Modeling strategies in Stroop with a general architecture of executive control

Tomasz Smoleń (bertrand@o2.pl)
Adam Chuderski (a.chuderski@emapa.pl)
Institute of Psychology, Jagiellonian University
al. Mickiewicza 3, 31-120 Krakow, Poland

Abstract

This paper presents a preliminary work on a new architecture of executive control (DUCCA), aimed at integration and extension of some leading approaches to executive control. We present DUCCA assumptions and operation and use the architecture to simulate a few effects observed in Stroop-like task, a hallmark test of how control deals with interference. The focus of DUCCA is on how strategical use of general executive mechanisms contributes to Stroop effect. We explain also what is usually neglected in Stroop modeling: the significant individual differences in task performance.

Introduction

Executive control is implemented via numerous brain mechanisms and on different levels of neuronal organization. However, a few general flexible control mechanisms, which are involved in most of situations that require control, were also proposed in control literature (Anderson, Fincham, Qui, & Stocco, 2008; Braver, Gray & Burgess, 2007; Kane & Engle, 2003; Koechlin & Summerfield, 2007). The goal of this paper is to present a new model of executive control, called Dual Cognitive Control Architecture (DUCCA), which integrates several recent theoretical approaches to control and extends them with a few original control mechanisms. The model explains crucial effects related to interference control in Stroop-like tasks with an appeal only to general mechanisms of control, while abstracting from specific (e.g., semantical or stimulus-related) ones.

The first general function of control regards using contextual, episodic, or goal information in order to change the probability distribution of alternative actions into a one that maximizes their task-relevance (Anderson et al., 2008; Koechlin & Summerfield, 2007). Such a function is implemented in cognitive models of executive processing in two ways. In most of connectionist models, a network carrying out non-executive processing is supplemented with task (or goal) units, which modulate processing by propagating additional activation to nodes relevant to a respective task (e.g., Altmann & Davidson, 2001; Cohen, Dunbar & McClelland, 1990; Verguts & Notebaert, 2008). The control in symbolic architectures is usually implemented as control signals, stored in a goal or working memory buffer, which are matched to possible actions in order to select a next operation (Anderson et al., 2008; Meyer & Kieras, 1997). The first original aspect of DUCCA is that it integrates these two approaches into the unitary, general mechanism of top-down control, which may either directly select an action or just modulate a chance of its selection.

The second important function of executive control deals with regulation of its strength, as maintaining control for long periods of time is metabolically costly and often cognitively inefficient. Early observations indicated that control is amplified after errors. However, results like Gratton effect (i.e., the interference cost in a flanker task is 20 ms smaller in trials following incongruent stimuli, compared to ones following congruent stimuli; Gratton, Coles, & Donchin, 1992), usually observed even if errors are rare, suggested that control can be dynamically modulated on some other basis. Botvinick et al.'s (2001) conflict monitoring theory states that specialized brain mechanism (anterior cingulate cortex; ACC) performs online computing of the level of conflict between alternative responses and it increases the strength of top-down control as such a conflict arises. A more general idea is that ACC learns and reacts to a level of "risk" – conflict related error likelihood and its real-world consequences (Brown & Braver, 2007). Both cited models, however, evaluate only response representations in performing need-for-control monitoring, while conflicts can also be found between covert cognitive processes, which just influence next steps of cognition. Another new mechanism implemented in DUCCA is such a conflict monitoring procedure, which evaluates conflicts in cognitive processing (e.g., between opposing goals), which need not lead directly to any response.

Finally, DUCCA is aimed at taking into account the individual differences in control. Even healthy people differ in efficiency of control, which seem to be correlated with working memory capacity and fluid intelligence (Chuderski & Nęcka, 2010). Moreover, humans are able to regulate their mode of control by switching between top-down, proactive control and bottom-up, reactive one (Braver et al., 2007). All these differences can be expressed as differences in values of DUCCA internal control parameters, which yield qualitative changes in its simulated behaviour.

Overview of DUCCA

Cognitive operations

DUCCA is modeled as a hybrid production system. Coordination of the working of its modules is inspired by ACT-R architecture (Anderson et al., 2008). However, as the system is focused on executive functioning, "ordinary" cognitive operations have been very simplified. The system stores information received from the environment in a visual attention module, which recognizes 25 (5×5) locations on the computer screen and attends to one of them (via a focus of visual attention) at a time. Model can read symbols and some of their features (e.g., colors) from the focus. Long-term declarative knowledge is organized as a semantic network, which consists of information chunks of defined

categories (which are also chunks). A chunk contains a few slots. Each slot can contain either an atomic symbol or a reference to another chunk. One chunk can be retrieved at a time and placed in a retrieval buffer. Some information relevant to the task is actively maintained in system's focus of working memory (WM). The capacity of the focus is limited to a few (DUCCA's parameter) chunks. Contents of WM focus constitute a *context* of cognitive operations. Another structure is a goal module, which can do only one thing: it represents one chunk as a current goal of the system. Finally, a simplified motor module simulates reactions. Each response is registered and processed by a virtual key set.

Crucial for how DUCCA behaves is its procedural module, consisting of production rules, their utilities, and the mechanism for adapting utilities. Each rule is defined as a collection of conditions and a collection of actions. Conditions are imposed on both foci and the retrieval buffer. For each rule (i), a utility value (U_i) is assigned, which is updated on the basis of feedback. The utility of i tends to the expected value of feedback received after the action i. The higher U is, the more probable is the execution of a respective rule (see below).

DUCCA adapts the value of a recently executed rule in a reinforcement learning procedure, according to formula (1):

$$U_{i,t} = U_{i,t-1} + \frac{f - U_{i,t-1}}{1 + L_i}, \qquad (1)$$

where $U_{i,t}$ is a new value of utility of rule i, f is a feedback value (in range zero to one, where zero reflects "complete failure" while one means "full success"), and L_i is the reliability of a recent value of utility ($U_{i,t-1}$), estimated as the number of trials in which reinforcement of rule i has been applied. The rationale for equation (1) is that the more reliable a utility is, the less a current feedback alters this utility value. If a rule is new and L_i equals to zero, then after the first execution of a rule its utility reflects exact value of a feedback. After numerous rule's executions, its utility becomes very reliable and feedback can change it minimally. U values (in [0,1] range) reflect expected probability of reaching a goal if a rule is executed. In simple executive tasks, the reinforcement value f may be usually operationalized as the extent to which a task instruction was fullfiled, as perceived by a subject or signaled by a task.

If the environment and a context unambiguously determine an adequate action, then one rule will be matched and executed in time inversely proportional to its utility. Execution of the rule may: change the goal and/or contents of WM focus, redirect the focus of visual attention, add a chunk to the declarative memory, and send a motor command to the motor module. Then a next cycle of operation starts, until the goal is reached. However, if at least two alternative rules match (i.e., DUCCA detects a conflict related to rule selection), then executive control has to be involved in the choice of one rule from a set of matching ones (*conflict set*).

Control of cognitive operations

The first mechanism of executive control deals with evaluation of the level of detected conflict C, which is calculated according to formula (2) based on nonlinear Luce's ratio:

$$C_i = \left(\frac{\sum_{j \neq i} e^{U'_j/n}}{\sum_k e^{U'_k/n}} \right)^c, \qquad (2)$$

where j indexes all production rules in a conflict set, which yield different cognitive or behavioural consequences than a rule i of maximum utility in a conflict set, k indexes all rules in a conflict set, and n is a noise parameter. Conflict measure is thus a proportion of utilities of matching rules which are alternative to the dominant tendency for cognitive or motor processing. Parameter n controls how nonlinear is the computation of C. Note that U's instead of Us are used (the calculation of U' is explained below).

The C value determines the strength of top-down control (G_t) exerted from the goal, according to formula (3):

$$G_t = ag\big(C + E(1-C)\big) + (1-a)G_{t-1}, \quad (3)$$

where G_{t-1} denotes the strength of control in a previous cycle, E is an error value (meaning the probability that the system committed an error in a previous cycle), g is the maximum strength of control that DUCCA can exert, and a is a control adaptation parameter. C and E work in underadditive interaction. Parameter a can vary between zero (DUCCA exerts fixed strength of control and ignores conflicts and errors) and one (system uses a proportion of its maximum control strength relative to the conflict level). Theoretically plausible values of a lay above zero and below one and they mean that DUCCA adapts control to conflicts and errors, but it does so with some inertia.

The set of DUCCA's rules and their utilities may be understood as a strategy, which maps a set of possible cognitive operations onto a set of probabilities of executing these operations, in a given state of the environment and a given goal and context. Without executive control, a distribution of these probabilities reflects the effects of learning (via Us). The operation of control consists in changing this distribution into one independent on learning but dependent on how these actions are adequate to a current goal. Due to control, an agent can undertake some arbitrary behavior, even if other well-learned behavioral patterns conflict with it. The second control mechanism operates thus as modifier of rules' utilities, according to formula (4):

$$U'_i = \frac{U_i}{e^{G(1-A_{ij})}}, \qquad (4)$$

where modified utility U'_i of rule i, which is used is for conflict evaluation and conflict resolution (see below), is decreased in a function of a current control strength (G) and a value of association A_{ij} between rule i and current goal j. If either rule i is perfectly adequate to goal j (A_{ij} equals one) or control strength G is null, then U'_i equals U_i. In all other cases U_i is decreased in a nonlinear function of G and A_{ij}. If G is very high, the system just selects the rule closest to a goal. Though such a control mechanism can be judged inhibitory, our model is not committed to either an inhibitory or activational nature of control. In terms of probabilities, inhibition of one set of rules is conceptually indistinguishable from activation of an alternative set of rules.

Finally, DUCCA uses modified utilities in order to resolve a conflict among rules present in a conflict set. Analogously as in conflict evaluation formula, nonlinear Luce's ratio is exploited in formula (5) for the calculation of a probability P_i of rule i execution:

$$P_i = \frac{e^{U_i'/n}}{\sum_j e^{U_j'/n}}, \qquad (5)$$

where j denotes all rules in a conflict set, and n is a noise parameter (the same as in formula [2]). When n is very high, the rule with maximum U' always wins, while at n close to zero P equals to one divided by a number of rules in a conflict set. An important DUCCA's assumption (opposite to ACT-R theory) is that conflict resolution consumes time relative to the conflict level. Latency of conflict resolution is a multiplication of conflict value C and a scaling parameter s (i.e., $Lat = s \times C$).

Executive control in DUCCA stems from a dynamical interaction of external stimulation and its consequences (rules' utilities and goal-rule associations) and two internal mechanisms strategically adapting to the pattern of cognitive processing (conflict evaluation plus control strength modification and utility learning).

Modeling of Stroop

Stroop-like tasks, which are widely used to examine operations of executive control (MacLeod, 1991), impose interference by presenting bivalent, incongruent stimuli, which activate two cognitive processes: one dominant and the other much weaker. The task is to complete the non-dominant process. The well-known example is naming a color of a colored word that itself means an incongruent color. Interference effect, namely a positive difference between RTs for incongruent stimuli and neutral ones (e.g., colored letters X), reflects the unavoidable additional time needed for control processes to override interference from a dominant process. At the same time, control processes are usually successful, as error rates in Stroop-like tasks are low (2-10% on average). Often, a facilitation effect is also observed: people are faster for congruent stimuli (e.g., when word and its color match) than for neutral ones (MacLeod, 1991).

Some existing models

A seminal connectionist model (Cohen et al., 1990) represented alternative processing pathways as interconnected nodes in a network. Nodes for non-dominant process were associated more weakly than those of the dominant one. For the non-dominant pathway to win, an additional task-unit had to activate that pathway. A version of the model supplemented with conflict monitoring node (Botvinick et al., 2001), which controlled the amount of activation spread by the task-node in a function of conflict within a response layer, replicated above mentioned Gratton effect. It was also able to simulate an observed decrease in interference with increase in proportion of non-neutral (congruent plus incongruent) stimuli as well as smaller than interference a facilitation effect (Tzelgov, Henik, & Berger, 1992). In another

model, Verguts and Notebaert (2008) implemented conflict-modulated Hebbian learning rule, which adapted specific network connections involved in conflict resolution. The model was able to account for a decrease in interference for items often presented in incongruent contexts, in comparison to stimuli usually presented as congruent (i.e., for a so-called item-specific proportion congruency effect).

However, connectionist models are often judged atheoretical (e.g., Altmann & Davidson, 2001). They represent a modeled mechanism as just a several links between a few abstract nodes of no internal structure. A node for "redness" would be exactly the same as a node for "left keypress", even if they belong to different categories of phenomena. These models are not related to any cognitive theory (e.g., of language or memory) either. In consequence, models of tasks imposing different constraints (e.g., Stroop, flanker, or antisaccade tasks) may be described by the same network.

Some other Stroop-like models do make assumptions on related cognitive processing and focus also on more specific aspects of Stroop performance. Altmann and Davidson (2001) modeled Stroop interference as an effect of the competition between syntactic properties of the words (lemmas) and embedded this linguistic mechanism in a broader cognitive architecture (i.e., ACT-R). The model was able to explain why the separation of incongruent aspects of stimulus in time decreased interference. Lovett (2005), exploiting ACT-R's idea of utility learning of production rules, was able to explain strategical preferences of participants in chosing dominant and non-dominant processes. However, all these models would have difficulty in explaining interference effects in Stroop isomorphic tasks, which do not relate so much on linguistic properties (e.g., flankers task) or memory retrievals (e.g., Navon task).

Specific processes surely explain some part of a variance in Stroop interference, but the general executive mechanisms beyond specific processes may be responsible for the significant part of that variance. Our architecture is aimed to describe these mechanisms. However, it explains them with higher theoretical plausibility than most of connectionist models do. The model identifies different categories of cognitive structures (e.g., rules, chunks, goals) and it can ascribe meaningful contents to particular representations. Moreover, the architecture isolates executive aspects common to different tasks from task-specific characteristics. Finally, it can easily be extended with additional theoretical assumptions (e.g., ones concerning language or memory).

DUCCA's model of Stroop

We developed a model of a *generalized* Stroop-like task in order to account for a variety of results, observed within different experimental conditions and numerous versions of Stroop tasks (i.e., we abstracted from task-specific aspects).

DUCCA was supplemented with task-specific rules and chunks. There are three crucial rules for response choice: "trained", "target", and "others". The first rule leads to a skilled action, which is not proper for a task instruction. For this rule, the maximum utility ($U_{trained} = 1.0$) was set, reflecting that for adult participants such a rule had received millions of positive feedbacks. The second rule leads to instructed, but relatively poorly trained action. Its utility should be

much lower than $U_{trained}$ but still significantly above 0 (here, U_{target} = .6). "Others" represents all task-unrelevant possible processes, including ruminations and mental slips, and it should have a utility close to 0 (here, U_{others} = .1), as ruminations and slips rarely lead to positive feedbacks.

The model contains some visual and memory chunks. One important aspect of perceived stimuli is that each congruent and incongruent stimulus is bivalent: one its aspect is matched by the rule "trained", while the other aspect is matched by the rule "target". Rule "others" matches any stimulus. Memory chunks associate stimuli with proper responses. We skip other details of chunks' description.

Though the rule "target" has a low utility, it is fully associated with the goal (A_{target} = 1.0). The rule "trained" has goal association much lower than A_{target}, but still significantly above 0 (here, $A_{trained}$ = .2), as it is somehow related to what happens during the task (e.g., when congruent stimuli are frequent, it may be beneficial to use sometimes the dominant rule). Thus, in every congruent and incongruent trial there is a competition between useful rule "trained" and goal-relevant rule "target". This is modulated by the strength of control (G): the stronger control is the higher is choice probability of the rule "target". Though the rule "others" is not associated with the goal (A_{others} = .01), it may sometimes be chosen, depending on the amount of noise. When the model perceives a neutral stimulus, the rule "trained" cannot be effectively applied and only the rules "target" and "others" fall into the conflict set.

Choosing a reaction means that either the rule "trained" or the rule "target" retrieves a chunk from the declarative memory, according to stimulus features present in the visual buffer. Perceiving a feedback is applied in a simplified form, as the information about correctness of the response is displayed on the screen and processed directly.

Simulation results and discussion

The noise was set to relatively low value of 0.15, as all modeled experiments involved young and healthy participants. Parameter g equalled to 3.625 (i.e., the mean value between high- and low-WM groups, see last section). Value of c was set to 0.6, reflecting relative sensitivity to conflicts. Two time scaling parameters for each simulation were optimized to fit observed data. As these data come from differing tasks (a flanker task and two different versions of Stroop task) and experimental conditions, we did not try to fit data precisely, but we were looking for qualitative replication of the wide range of effects, instead.

Gratton effect The original Gratton et al.'s (1992) effect in flanker task is often replicated within Stroop paradigm (e.g., Kerns et al., 2004). However, for comparision with other models, we aimed to replicate the original effect (see Figure 1, left panel). In the first simulation study, 5000 runs of the model were administered with 50/50 proportion of congrent vs. incongruent trials. The ordering of trials was random. The simulated Gratton effects is presented in Figure 1, right panel. Though the model generated slightly larger interference effect, influence of previous trial was the same as in the original experiment. The Gratton effect in DUCCA comes from the rise in conflict level (C) after incongruent

trial. In a subsequent trial, C is higher than it would be if a previous trial was congruent. So, the control strength (G) is higher and it makes (via U's) the firing of the rule "target" faster, leading to decrease in RT in incongruent trials. It also makes the execution of the rule "trained" slower. As this rule may often be fired in congruent trials, it thus results in increased RT in these trials.
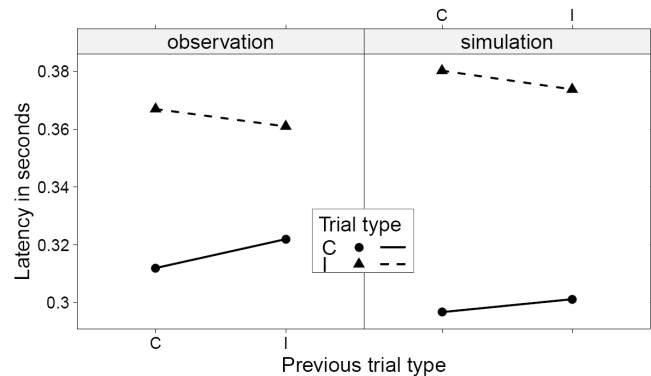

Figure 1: Left panel: data adapted from Gratton et al. (1992) on latency in congruent (C) and incongruent (I) trials as a function of a previous trial. Right panel: simulated data.

Practice on a non-dominant process The seminal study on a relation between the level of automaticity of a non-dominant process and Stroop interference was administered by MacLeod and Dunbar (1988). The participants were asked to name colors arbitrarily associated with shapes by an instruction (a task to be practiced). The shapes were colored. As expected, when color-to-name and actual color mismatched, responses took longer when colors matched or a shape was non-colored. On some days, only practice trials (naming shapes) were applied. General result was that practice on non-dominant process deacresed (and after some enourmous number of practice trials – even reversed) an interference cost. Here, we replicated the effect of five days of training (about 2000 practice trials) on interference (see Figure 2).
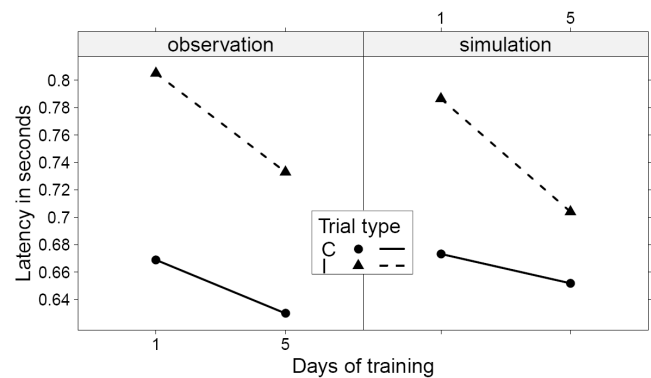

Figure 2: Left panel: data adapted from Experiment 3 by MacLeod and Dunbar (1988) on latency in congruent (C) and incongruent (I) trials as a function of practice on a non-dominant task. Right panel: simulated data.

In this simulation, which regarded a task with highly artificial non-dominant action, we used a lower value of U_{target} equal to 0.1. The practice runs resulted in decrease in utility of the rule "trained" (as it lead to errors during practice) and in increase in U_{target}. This "automatization" effect was caused by model equation (1). The change in Us caused the decrease in an interference cost, as the lower difference in utility between both rules increased a conflict value C. The increased conflict engaged more efficient control because of larger value of G. Then, 480 test runs were carried to simulate the presented data.

Proportion of incongruent stimuli, facilitation, and individual differences in Stroop performance Kane and Engle (2003; Experiment 4) observed decrease in Stroop interference as a result of decreasing proportion of congruent simuli, when neutral stimuli were absent. Moreover, it appeared that this proportion influenced the difference in accuracy in incongruent trials between low- and high-working memory capacity (WMC) participants, screened with operation span task. When proportion was low (20% congruent), both WMC groups scored around six percent of errors, with no significant advantage of WMC-high group. When incongruent trials were rare (80% congruent), error rate increased, but much more for WMC-low subjects (see Figure 3, left panel). Kane and Engle interpreted this as a result of more freqent slips of attention control of WMC-low group. In 20% congruent sequence, stimuli exogenously kept the control focused on non-dominant process and the differences in quality of internal control did not matter much. When incongruent trials were rare, only internal control could keep focus on non-dominant process and weak control of WMC-low group more often made it loose the task goal and commit more errors on incongruent trials.
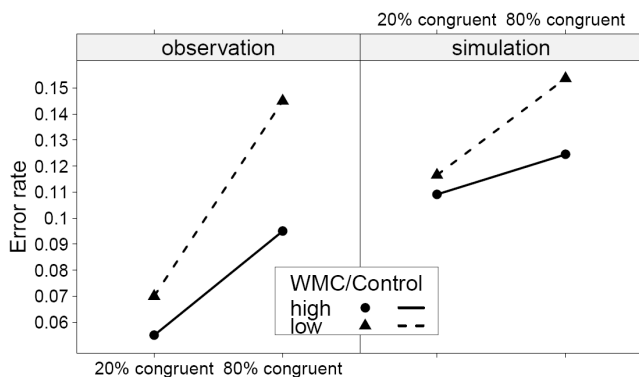
Figure 3: Left panel: data adapted from Kane and Engle (2003) on error rate in incongruent trials as a function of proportion congruent and WMC. Right panel: relevant data simulated with high/low parameter g value.

Interestingly, WMC differences did not interact with the effect of congruent trials proportion on latency: interference effect increased with increasing proportion of congruent trails, but WMC-low participants presented higher effect than WMC-high ones in both conditions of proportion congruent (see Figure 4, left panel). Kane and Engle observed also (Experiment 2) the differences in facilitation effect.

Surprisingly, WMC-low persons exhibited a larger effect (72 ms) than WMC-high ones (41 ms). On congruent trials, WMC-low participants might have more often used the dominant process to emit a response. Although use of this process did not cause errors in congruent trials, as both processes lead to the same response, it could have speeded up WMC-low participants' RTs comparing to RTs of WMC-high ones (who probably avoided the dominant process).
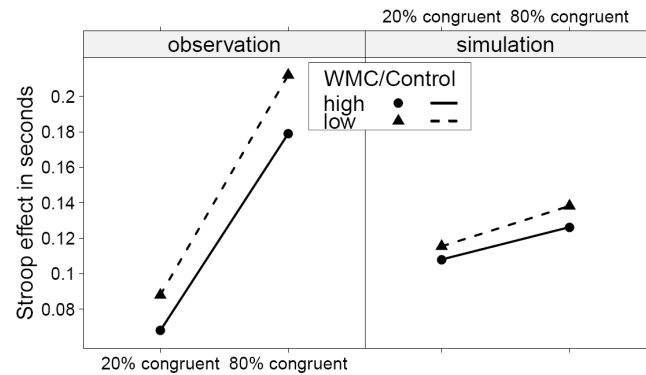
Figure 4: Left panel: data adapted from Kane and Engle (2003) on interference effect a function of proportion congruent trials and WMC. Right panel: data simulated with high/low parameter g value.

The complicated pattern of results presented in this subsection constitutes a tough test for any Stroop model. We simulated those data using either 36 congruent and 144 incongruent trials (20% congruent condition) or vice versa (80% congruent condition), following Kane and Engle's procedure in Experiment 4. The value of g parameter was set to lower value of $g = 3.5$ in order to reflect WMC-low group or set to higher value of $g = 3.75$, to reflect WMC-high group. 4320 runs of the model yielded simulated data.

All observed effects were qualitatively replicated. As in Kane and Engle's study, the effect of the proportion congruent was observed in latencies as well as in errors. In all conditions, increase in parameter g caused reasonably lower interference effects in latencies. However, the difference in g resulted in difference in accuracy on incongruent trials only when incongruent stimuli were rare. Simulated data are presented in right panels of Figures 3 and 4. In a simulation of Experiment 2, which differed slightly from Exp. 4, neutral trials were included and the values of $g = 3$ and $g = 4$ were set for WMC-low and WMC-high groups, respectively. The facilitation effect (68 ms) appeared much smaller than the interference effect (137 ms) and it fitted observed results. Also, WMC-low group scored larger facilitation effect (76 ms) than WMC-high persons (60 ms).

Figure 5 presents the indices of strategical adaptation to different (20% vs 80% congruent) task conditions. The model adapted mean level of control, rising its average level from 80% congruent to 20% congruent condition. Due to utility learning, in the more difficult condition the model amplified a utility of non-dominant rule and lowered the one of dominant rule, what increased internal conflict and thus recruited additional control. Such a strategical adaptation was less efficient when maximum strength of control was

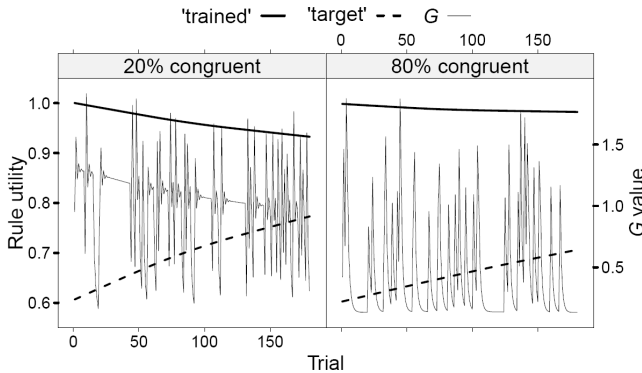limited (i.e., when *g* value was low), matching the results of WMC-low participants.

Figure 5: Internal dynamics of the model expressed as fluctuations in exerted control (*G*) and changes in utilities of the rules "trained" and "target" in two task conditions.

Two major quantitative deviations from data may be noticed: much smaller effect of proportion of congruent stimuli on latency interference and more errors committed by the model than by participants. These deviations probably result from the fact that our model captures only general aspects of control, while experimental situation involve many other general processes (e.g., expectations about the probability of events, changes in speed-accuracy trade-offs, decreased vigilance, and so on) as well as some task specific processes, all influencing interference effects. However, as a hybrid and general architecture, DUCCA can potentailly implement all these processes within more complex models.

Summary and conclusions

DUCCA, a new general architecture of executive control was presented. It was applied in order to simulate Stroop-like task. We used only a few simple assumptions of how control operates and still were able to replicate most of general effects observed in Stroop paradigm: asymetrical interference and facilitation effects, the Gratton effect, an influence of practice on Stroop effect, decrease in interference as proportion of congruent trials decreases, and the complex pattern of individual differences related to WMC.

The presented work is on a preliminary stage. Taking into account semantics and item-specific effects in executive control, linking executive mechanism to brain structures, and explaining the common variance in several executive tasks and its role in complex cognition constitute the most important future directions of DUCCA development.

Acknowledgments

References

Altmann, E. M., Davidson, D. J. (2001). An integrative approach to Stroop: Combining a language model and a unified cognitive theory. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 21-26): Hillsdale, NJ: Laurence Erlbaum.

Anderson, J. R., Fincham, J. M., Qui, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences, 12,* 136-143.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway. C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse, *Variation in working memory* (pp. 76-108). Oxford: Oxford University Press.

Brown, J. W., & Braver, T. S. (2007). Risk prediction and aversion by anterior cingulate cortex. *Cognitive, Affective, & Behavioral Neuroscience, 7,* 266, 277.

Chuderski, A., & Nęcka, E. (2010). Intelligence and cognitive control. In A. Gruszka, G. Matthews, & B. Szymura (Eds.), *Handbook on individual differences in cognition* (pp. 263-281). New York: Springer Verlag.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing model of the Stroop effect. *Psychological Review, 97,* 332-361.

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,* 480–506.

Kane, M. J., Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47-70.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., 3rd, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023–1026.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences, 11,* 229-235.

Lovett, M. C. (2005). A strategy-based interpretation of Stroop. *Cognitive Science, 29,* 493-524.

MacLeod, C. M. (1991). Half a century of a research on the Stroop Effects: An integrative review. *Psychological Bulletin, 109,* 163-203.

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 126-135.

Meyer D., Kieras D. (1997) A computational theory of executive cognitive processes and multiple-task performance. Part 1: Basic mechanisms. *Psychological Review, 104, 3-65.*

Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Memory & Cognition, 20,* 727-735.

Verguts, T., & Notebaert, W. (2008). Hebbian learning of cognitive control. *Psychological Review, 115,* 518-525.