Psychological Research (2007) 71: 1–3
DOI 10.1007/s00426-005-0030-6

EDITORIAL

Bernhard Hommel · Bruce Milliken

Taking the brain serious: introduction to the special issue on integration in and across perception and action

Received: 1 January 2005 / Accepted: 14 August 2005 / Published online: 30 November 2005

Our world consists of object and events. As an almost defining criterion of how an object is constituted, the features of an object are all attached to each other: the redness of the cherry is exactly where I find its roundness and the softness of its skin, and all these features are there at the same point in time. Things get a bit more complicated if we consider larger objects, like a sailing ship—where the whiteness may be restricted to the sail, and temporally extended events, where features may gradually change over time. But the basic organizing structure is always the same, namely, things that belong together share spatiotemporal coordinates. This observation may seem trivial, but it is in fact striking when one considers that our brain has a very different organizational structure.

One important difference is that the brain codes the different features of an object or event in different modules: color is coded in color maps, aspects of shapes in shape maps, motion in motion maps, pitch in pitch maps, and so on (e.g., DeYoe & Van Essen, 1988; Schreiner, 1995). Some of these maps are spatially organized, and thus retain at least this aspect of how our world is organized, but the sheer multiplicity of representational maps demonstrates that properties that are perceived together in the world are not represented together in the brain. A second important difference concerns the temporal aspect of perception. The computation and transmission of different features within the same modality and in different modalities varies widely, so that the time at which a representational code is activated is an unreliable estimate of when the represented event occurred in the world (Zeki, 1993). The implication is that things that occur at the same point in time are not reliably represented by temporal aspects of their representations.

Our brain is thus facing the problem that the spatiotemporal characteristics of objects and events in our world are distorted in the process of coding them, so that representations of things cannot be mere copies of their attributes in the world. But how else does our brain keep things together? This question has become known as the *binding problem*, which was introduced to cognitive psychology by Alan Allport (e.g., Allport, Tipper, & Chmiel, 1985) and Anne Treisman (e.g., Treisman & Gelade, 1980). As these authors pointed out, distributed representations such as those apparently preferred by the human brain need some mechanism that cross-links or binds the representational codes that refer to the same object or event. Assume, for instance, you are facing a green square and a red circle. Processing them will activate the codes <green>, <red>, <rectangular>, and <round> in your perceptual system, so that some mechanism must sort those codes in such a way that <green> will be related to <rectangular> but not <round>.

The binding problem has been addressed in numerous ways and we will not be able to discuss them in detail (cf., Treisman, 1996). Yet it is interesting to see that the most accepted solutions to the binding problem suggest that the brain in some sense tries to imitate or reconstruct the spatiotemporal organization of the world. Spatial organization is reintroduced by restricting a hypothetical attentional spotlight to a single object and/or location in space, so that only those feature codes that belong to this object and/or the location it occupies get (strongly) activated (e.g., van der Heijden, 1992; Treisman & Gelade, 1980). Temporal organization is reintroduced by creating temporal integration windows (implied but commonly not discussed by space-related models), during which the available information about the attended object, or at the attended location, is sampled. The length of the assumed integration window

B. Hommel (✉)
Department of Psychology, Cognitive Psychology Unit,
Leiden University, P.O. Box 9555, 2300 Leiden, The Netherlands
E-mail: hommel@fsw.leidenuniv.nl

B. Milliken
Department of Psychology, McMaster University, Hamilton,
Ontario, Canada

is commonly (implicitly or explicitly) assumed to exceed the differences in the times required for registration and preprocessing of the to-be-integrated features, so that all of their codes are simultaneously available for integration.

The binding problem or, better, the binding problems (Treisman, 1996) are usually discussed in the context of visual perception, presumably because the visual system is particularly well studied and because the multiplicity of specialized feature maps it apparently houses makes the consequences of distributed processing particularly obvious. But there is no reason to believe that binding problems are restricted to visual perception (Hommel, 2004). First, there are other sensory modalities that code incoming information in a distributed fashion, which raises the question of how that information is bound. Second, we commonly face multimodal events, such as the people we communicate with, which raises the question of how features in different modalities interact with each other to create coherent representations. Third, planning an intentional action, such as grasping a cup of coffee, requires the integration of multiple codes specifying, say, the effector to be used, the length and endpoint of the reach, the force of the grasp, etc., which raises the question of how they are related to each other. And, fourth, the things we perceive and the actions we perform must be systematically related to each other to make sure that, say, the appropriate information is channeled to our feet and mouth when walking and talking at the same time. Hence, some integration spanning perception and action is necessary, and this integration somehow needs to be tailored to the current task and context.

Obviously, then, our brain's strategy to code information for perception and action in a distributed fashion poses numerous problems, which apparently are smoothly solved by integration mechanisms that allow us to function as efficiently as we do. But it is fair to say that we know extremely little about how these integration mechanisms work. Some progress has been made in the area where it all began, in visual perception (e.g., Robertson, 2003), and first steps have been taken to tackle perceptual integration spanning more than one modality (e.g., Calvert, Spence, & Stein, 2004). However, hardly anything is known about how action parameters are integrated (Stoet & Hommel, 1999) or how actions are bound to perceptual objects. Evidently, these blind spots imply others, so that we have no idea whether the integration mechanisms are related (e.g., does integration within a modality precede cross-modal integration, does cross-modal integration precede sensorimotor integration?) and how, if at all, the integration processes are controlled and adapted to the task at hand.

This state of affairs or, perhaps better, this widespread ignorance has motivated the present special issue on feature integration in and across perception and action. This special issue has three aims. First, we intend to draw attention to the fact that feature-binding problems are by no means restricted to visual perception but must also be solved in other modalities, in intermodal perception, in action planning, and in perceptuo-motor organization. Second, we want to provide an overview of how integration processes in other than the visual domain are currently investigated. Given the space limitations we of course do not pretend that this overview is exhaustive and it is likely to reflect our empirical preferences and theoretical biases. However, we do hope that the present contributions provide a stimulating potpourri of research strategies that demonstrate how feature integration in a wider sense can be fruitfully investigated. And, third, we consider the work presented in this issue to show how neuroscientific insights can successfully stimulate and shape cognitive-psychological research, and thus make cognitive theorizing more biologically realistic, without requiring one to give up strictly functional explanatory goals. In this sense, the present issue is a straightforward extension of Allport's and Treisman's considerations.

The issue is opened by Navarra and Soto-Faraco, who investigate whether and how audiovisual feature integration supports the processing of a second language. They show that intermodal integration provides additional information about how heard sounds are produced by a speaker and, thus, allows perception to be informed by action-related information. A similar topic is addressed by Hamilton, Joyce, Flanagan, Frith, and Wolpert, who studied the ability of observers to judge the weight of a box from watching another person lifting it. They find that observers rely most on the duration of the lifting movement and make less use of the durations of the grasp, which however is the best predictor of lifted weight in the kinematics of natural box lifting behavior. This outcome suggests that perception is informed by action features (i.e., through information derived from what action looks like) but that weight perception relies on more than kinematically derived cues.

Fagioli, Hommel, and Schubotz investigate how planning an action affects the integration of visual features. They provide evidence that planning a particular action primes the processing of features that are informative for that action: planning to grasp an object facilitates the processing of shape, whereas planning to point to an object facilitates the processing of location information. Action plans not only bias the processing of perceptual events, they also become integrated with these events into an episodic memory trace. This is demonstrated by Oriet, Stevanovski, and Jolicœur, who show that visual stimuli that are perceived while the plan to carry out a manual key press is maintained become linked to this plan, so that changing the relationship between plan and stimulus in the next trial impairs performance.

This theme is continued in the contributions of Hommel and of Milliken and Lupiàñez. Hommel shows that repeating task-relevant stimulus features from one trial to another increases the likelihood of repeating the response in a free-choice task. This suggests that performing a response to, or in the context of, a stimulus creates a binding between the codes of the response and the stimulus features. It also suggests that task-relevant features are more likely to be selected for integration and/or more effective in retrieving episodic bindings. Further constraints are investigated in the study of Milliken and Lupiáñez. They study the facilitatory and inhibitory effects of word primes on word identification and provide evidence that such primes yield two types of effects: a positive effect due to activating the same word representation twice within a short time interval and a negative effect due to presenting the same stimulus twice but requiring different responses to it (i.e., having participants respond to the target but not to the prime). This latter effect supports the idea that stimulus events are integrated with the responses (or nonresponses) they accompany, but Milliken and Lupiáñez provide evidence that this integration requires the stimulus event to appear for quite a while—that is, stimulus–response integration takes time.

The last three papers address (more directly) the question of whether and how feature integration is controlled. From a Gestalt perspective, one may assume that constructing the representation of a perceptual event proceeds more or less automatically, and judging from our apparently shared phenomenal experiences it does indeed make sense to believe that people integrate more or less the same features in more or less the same way. And yet, this does not exclude that integration is tailored to the current goals of the perceiver/actor and the task he or she is busy with. Indeed, Lupiáñez, Ruz, Funes, and Milliken use a spatial cuing procedure to demonstrate that task demands can shape the integration of two events, cue and target, that appear at the same spatial location. By varying task demands between two target types that appear unpredictably following a spatial cue, their results point to the importance of retrospective cue–target integration processes, rather than prospective cue-initiated processes alone, in explaining spatial cuing effects. Wenke, Gaschler, and Nattkemper studied the impact of the stimulus–response instruction for one task on feature integration in another, logically unrelated task embedded into the first. They show that the mere instruction to apply a particular stimulus–response translation in one task changes the way stimuli and responses are integrated, and the way their codes interact, in the other task. This suggests that instructing someone (including oneself) to carry out particular responses to particular stimuli induces corresponding stimulus–response bindings that are maintained during intervening tasks. Thus, a task set can induce bindings. But the opposite is also true, as demonstrated in the final paper by Mayr and Bryck. They had participants switch between two tasks, which were either carried out on different objects (in different locations) or on the same object. The results show that switching is much easier in the first case, suggesting that objects and/or locations can become integrated with the tasks they "afford". If so, the retrieval of the control structures needed to organize stimulus and response processing can be left to the external events they operated on in the past—a clever way to outsource action control, as Mayr and Bryck put it.

Taken together, although the contributions to this special issue address a wide array of questions, they certainly do not cover the whole ground of what one may consider problems of feature integration. Yet we do think that the issues addressed in the contributions have widespread implications, and that they constitute a strong argument for the richness of the feature integration domain. Our hope is that others will agree, and will be motivated to answer some of the questions about integration in and across perception and action that the present studies have left unanswered.

In closing, we would like to thank Richard Abrams, Harold Bekkering, Matt Crump, Dirk Kerzel, Klaus Kessler, Günther Knoblich, Wilfried Kunde, Wido La Heij, Jason Leboe, Jochen Müsseler, Mike Wendt, Peter Wühr, and four anonymous reviewers for providing comments on earlier versions of the papers in this special issue.

References

Allport, D. A., Tipper, S. P., & Chmiel, N. R. J. (1985). Perceptual integration and postcategorical filtering. In M. I. Posner, & O. S. M. Marin (Eds), *Attention and performance XI* (pp. 107–132). Hillsdale: Erlbaum Associates Inc.

Calvert, G., Spence, C., & Stein, B. E. (2004). (Eds.), *Handbook of multisensory processes*. Cambridge: MIT Press.

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences, 11*, 219–226.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*, 494–500.

Robertson, L. C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience, 4*, 93–102.

Schreiner, C. E. (1995). Order and disorder in auditory cortical maps. *Current Opinion in Neurobiology, 5*, 489–496.

Stoet, G., & Hommel, B. (1999). Action planning and the temporal binding of response codes. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1625–1640.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*, 171–178.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Van der Heijden, A. H. C. (1992). *Selective attention in vision*. London: Routledge.

Zeki, E. (1993). *A vision of the brain*. New York: Oxford University Press.