In this condition unit interference in the frequent incompatible pairs was less pronounced, because the cognitive system has been adapted to the case of incompatible number pairs.

Additionally, we were able to make specific predictions concerning the Gratton effect for multi-digit number processing. From the current computational modeling results we hypothesize that the unit-decade compatibility effect should be modulated by the preceding number pairs being either compatible or incompatible. For sequences of three trials (strictly controlled for numerical distance and problem size) we predict the unit-decade compatibility effect to be less pronounced when at least one of the preceding two items was incompatible. Macizo and Herrera (2011) already reported first empirical evidence for a Gratton effect – however, for number words, but not for Arabic numbers without systematic control of stimulus sequences and their numerical properties.

Together with the current simulation data this is informative on how cognitive control may influence number processing. Adaptation in our simulations was a result of continuous adaptation through stimulus characteristics (unit-decade compatibility) on a trial by trial basis, also described as cognitive control acting locally. More specifically, after each number pair connection weights between task demand layer and comparison layer were adapted. This is important to note because in the study by Moeller et al. (2011) effects of cognitive control (i.e., adaptation to different proportions of within-decade fillers) were simulated by adjusting weights beforehand also called global cognitive control. However, adaptation effects were observed only after a large number of trials (i.e., 1000 cycles). Using the modified and extended model in the present study, adaptation happened very fast. This replicates adaptation effects found in studies with human participants more closely than the global mechanism suggested in the previous model (Moeller et al., 2011). Thus, the model architecture and the observation of the Gratton effect provide further evidence for cognitive control to act locally – even on processes as basic as number magnitude processing.

References

Brysbaert, M. (1995). Arabic Number Reading - on the Nature of the Numerical Scale and the Origin of Phonological Recoding. *Journal of Experimental Psychology-General, 124*(4), 434-452. doi: 10.1037//0096-3445.124.4.434

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is Numerical Comparison Digital - Analogical and Symbolic Effects in 2-Digit Number Comparison. *Journal of Experimental Psychology-Human Perception and Performance, 16*(3), 626-641. doi: 10.1037/0096-1523.16.3.626

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the Use of Information - Strategic Control of Activation of Responses. *Journal of Experimental Psychology-General, 121*(4), 480-506. doi: 10.1037//0096-3445.121.4.480

Macizo, P., & Herrera, A. (2010). Two-Digit Number Comparison: Decade-Unit and Unit-Decade Produce the Same Compatibility Effect With Number Words. *Canadian Journal of Experimental Psychology, 64*(1), 17-24. doi: 10.1037/A0015803

Macizo, P., & Herrera, A. (2011). Cognitive control in number processing: Evidence from the unit-decade compatibility effect. *Acta Psychologica, 136*(1), 112-118. doi: 10.1016/j.actpsy.2010.10.008

Macizo, P., & Herrera, A. (2012). The processing of Arabic numbers is under cognitive control. *Psychol Res*. doi: 10.1007/s00426-012-0456-6

Moeller, K., Huber, S., Nuerk, H. C., & Willmes, K. (2011). Two-digit number processing: holistic, decomposed or hybrid? A computational modelling approach. *Psychol Res, 75*(4), 290-306. doi: 10.1007/s00426-010-0307-2

Moeller, K., Klein, E., & Nuerk, H. C. (2013). Influences of Cognitive Control on Numerical Cognition-Adaptation by Binding for Implicit Learning. *Top Cogn Sci*. doi: 10.1111/tops.12015

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature, 215*(5109), 1519-1520.

Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition, 82*(1), B25-B33. doi: 10.1016/S0010-0277(01)00142-1

Nuerk, H.-C., & Willmes, K. (2005). On the magnitude representations of two-digit numbers. *Psychology Science, 47*(1), 52 - 72.

Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop Effects by Manipulating Expectations for Color Words. *Memory & Cognition, 20*(6), 727-735. doi: 10.3758/Bf03202722

Verguts, T., & De Moor, W. (2005). Two-digit comparison - Decomposed, holistic, or hybrid? *Experimental Psychology, 52*(3), 195-200. doi: 10.1027/1618-3169.52.3.195

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review, 12*(1), 66-80. doi: 10.3758/Bf03196349

Verguts, T., & Notebaert, W. (2008). Hebbian learning of cognitive control: Dealing with specific and nonspecific adaptation. *Psychological Review, 115*(2), 518-525. doi: 10.1037/0033-295x.115.2.518

Verguts, T., & Notebaert, W. (2009). Adaptation by binding: a learning account of cognitive control. *Trends in Cognitive Sciences, 13*(6), 252-257. doi: 10.1016/j.tics.2009.02.007