Available online at www.sciencedirect.com
ScienceDirect
Neural Networks 19 (2006) 1455–1457
Neural Networks
www.elsevier.com/locate/neunet
2006 Special Issue

On the social psychology of modelling

Bernhard Hommel

Department of Psychology, Cognitive Psychology Unit, Leiden University and Leiden Institute for Brain and Cognition, Postbus 9555, 2300 RB Leiden, The Netherlands

Modelling is often considered to represent the most scientific way to theorize, and the amount of modelling in theorizing is taken to reflect the maturity of a scientific discipline. Indeed, modelling papers are often introduced by pointing out the great things models can do for us: they make theoretical assumptions (more) explicit, they can be tested against numerous aspects of the data (and not just, say, reaction time differences in condition means) and generate new, possibly (and ideally) counterintuitive hypotheses that motivate novel experiments and research directions. But if this is the case, how come that models play such a minor role in guiding our research and our thinking, which is true at least for the areas I am working in? In the following, I will briefly attempt to substantiate my impression that most (but not all) models do play a minor role and then consider the possible reasons for that. My (hopelessly overgeneralizing) claim will be that the social context of modelling may often limit the modeller's creativity and the impact of modelling on empirical research.

1. More than you want to know: Which assumptions do models make explicit?

It is certainly true that running a network model or simulation of cognitive processes requires numerous assumptions, and that the simulation would not work without making them explicit. Unfortunately, however, many of the assumptions a simulation needs to run tend to be of little theoretical interest, whereas assumptions that would allow for a deeper theoretical insight into an empirical phenomenon and its underlying mechanism are surprisingly often underspecified.

For instance, models of stimulus–response compatibility (i.e., of the observation that some stimulus–response combinations allow for better performance than others) have done a good job in accounting for the outcome patterns of studies using various combinations of stimulus and response sets and stimulus–response pairings (e.g., Kornblum, Stevens, Whipple, and Requin (1999), Zorzi and Umiltà (1995)). However, one of the main factors that discriminates between alternative models turned out to be whether Gaussian or non-Gaussian noise was added to induce variability in the simulated reaction times and whether it is added to input or output nodes (e.g., Kornblum et al. (1999)). Even though this is certainly an exciting issue, I'm not convinced that this is what compatibility researchers are concerned about most. At the same time, all models "explain" the central theoretical issue of why and under which circumstances task-irrelevant stimuli can automatically trigger action tendencies by simply drawing arrows between stimulus and response representations. But *that* stimuli can prime feature-overlapping actions (as indicated by such arrows) is something we already knew way before computational models entered the scene, so that the fact that they contain a link between the corresponding internal codes does not seem to add much to our theoretical insight. What we still need to know is *how* stimuli and actions are cognitively represented and the mechanism by which their representations bring about compatibility phenomena (Hommel, Müsseler, Aschersleben, & Prinz, 2001). There are more examples, such as Logan's (1988) instance theory, which devotes only a few sentences to what an instance actually is and what kind of codes it comprises. My main point here is that computational models are often mathematically much richer and much more ambitious than they are conceptually. If one aims at predicting data, this is certainly the most appropriate strategy, but if one is looking for theoretical insight beyond the algorithmic level, models are often disappointing.

2. What do models buy us really?

Even though it is true that complex models make it possible to test complex aspects of the data, their validity is notoriously difficult to objectify. Model fits, as such, tell us little about a

E-mail address: hommel@fsw.leidenuniv.nl.

0893-6080/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2006.09.008

model's flexibility and its restrictions, and the frequent lack of systematic cross-model comparisons makes it difficult to relate them to each other (Roberts & Pashler, 2000). Indeed, models are commonly tested against data from extremely restricted and artificial tasks, and it remains unclear how they would do in slightly different (but theoretically related) tasks and paradigms, not to speak of the "real world".

But let us leave aside such technicalities for a moment and concentrate on the often proposed heuristic potential of models. The more complex theoretical ideas become, the more difficult it is to think them through fully and to derive clear-cut predictions. The strength of a computational model, so it is often argued, is that it can do the thinking-through for us and generate the predictions automatically. If this were true, computational models would be of enormous value and it should be easy to predict the most surprising, counterintuitive results. But such predictions are the exception rather than the rule. The field of stimulus–response compatibility is full of surprising observations, such as the impact of task-irrelevant stimuli (Simon & Rudell, 1967), the strong impact of task intention (Hommel, 1993) and task preparation (Valle-Inclan & Redondo, 1998), or the impact of action planning (Müsseler & Hommel, 1997) and action execution (Bekkering & Neggers, 2002) on perception. Some, but not all, of these observations were predicted from theoretical approaches, but computational models were not among them. Indeed, it is probably fair to say that the available computational models did not generate a single finding that could rightly be described as counterintuitive or unexpected (from alternative theoretical frameworks), and I'm afraid that the same goes for quite a number of models in other cognitive areas. Hence, if we follow the suggestion of Honing (in press) to replace the goodness-of-fit criterion of model testing by a measure of "theoretical surprise", the average computational model does not score highly.

3. Great exceptions from a disappointing rule

Does all that mean that models are of no use? Of course not! Indeed, there are computational models that seem to do exactly the job that modellers sell us as the main function of a model: to guide empirical research and make interesting predictions, integrate its results to improve the model, make even more interesting predictions, and so on. One example is the Princeton group around Jonathan Cohen, which over the years has built up a very intense and interactive working relationship between researchers involved in theorizing, modelling, behavioral experiments and brain imaging studies (e.g., Botvinick, Braver, Carter, Barch, and Cohen (2001)). Empirical observations have motivated the creation of models, which are systematically tested and refined by model-guided empirical investigations. Another example is the increasing collaborative network around Claus Bundesen, which puts Bundesen's computational theory of visual attention to all sorts of tests, such as predicting the performance of patient groups by lesioning the computational model in systematic ways (Duncan, Bundesen, Olson, Humphreys, Chavda, & Shibuya, 1999). What went right in these groups that elsewhere went wrong?

In my view, computational modelling is like socialism: being merely a method, it is neither good nor bad—the question is only of what you do with it. And what you can do with it depends on your scientific environment and social network. Lonely thinkers are no good modellers; they may publish their model, but it will not likely be picked up and systematically tested by empirical researchers. This is simply because it is often way too much trouble for a non-modeller to make all the modifications necessary to derive predictions for tasks that only slightly deviate from those originally used to fit the model. However, modellers that either conduct experimental research themselves or that are integrated into a wider research network that also includes experimental researchers and, ideally, non-modelling theoreticians, may well be very influential and of great help in both guiding and integrating empirical research. Mainly two factors may be responsible for that: the possibility for experimentalists to interact directly with the modeller and get the necessary advice in deriving concrete predictions and the personal commitment on both sides to make the model transparent and flexible enough to adapt it to new tasks and theoretical challenges, and to tailor the empirical designs so to provide systematic tests, especially of the model's weak spots. My claim is thus that it may not be so much the (original) quality of a given model that determines its impact on empirical research, but more the social environment in which the modeller is embedded. The tighter this network, the better the aimed-at cycle of modelling–testing–modelling–testing ... really works, which in the end will also determine the quality of the model.

Acknowledgments

Support for this research by the European Commission (PACO+, IST-FP6-IP-027657) is gratefully acknowledged.
Correspondence and requests for materials should be send to Bernhard Hommel, Leiden University, Department of Psychology, Cognitive Psychology Unit, Postbus 9555, 2300 RB Leiden, The Netherlands, hommel@fsw.leidenuniv.nl.

References

Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science*, *13*, 370–374.

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

Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Chavda, S., & Shibuya, H. (1999). Systematic analysis of deficits in visual attention. *Journal of Experimental Psychology: General*, *128*, 450–478.

Hommel, B. (1993). Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information. *Psychological Research*, *55*, 270–279.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–878.

Honing, H. The role of surprise in theory testing: Some preliminary observations. In *Proceedings of the international conference on music perception and cognition* (in press).

Kornblum, S., Stevens, G., Whipple, A., & Requin, J. (1999). The effects of irrelevant stimuli I: The time course of S-S and S-R consistency effects with Stroop-like stimuli (DO Type 4 task), Simon-like tasks (DO Type 3 task), and their factorial combinations (DO Type 7 task). *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 688–714.

Logan, G. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.

Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 861–872.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, *107*, 358–367.

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*, 300–304.

Valle-Inclan, F., & Redondo, M. (1998). On the automaticity of ipsilateral response activation in the Simon effect. *Psychophysiology*, *35*, 366–371.

Zorzi, M., & Umiltà, C. (1995). A computational model of the Simon effect. *Psychological Research*, *58*, 193–205.