Dynamic competition and binding of concepts through time and space

Jean-Charles Quinton, Annique Smeding

Dynamic competition and binding of concepts through time and space

Jean-Charles Quinton[12], Annique Smeding[3]

[1] *Clermont University, Pascal Institute, Clermont-Ferrand, France*
[2] *CNRS (UMR 6602), Pascal Institute, Aubière, France*
[3] *Savoie Mont Blanc University, LIP/PC2S, Chambéry, France*

Abstract

Models of implicit stereotypes (e.g. association of male with math, or female with language) usually explain the faster responses observed for stereotype congruent trials in the Implicit Association Test (IAT) by requiring a fundamental opposition between the male-female concepts (or math-language), limiting the decision-making dynamics to abstract dimensions. This paper introduces alternate models exploiting the sensorimotor dimensions of the IAT, which naturally account for the opposition between concepts, because typically mapped on opposite corners of the screenspace and different response actions. In addition to the emergence of the IAT effect, dynamic characteristics of the decision-making process within these models are tested against human data, obtained with a mouse-tracking adapted IAT procedure.

Introduction

Stereotypes have been defined as "associations of a social group concept with one or more (nonvalence) attribute concepts" (Greenwald et al., 2002). For instance, in the case of a gender-math stereotype which will be the target illustration of this paper, *math* is more strongly linked to the social group of males than to females (Nosek, Banaji, & Greenwald, 2002). Such associations are constitutive of a social knowledge structure, formed and reinforced on the long term, and can be measured at the implicit level. This is the case in the Implicit Association Test (IAT; Greenwald et al., 2002), where stimulus words such as *man* and *equation* (appearing in the middle of a computer screen) must respectively and quickly be categorized into target categories *male* and *math* (appearing at the top left or top-right corner of the screen, depending on experimental condition). The IAT effect corresponds to the slowing down of responses and increase in errors when moving from stereotype-congruent trials (*male* and *math* categories grouped on one corner of the screen and opposed to *female* and *language* categories, grouped on the other corner) to stereotype-incongruent trials (*male* and *language* grouped and opposed to *female* and *math*). This opposition occurs both at the visual level (target categories being located on opposite sides of the computer screen) and at the motor level (a single key press being used for categories on the same side of the screen). In the theoretical framework behind IAT effects (i.e. Unified Theory, Greenwald et al. (2002)), it is posited that associations between stereotypical concepts are reinforced (balance-congruity principle), and that bipolar opposition of concepts such as *male* and *female* (imbalance-dissonance principle) must be present. A direct competition is also invoked in more recent developments, where attribute concepts as well as social group concepts are made mutually inhibitory in neural network based computational models (Freeman & Ambady, 2011).

On the one hand, the IAT literature thus supports that social knowledge (i.e., stereotypes) can influence our behavior and decisions. On the other hand, the dynamic categorization literature shows that categories themselves, or at least the binding or competition of their constitutive features or elements, directly depend on the task to be performed. Words such as *grandmother*, *grandfather*, *daughter* and *son* (all part of the gender-math IAT stimuli set) will be bound differently depending on whether target categories are *male* vs. *female* or *young* vs. *old*. Even if we cannot rule out the possibility that competition between concepts (such as *male* and *female*) may be permanent, competition must be modulated by the context to not hinder performance when target categories change. For categorization tasks with non linguistic stimuli, including stereotype related tasks (e.g. Freeman & Ambady, 2011), the competition and binding of features occur in space and time. Indeed, and especially for visual stimuli, objects are often described by a set of more or less complex features, and active capabilities of the human perception system allow to sequentially focus on, or search for, discriminative features (Quinton, Catenacci Volpi, Barca, & Pezzulo,2014). These discriminative features depend on the target categories, so that even the exact same stimuli might be explored, processed and categorized differently depending on context and a prioris (Kietzmann, Geuter, & König 2011). Therefore both the sensorimotor apparatus and the experimental design put constraints on the dynamic decision process.

Taken together, these two literatures acknowledge the wide range and huge amount of information that can be quickly integrated, as well as the complementary processes that influence decisions. However, they diverge when it comes to the characteristics of competition. Indeed, the literature on social stereotypes assumes a permanent *male-female* competition at the conceptual level, while dynamic categorization research indicates that this competition may be temporary and may depend on task constraints. This partial incompatibility is not surprising, as Unified Theory was developed to provide theoretical support to the IAT effect, with a disciplinary focus from social psychology. Reciprocally, research on dynamic categorization originated in cognitive psychology and neuroscience, and is only progressively expanding to encompass social factors.

Our aim is to reconcile the two literatures around these notions of binding and competition, by adopting a dynamical system perspective and considering the different levels at which competition might occur (i.e., conceptual *vs.* perceptual). In the present research, we combine an empirical study with a computational study to test the hypothesis that dynamic competition is required to correctly model human behavior, but that such competition is not required at the conceptual level to explain the emergence of the IAT effect. The next section introduces the empirical procedure of the present research. This procedure relies on a mouse-tracking adapted IAT (Yu, Wang, Wang, & Bastin, 2012) specifically designed to collect continuous data and to allow the analysis of dynamical characteristics of categorization processes under the influence of stereotypes. A set of connectionist models are then introduced as proofs of concepts. These models are tested against human data, demonstrating that the IAT effect might actually depend on the experimental design and response modalities, involving sensorimotor dimensions and not being limited to abstract concepts. Nevertheless, the same models also suggest that the effect is not an artifact of the procedure, as it also requires and reflects stereotypical associations between concepts.

Empirical Study

Method

The standard IAT procedure involves a comparison of latencies and error rates between congruent and incongruent trials. It therefore provides little information about the dynamics of the decision, nevertheless required to reach a deep understanding of the processes underlying categorization under the influence of stereotypes, and thus to test the associated theoretical or computational models. On the contrary, mouse-tracking, i.e. the tracking of computer mouse movements, has proven to be effective in revealing the online decision-making dynamics underlying (social) categorization (Freeman & Ambady, 2011). Except for a research targeting valenced associations (Yu et al., 2012), the IAT literature has remained silent with respect to these theoretical and methodological advances.

To investigate the dynamics of categorization and obtain human data to confront computational models to, we used a mouse-tracking adapted IAT procedure investigating gender-math associations (see Fig.1 for the setup, (Nosek et al. 2002) for details on the original procedure). Twenty-two female humanities undergraduates took part in the study and were chosen because of the typical IAT effect observed in this sample (Smeding 2012). Data were part of a larger study.

Results

Trials where the final categorization was erroneous (6.7%) were removed. Coordinates (x,y) of mouse trajectories were remapped rightward in [-5, 5]×[-0.5, 9.5] starting from (0,-0.5) Analyses on global statistics over trajectories were performed using mixed-effects models, with standard temporal and geometrical mouse-tracking statistics (reflecting attraction toward the non-selected response) regressed on trial congruency: RT (Response Time), AUC (Area Under the Curve) and MD (Maximum Deviation). The effect was significant for RT, $B = 65.5$, $t(2440) = 3.21$, $p = .001$. As expected, RT for female humanities students was higher for incongruent trials (M = 901ms, SD = 452) than for congruent trials (M = 832ms, SD = 347). Consistent yet not significant results were obtained for MD, $B = 0.18$, $p = .05$, and AUC, $B = 0.07$, $p = .07$, reflecting a larger deviation for incongruent trials(see Fig.2).

Turning to the dynamical characteristics of trajectories, we computed the average speed profile on the X axis, on which categories are opposed (Fig.3). ANOVA was performed between congruent and incongruent trials for all participants. Speed profiles showed a significant difference in mean speed amplitude, $B = -0.21$, $t(42) = -2.22$, $p = 0.03$, with slower movements for incongruent trials ($M = 1.44$, $SD = 0.35$) than congruent trials ($M = 1.65$, $SD = 0.26$). Similar results were obtained for the center of mass of the distribution, reflecting a shift in time $B = 130.1$, $t(42) = 3.27$, $p = .002$, with slower onset for incongruent trials (M = 815ms, $SD = 174.3$) than congruent trials (M = 685ms, $SD = 66.9$).

Computational Study

Method

Human data from our mouse-tracking IAT demonstrate that hypotheses about the interpretation of stimuli influence the behavior in real-time and compete from an early stage of the decision-making process (Spivey, 2007). To develop the connexionist models introduced in this section, we have introduced representations of these hypotheses and made them interact. The behavior and final decision are thus emergents of the interactions between the task environment (e.g. stimuli) with a network of concept nodes $c_i \in$ {*math, language, male, female*...} (see Fig.4). Pairs of nodes (c_i, c_j) may interact through a link of strength w_{ij}, which can be related to a wide variety of representations, including (a) social knowledge (e.g. *math-male* stereotype), (b) linguistic knowledge and task constraints (e.g. *father-male* association) and (c) knowledge about experimental design constraints (e.g. *male-left* mapping on the screen). To each node c_i is associated a dynamic activity a_i, governed by Eq.1. This is a continuous time version of existing stereotype models based on recurrent neural network (Freeman & Ambady, 2011), including a relaxation term ($-a_i$), the nonlinear influence of connected nodes ($w_{ij}\, \sigma(a_j)$) and noise (ε_i).

$$\tau \frac{\partial a_i}{\partial t} = -a_i + \sum_j w_{ij} \sigma(a_j) + \varepsilon_i \qquad (1)$$

To make the comparison of models relevant to the stereotype literature, we only initially consider excitatory links between nodes (see Greenwald et al., 2002). As we are interested in generative models able to reproduce mouse trajectories toward responses, we however need to introduce additional nodes for *left* and *right* associated to sensorimotor behavior (linked to either target location or movement orientation). The choice of approximating continuous dimensions by two opposing nodes was made for simplification, but results could be extended by using dynamic neural field models (Quinton, 2010). The current model can indeed be considered as the limit case of the classical neural field equation, that is extensively used to account for decision dynamics on continuous feature spaces.

At any time during the simulations, the mouse position is updated based on a speed vector, computed as a weighted sum of vectors pointing to the targets from the current position, scaled to roughly reproduce the human hand dynamics, and with weights set to the activity of the respective *left* or *right* nodes (Quinton et al., 2014). Taking the maximum likelihood over any activity distribution (e.g. *left-right*) instead leads to initial jerky movements not found in human data. Adding an activity threshold from which movement is initiated in part solves this problem, but the simulated data then cannot account for the progressive bifurcation (with initial straight-up movement) and hesitations (i.e. changes in direction) observed in human trajectories, showing that a continuous integration and competition of hypotheses seems required, either at the neuronal or body apparatus level.

To further test this hypothesis, we introduce 3 different models in addition to the one previously described, that lacked inhibition (model A). These models simply differ by the level at which competition takes place, introducing an inhibitory link between nodes either at the conceptual level (model B), sensorimotor level (C), or both (D).

Results

In order to ensure a fair comparison, a metaheuristic was applied to select the best parameters for each individual model (see (Quinton, 2010) for methodological details). Optimization criteria were selected to reflect the capability to generate: (i) correct responses, (ii) convergence to a clear decision when the time constraint is relaxed, (iii) smooth trajectories with progressive bifurcation, (iv)significant differences between the congruent and incongruent blocks on RT, AUC and MD.

Taken separately, criteria were correctly satisfied by all models, yet optimization led to different parameter values. Satisfying criteria (iv) alone for instance means that all models are able to account for the IAT effect. However, taking all criteria together leads to different results, with model A generating wrong decisions or chaotic trajectories, and model B not satisfying criteria (ii). The time constraint is a requirement of the mouse-tracking paradigm, because if relaxed, the participant may decide before engaging in action. With this in mind, authors of previous computational studies argue that conceptual nodes could settle into an attractor with a mixed activity, allowing non stereotypical attributes to remain marginally active, while still being able to reach a decision (see Freeman & Ambady, 2011 for a type B model, Greenwald et al., 2002 for type A). Although we defend the same position, any model willing to account for human decision-making should also be able to generate straight trajectories if movement starts once the dynamical system has converged to its fixed attractor. This is only the case for model C and D, where a distributed activity is maintained at the conceptual level (e.g. $a_{male}=0.8$, $a_{female}=0.2$), but where competition at the sensorimotor level guarantees a clear decision after an infinite amount of time (e.g. $a_{left}=0$, $a_{right}=1$).

Discussion

The present research relied on a mouse-tracking adapted IAT to investigate the dynamical processes underlying implicit gender-math associations. In addition to replicating the IAT effect, our results demonstrate that, beyond reaction times, the influence of congruency influences the whole behavior, as reflected by early mouse movements. Mouse trajectories were correctly simulated by a set of generative dynamical connectionist models, using a few conceptual and sensorimotor nodes, with block dependent binding.

However, only models including competition at the sensorimotor level were able to satisfy all criteria characterizing human decision-making. Although we cannot rule out that the modeling choices are responsible for the lack of solution satisfying all criteria when competition is absent or limited to the conceptual level, the computational models were designed to be as similar as possible to references from the literature. In virtue of Occam's razor, also known as the law of parsimony, competition at the sensorimotor level alone should be preferred, as it is sufficient to explain the decision dynamics. Additionally, not requiring fixed inhibitory links between abstract concepts (such as *male* and *female*) reconciles the IAT literature with the flexible binding capabilities demonstrated by humans in dynamic decision-making.

Acknowledgments

This research was supported by the French National Center for Scientific Research (CNRS) through the PEPS « Humanités – Mathématiques – Sciences de l'information » scheme and Université Savoie Mont Blanc through the International Relations Grant (RI)

References

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, 118, 247-279. doi: 10.1037/a0022327

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3–25. doi: 10.1037/0033-295X.109.1.3

Kietzmann, T. C., Geuter, S., & König, P. (2011). Overt visual attention as a causal factor of perceptual awareness. *PloS one,* 6, e22614.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = Male, Me = Female, therefore Math≠Me. *Journal of Personality and Social Psychology, 83,* 44–59. doi: 10.1037/0022-3514.83.1.44

Quinton, J-C. (2010, July). Exploring and optimizing dynamic neural fields parameters using genetic algorithms. In 2010 International Joint Conference on Neural Networks (IJCNN), 1-7. IEEE. 10.1109/IJCNN.2010.5596706

Quinton, J-C., Catenacci Volpi, N., Barca, L., & Pezzulo, G. (2014). The cat is on the mat. Or is it a dog? Dynamic competition in perceptual decision making. *IEEE Transactions on Systems, Man and Cybernetics: Systems, 44,* 539–551. doi: 10.1109/TSMC.2013.2279664

Smeding, A. (2012). Women in Science, Technology, Engineering, and Mathematics (STEM): An Investigation of Their Implicit Gender Stereotypes and Stereotypes' Connectedness to Math Performance. *Sex Roles, 67, 617-629*. doi: 10.1007/s11199-012-0209-4

Spivey, M.J. (2007). *The Continuity of Mind*. New York: Oxford University Press.

Yu, Z., Wang, F., Wang, D., & Bastin, M. (2012). Beyond reaction times: Incorporating mouse-tracking measures into the implicit association test to examine its underlying process. *Social Cognition, 30*, 289–306. doi: 10.1521/soco.2012.30.3.289

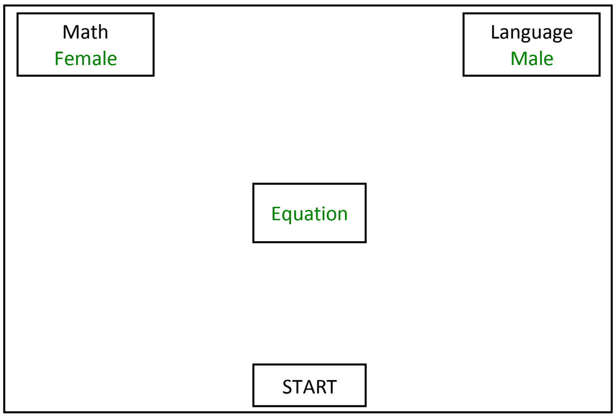

Figure 1. Setup used in both experimental and computational studies. Participants were instructed to click the START button and to categorize as fast as possible each stimulus word (appearing at the center of screen). For this purpose, participants had to move the computer mouse over the chosen category (top-left or top-right corners of the screen).

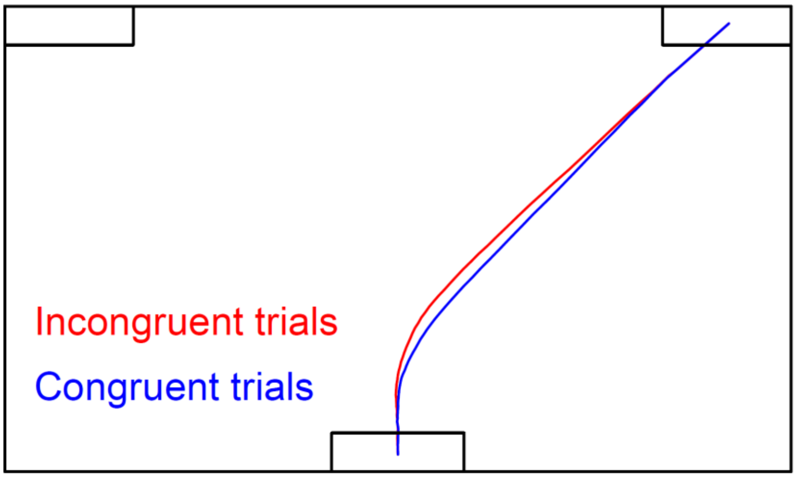

Figure 2. Averaged trajectories for congruent and incongruent trials for all participants, superimposed on a schematic representation of the setup.

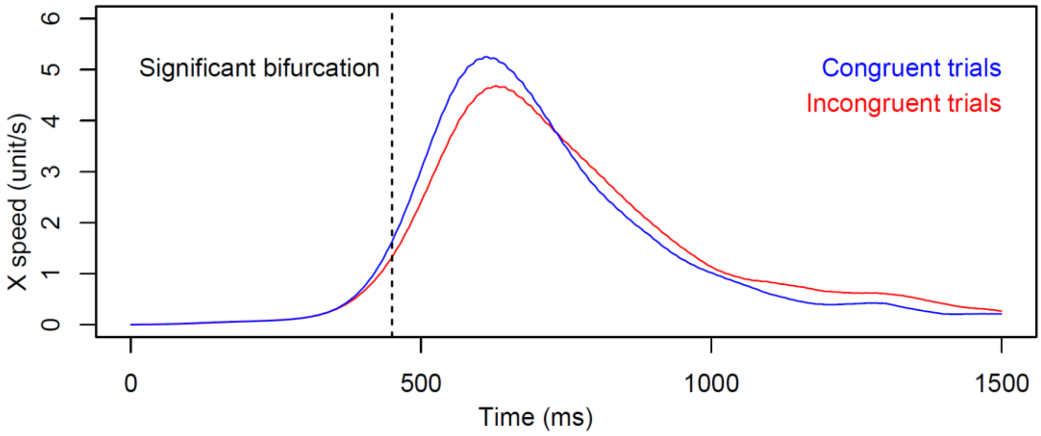

Figure 3. Speed profiles of mouse trajectories for congruent and incongruent blocks, computed along the X axis and averaged over all participants and block trials. These profiles show significant differences in amplitude and onset. Applying a mixed effects ANOVA model on each timestep also reveals that the bifurcation in speed profiles becomes significant from an early stage of the movement, at 450ms, $B = -0.54$, $t(2435) = -2.04$, $p = 0.04$.

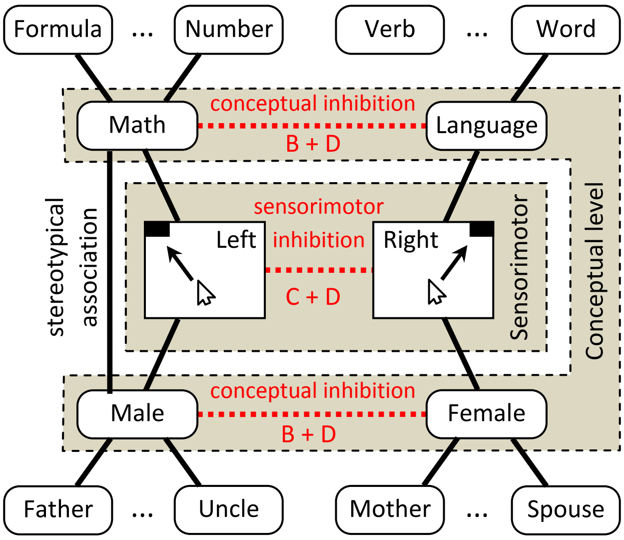

Figure 4. Architecture of the dynamical connectionist model used in the computational study. During the simulation, activity of the stimulus node is maintained to 1.0, generating indirect activations from first-order links (e.g. probability of *math* category to be relevant when observing the word *father*, even if there is no direct *math-father* link). In the end, mouse movements are generated from the activities of the *left* and *right* nodes associated with the onscreen targets. The dynamical binding between the *math* and *male* through *left* illustrated here corresponds to stereotypical trials. Models A, B, C and D differ by the inhibitory links introduced (dashed lines).