

The Role of the Dorsal Visual Processing Stream in Tool Identification

Psychological Science
XX(X) 1–7
© The Author(s) 2010
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797610371343
http://pss.sagepub.com


Jorge Almeida^{1,2}, Bradford Z. Mahon^{3,4}, and Alfonso Caramazza^{1,4}

¹Cognitive Neuropsychology Laboratory and ²Harvard University Vision Sciences Laboratory, Harvard University;

³Department of Brain & Cognitive Sciences, University of Rochester; and ⁴Center for Mind/Brain Sciences (CIMeC), University of Trento

Abstract

The dorsal visual processing stream subserves object-directed action, whereas the ventral visual processing stream subserves visual object recognition. Little is known about how information computed by dorsal-stream structures influences object recognition. We used continuous flash suppression to functionally separate information computed by the dorsal stream from that computed by the ventral stream. We show that information originating from the dorsal stream influences not only decisions requiring the selection of superordinate category labels, but also decisions that entail the selection of a basic-level object. We further show that information computed by the dorsal stream does not carry specific functional information about objects. Our results indicate that the dorsal stream, in isolation from the ventral stream, is agnostic as to the identity of the objects that it processes. We suggest that structures within the dorsal visual processing stream compute motor-relevant information (e.g., graspability), which influences the identification of manipulable objects, and is not either about the function of the object or function-specific.

Keywords

dorsal processing stream, graspability, object recognition, semantic memory, tools, manipulable objects

Received 8/27/09; Revision accepted 11/8/09

Information about an object's shape, its surface material, and how it can be grasped, among other information, becomes available soon after that object engages the visual system. Convergent evidence from neuroimaging and neuropsychology studies indicates that different types of information are differentially relevant to object recognition. For instance, brain-damaged patients may show deficits in their ability to reach, grasp or manipulate objects but retain the ability to identify these objects (e.g., Buxbaum & Saffran, 2002; Buxbaum, Sirigu, Schwartz, & Klatzky, 2003; Hodges, Spatt, & Patterson, 1999; Jeannerod, Decety, & Michel, 1994; Ochipa, Rothi, & Heilman, 1989). In contrast, patients with severe object recognition deficits may exhibit spared object manipulation (e.g., Buxbaum, Schwartz, & Carew, 1997; Goodale & Milner, 1992; Hodges, Bozeat, Lambon-Ralph, Patterson, & Spatt, 2000; Negri, Lunardelli, Reverberi, Gigli, & Rumiati, 2007). Moreover, the processing of these different types of information has been traced to independent anatomical pathways: Visuomotor knowledge is extracted via dorsal-stream structures, whereas information about object identity necessary for recognition is extracted via ventral-stream structures (e.g., Goodale & Milner, 1992; Johnson-Frey, 2004; Shmuelof & Zohary, 2005).

Little is known, however, about whether, and if so how, information processed by the dorsal stream informs object recognition processes, which are presumably mediated by ventral-stream structures. In a recent study (Almeida, Mahon, Nakayama, & Caramazza, 2008), we showed that dorsal-stream outputs are relevant to the process of semantic categorization. In that study, we used an interocular suppression technique—continuous flash suppression (CFS; Tsuchiya & Koch, 2005)—to visually suppress prime pictures. Stimuli presented under CFS do not reach structures in the ventral stream, but are processed by the dorsal stream (e.g., Fang & He, 2005; Sheinberg & Logothetis, 1997; Tong, Nakayama, Vaughan, & Kanwisher, 1998). Thus, if semantic priming effects can be observed for CFS-suppressed primes, then these effects must be mediated by dorsal-stream computations. We found that CFS-suppressed categorically congruent prime pictures facilitated subsequent categorization responses of manipulable objects, but not of animals or vehicles.

Corresponding Author:

Jorge Almeida, 33 Kirkland St., Room 940, Harvard University, Cambridge, MA 02138

E-mail: jalmeida@wjh.harvard.edu

The critical open issue raised by our study (Almeida et al., 2008) concerns the nature of the information originating from the dorsal stream that is able to influence object recognition. Our 2008 study shows that this information can influence superordinate categorization, but does not address whether it can influence narrower categorizations (e.g., categorizing an object as a hammer). In addition, previous findings are not informative about the specificity of the information extracted by dorsal-stream structures.

Basic-level picture naming offers a way to study the processing level at which information computed by dorsal-stream structures influences object recognition. Basic-level naming requires precise information to isolate a particular target from other similar within-category alternatives. If information originating from the dorsal stream is relevant in the selection of the particular to-be-named target, we would expect CFS-suppressed tool primes, when compared to animal primes, to facilitate basic-level naming of tool targets. This categorical priming effect should be specific for tool targets, as suggested in our (Almeida et al.) 2008 study.

The use of identical primes and targets (i.e., that are the same exact picture) can shed light on the specificity of information extracted by dorsal-stream structures. If such information is specific to the presented object, then the more information the prime and target share, the stronger the effect should be. Primes that are identical to the subsequent targets should therefore give rise to an identity effect that exceeds categorical priming.

In this article, we report a study showing that basic-level naming latencies for pictures of manipulable objects (but not animals) are influenced by information computed by dorsal-stream structures. We also show that under CFS, identity conditions (i.e., when prime and target are identical) do not lead to a priming effect that exceeds categorical priming. In contrast, when we used a masking technique that allows information to reach the ventral stream (backward masking; e.g., Dehaene et al., 2001), we obtained an identity effect that surpassed categorical priming. These results show that information originating from the dorsal stream can influence performance on tasks that require different degrees of processing—from determining the target's unique name to extracting its superordinate category. Our results also suggest that these effects are not dependent on detailed information about the presented object, but rather are based on the extraction of relatively coarse motor-relevant information.

Experiment 1

In Experiment 1, we used CFS-based priming in a picture-naming task to determine whether information originating from the dorsal stream produces a priming effect on picture naming at the basic level. Essentially, we would expect to see a priming effect under CFS if dorsal stream information can be used to affect basic level decisions. We also sought to replicate the previous finding that dorsal-stream information produces a priming effect in a categorization task.

To ensure reliability of the results, we conducted two separate experiments (1a and 1b). We enlisted different participants and used different sets of pictures for each experiment. We also varied the timing of the naming and categorization sessions (from no delay between sessions to a delay of a week).

Method

Participants. Twenty Harvard University undergraduates participated in each experiment) in exchange for course credit or a payment of \$7. All participants were right-handed, had normal or corrected-to-normal vision, and gave written informed consent. The project was approved by the appropriate institutional review board.

Materials and procedure. We used black-and-white pictures of animals and tools (described in Almeida et al., 2008). For each category, and separately for each experiment, we selected four pictures as targets and four different pictures as primes. In this way, the primes were categorically congruent with one group of targets and categorically incongruent with the other (e.g., animal primes were categorically congruent with animal target pictures and categorically incongruent with tool target pictures). The primes were rendered invisible via CFS (see Fig. 1a), a technique in which a static image competes with a dynamic image, with the latter reliably suppressing the former for a prolonged period of time. To induce CFS, on each trial we presented a low-luminance, low-contrast version of the prime stimulus to the participant's nondominant eye and a dynamic, high-contrast random-noise pattern that changed every 100 ms to the dominant eye. Red/green anaglyph glasses were worn by the participants to allow for dioptic presentation of the images. The stimuli were presented centrally and subtended 7° of visual angle.

The procedures we used were the same as those we followed in our earlier study (Almeida et al., 2008) except that the experiments proceeded in three independent stages: categorization, naming, and prime discrimination. The order of the categorization and naming tasks was counterbalanced across participants, and the prime-discrimination task was always performed at the end of the experiment. In the naming and categorization stages, participants saw a fixation cross (for 500 ms), the prime and the first random-noise pattern (for 100 ms), the prime and a second random-noise-pattern (100 ms), and finally the target picture (for 3 s or until the participant responded, whichever came first; see Fig. 1a). Depending on the task, participants were asked to either name the pictures or categorize them as "animals" or "tools" (button-press response). Experiments 1a and 1b differed in the time interval between naming and categorization: In Experiment 1a, the two tasks were separated by a week, whereas in Experiment 1b, one task immediately followed the other.

The prime-discrimination task provided independent data so we could assess awareness of the prime. In this task, participants were asked to place the prime into one of the two target

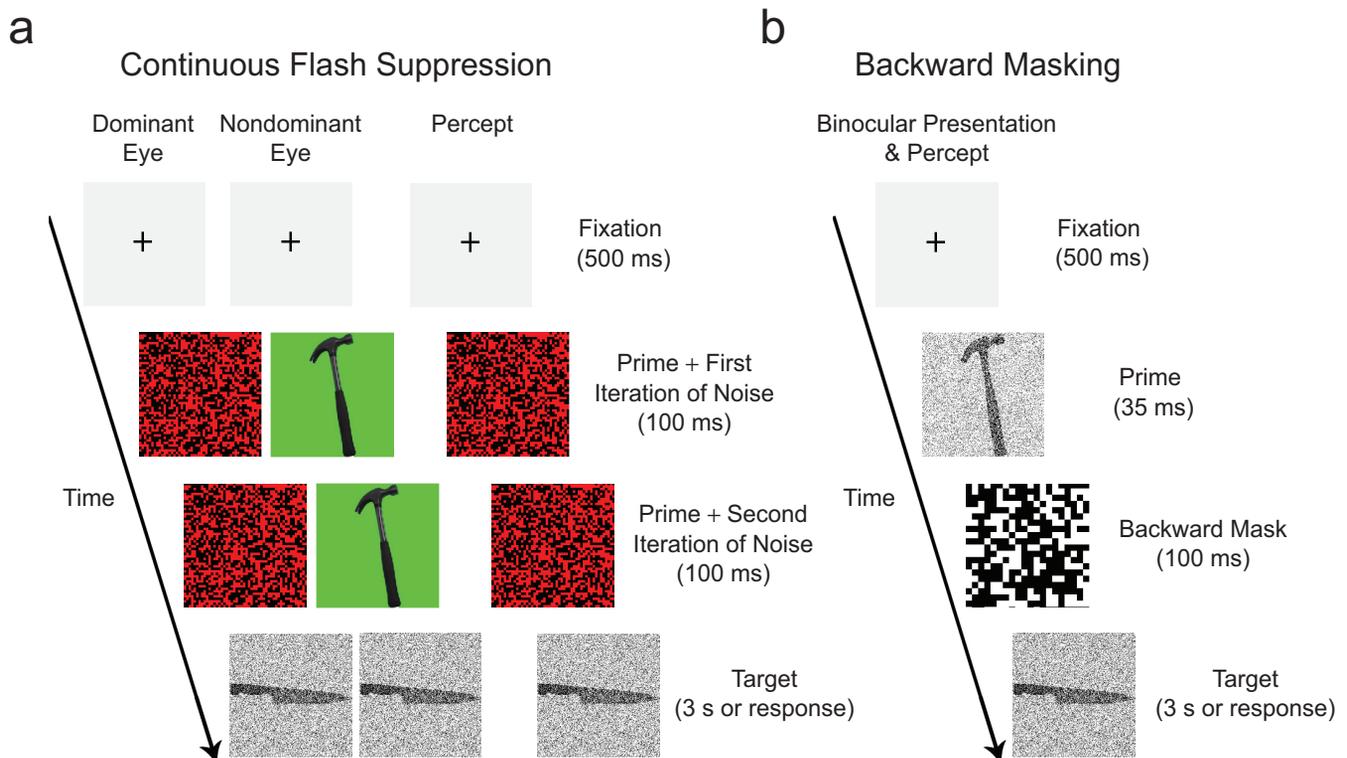


Fig. 1. Procedure used in (a) Experiments 1a, 1b, and 2a, which used continuous flash suppression to suppress the primes (the high luminance of the images is for visualization purposes), and (b) Experiment 2b, which used backward masking to suppress the primes. In Experiments 1a, 1b, and 2a, each trial included the following on-screen sequence: a fixation cross (500 ms), the prime picture accompanied by the first random-noise pattern (100 ms), the prime picture accompanied by the second random-noise pattern (100 ms), and the target picture (3 s or until response). In Experiment 2b, a fixation cross appeared on the screen for 500 ms, followed immediately by the prime picture (35 ms), the backward mask (100 ms), and the target picture (3 s or until response).

categories. The trial sequence remained the same as in the previous tasks except that the target was not presented. In Experiment 1a, participants performed the prime-discrimination task after the naming and categorization condition independently, as these conditions were separated by a week. Only participants who performed at chance level on the prime-discrimination task were included in the main analyses reported here (see Figs. S1a, S1b, and S1c in the Supplemental Material available online for more detailed analyses). Stimuli were presented using DisplayMaster DirectX (DMDX) software (Forster & Forster, 2003).

Analyses. We used planned contrasts to analyze response latencies (Rosenthal, Rosnow, & Rubin, 2000). For each category in each experiment, two pair-wise contrasts were employed to test the priming effect in the categorization and naming responses. The priming effect was calculated as the difference in reaction time between incongruent trials and congruent trials.

Results

In both experiments, significant priming effects were obtained for tool targets in both tasks (see Fig. 2a). Participants were

faster to categorize tools in the context of tool primes rather than animal primes. The mean priming effect was 11 ms ($SEM = 3$) in Experiment 1a, $F(1, 19) = 10.6$, $MSE = 227.8$, $p = .004$, $d = 0.73$, and 9 ms ($SEM = 4$) in Experiment 1b, $F(1, 19) = 5.7$, $MSE = 265.5$, $p = .028$, $d = 0.53$. Participants were also faster to name tools in the context of tool primes rather than animal primes. The mean priming effect was 18 ms ($SEM = 7$) in Experiment 1a, $F(1, 19) = 6.9$, $MSE = 923.7$, $p = .017$, $d = 0.59$, and 12 ms ($SEM = 6$) in Experiment 1b, $F(1, 19) = 4.8$, $MSE = 607.2$, $p = .04$, $d = 0.49$. The contrasts for animal targets did not reach significance in either task, all F s < 1 except in the case of the priming effect for naming in Experiment 1a, $F(1, 19) = 2.5$, $MSE = 876.2$, $p = .13$.

Discussion

In Experiment 1, we replicated the previous finding of category-specific priming under CFS in a categorization task (Almeida et al., 2008), and further showed that the effect can also be observed in basic-level naming. Participants were faster to categorize or name a target picture in the context of categorically congruent CFS-suppressed primes than in the context of incongruent primes. As expected, these results were obtained for tool targets only (and not for animal targets).

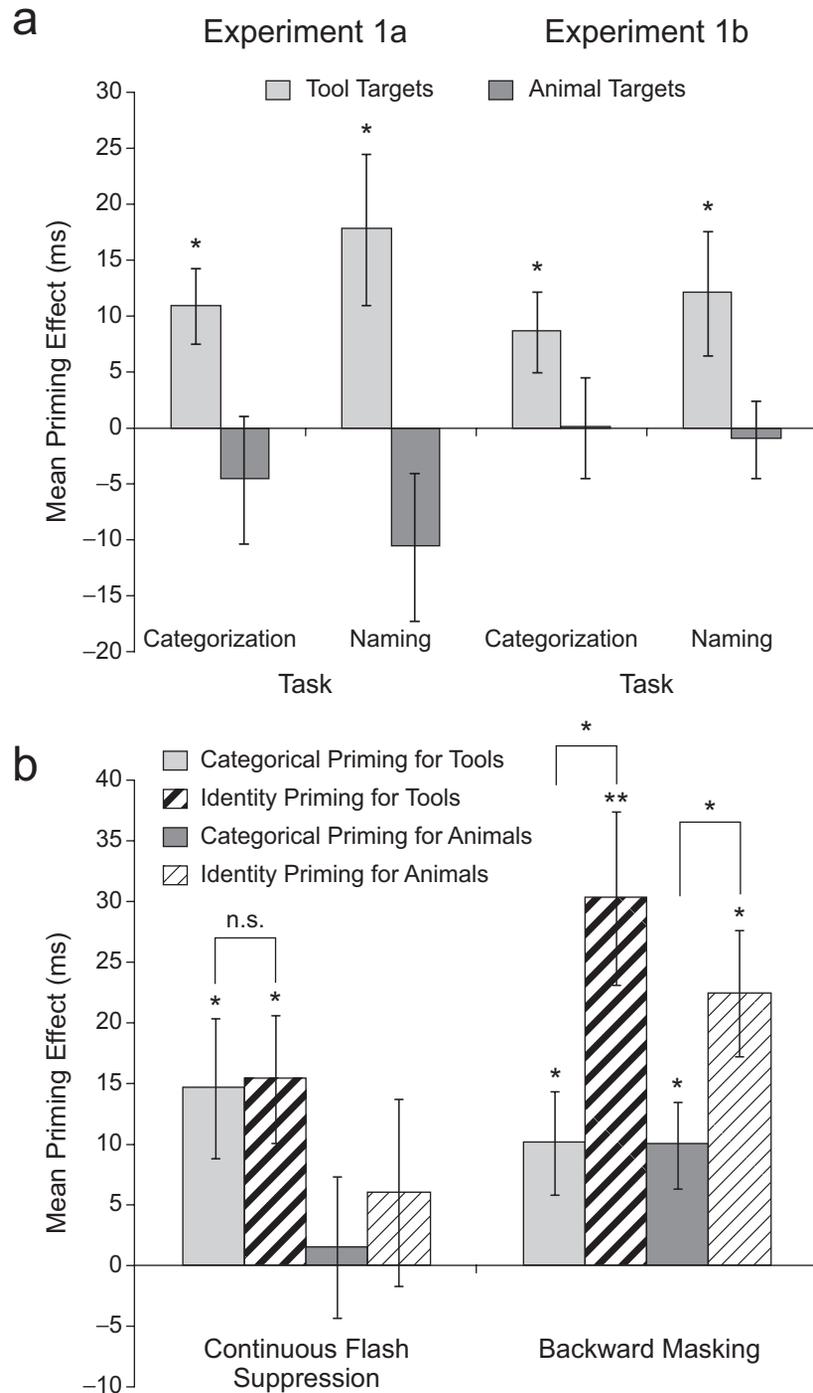


Fig. 2. Behavioral priming effects in (a) Experiments 1a and 1b and (b) Experiments 2a and 2b. In (a), the mean priming effect is plotted as a function of task, separately for tool and animal targets. In (b), mean categorical and identity priming effects for tool and animal targets are plotted as a function of experimental procedure. Error bars represent standard errors of the mean across subjects, and asterisks indicate significance (* $p < .05$, ** $p < .001$).

Experiment 2

In Experiment 2, we addressed the specificity of information originating from the dorsal stream by including an identity condition (in addition to the categorically congruent and incongruent conditions). If information originating from the

dorsal stream is specific to the prime picture, identity priming should surpass categorical priming for CFS-suppressed primes.

We used two techniques to mask prime pictures: CFS in Experiment 2a and backward masking in Experiment 2b. The latter technique allows information to reach the ventral stream

(Dehaene et al., 2001), and so we expected that the identity effect would exceed the categorical priming effect in Experiment 2b.

Method

Participants. Twenty-four undergraduates, who did not participate in Experiment 1, participated in Experiments 2a and 2b (12 per experiment).

Materials and procedure. We used the same pictures that were used in Experiment 1a, and we also used the target pictures as primes for the identity condition. Participants were asked to categorize the target pictures and indicate the selected category by pressing response buttons. We used the same procedures in Experiment 2a as in Experiment 1. In Experiment 2b, the prime picture (35 ms) was followed by a black-and-white random-noise mask (100 ms; see Fig. 1b). We added 70% noise to the prime stimuli (using Photoshop) to facilitate masking. The subsequent target presentation followed the procedures used in Experiment 1.

Analyses. We used planned contrasts to analyze response latencies. For each target category, we conducted three pair-wise contrasts to test for categorical and identity priming and to analyze whether identity priming exceeded categorical priming.

Results

In Experiment 2a, where we used CFS to render the primes invisible, participants were faster to categorize tools in the context of tool primes rather than animal primes (mean priming effect = 15 ms, $SEM = 6$), $F(1, 11) = 6.5$, $MSE = 397.5$, $p = .027$, $d = 0.74$. They were also faster to categorize tools in the context of identical tool primes rather than animal primes (mean identity effect = 15 ms, $SEM = 5$), $F(1, 11) = 8.6$, $MSE = 328.0$, $p = .014$, $d = 0.84$ (see Fig. 2b). However, response latencies did not differ between identical and nonidentical tool primes (mean identity effect above the categorical effect = 1 ms, $SEM = 4$), $F(1, 11) < 1$ (see Fig. 2b). None of the corresponding contrasts for animal targets reached significance (all $F_s < 1$).

In Experiment 2b (backward masking), categorical priming for tool targets was significant (mean priming effect = 10 ms, $SEM = 4$), $F(1, 11) = 5.7$, $MSE = 217.6$, $p = .037$, $d = 0.68$. The identity effect was also significant (mean identity effect = 30 ms, $SEM = 7$), $F(1, 11) = 18$, $MSE = 609.4$, $p = .001$, $d = 1.21$. Moreover, identity primes led to a larger effect than categorically congruent primes (mean identity effect above the categorical effect = 20 ms, $SEM = 7$), $F(1, 11) = 8.2$, $MSE = 589.4$, $p = .015$, $d = 0.82$ (see Fig. 2b). Contrasts for animal targets yielded significant category priming (mean priming effect = 10 ms, $SEM = 4$), $F(1, 11) = 7.6$, $MSE = 154.1$, $p = .019$, $d = 0.81$, as well as significant identity priming (mean identity effect = 22 ms, $SEM = 5$), $F(1, 11) = 18.9$, $MSE = 319.0$, $p =$

$.001$, $d = 1.23$. In addition, identity priming was greater than categorical priming (mean identity effect above the categorical effect = 20 ms, $SEM = 5$), $F(1, 11) = 6.2$, $MSE = 303.4$, $p = .03$, $d = 0.69$ (see Fig. 2b; see Figs. S2a and S2b in the Supplemental Material available online for prime awareness analyses).

Discussion

The identity effect in Experiment 2a did not exceed category-congruent priming. This indicates that the information originating from the dorsal stream, despite being able to influence broad (target categorization) and narrow (target naming) information processing, is not specific to the presented object. An enhanced identity effect was, however, observed in a backward-masking paradigm (Experiment 2b). Thus, when information reaches ventral-stream structures, primes that are identical to targets lead to the expected enhanced identity effect.

General Discussion

There are many unanswered questions about the processes underlying the cognitive task of recognizing objects. A central topic in visual object recognition research concerns the type of information used in object recognition. We recently suggested that information originating from the dorsal stream can influence object recognition (Almeida et al., 2008; see also Helbig, Graf, & Kiefer, 2006; and Mahon et al., 2007; for a similar result in the context of number processing, presumably also performed by dorsal-stream structures, see Bahrami et al., 2010). Characterizing this information and how it is used is critical in understanding the role of the dorsal stream in object recognition, as well as the interactions between the ventral and dorsal visual streams.

We found that dorsal-stream information influences object recognition over broad categorical processing such as superordinate categorization (i.e., categorization of targets on broad categories like animals vs. tools), as well as over narrow categorical processing such as basic-level naming. Moreover, our results show that this information is very coarse: When we used CFS to render primes invisible, response latencies did not differ between primes that were identical to the targets and primes that were only categorically congruent.

Our results show that the information computed by dorsal-stream structures can be used in recognizing manipulable objects, but not because it specifically identifies the functional properties of such objects. In fact, in isolation, the dorsal stream appears to be agnostic as to the identity of objects. Our findings suggest that dorsal-stream information is not about the object per se, but may be about more rudimentary motor-relevant knowledge, presumably knowledge related to whether the object is graspable and how it might be grasped, in a strict visuomotor sense.

This conclusion is supported by evidence from neuropsychological and neuroimaging studies suggesting that the

processes related to acting on an object (e.g., preparing a grasp for moving a hammer) and using an object (e.g., preparing a grasp for using a hammer to pound a nail) are dissociable (e.g., Johnson & Grafton, 2003; Rizzolatti & Matelli, 2003). For instance, optic-ataxic patients show deficits in reaching and grasping objects but may be able to manipulate familiar objects (Jeannerod et al., 1994). Conversely, apraxic patients are impaired in object use but may perform optimal grasps toward objects (e.g., Buxbaum et al., 2003; Heilman & Rothi, 1997). This latter performance pattern is presumably due to the processing of intrinsic physical characteristics of target objects (e.g., shape) in such a way as to facilitate hand/object interactions that is independent of the recognition of the target object or its canonical manipulation.

Whereas the processes involved in grasping an object depend on visuomotor transformations over the object's intrinsic physical properties, using an object requires the integration of stored knowledge about the object's function, its identity, and the motor programs associated with it. Studies suggest that dorsal-stream regions, such as the inferior parietal lobule, are fundamental for these processes (e.g., Buxbaum et al., 2003; Johnson & Grafton, 2003; Rizzolatti & Matelli, 2003). There is also evidence, however, that ventral-stream structures are important in object use. Patients with lesions that are restricted to ventral-stream structures may grasp objects in a way that is consistent with their physical structure, but that is not appropriate for subsequent use of the objects (Carey, Harvey, & Milner, 1996). The observation of grasps typically associated with a particular object use (e.g., grasping a hammer by its handle such that the head of the hammer is an extension of the arm), when compared with grasps that are possible but not typical (e.g., grasping a hammer by its handle such that the opposite tip of the hammer's head is an extension of the arm), leads to activations in ventral temporal regions (Valyear & Culham, 2010). Our results, taken together with these clinical and neuroimaging observations, indicate that visual dorsal-stream information can be interpreted conceptually but, on its own, is not conceptually defined.

Acknowledgments

We thank Veronica Zapater and Lukas Strnad for their help in data collection.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

A.C. was supported by National Institute on Deafness and Other Communication Disorders Grant R01-DC006842 and by Fondazione Cassa di Risparmio di Trento e Rovereto. J.A. was supported by Fundação para a Ciência e a Tecnologia, Portugal, Grant SFRH/BD/28994/2006. B.Z.M. was supported by Fondazione Cassa di Risparmio di Trento e Rovereto.

Supplemental Material

Additional supporting information may be found at <http://pss.sagepub.com/content/by/supplemental-data>

References

- Almeida, J., Mahon, B.Z., Nakayama, K., & Caramazza, A. (2008). Unconscious processing dissociates along categorical lines. *Proceedings of the National Academy of Sciences, USA*, *105*, 15214–15218.
- Bahrami, B., Vetter, P., Spolaore, E., Pagano, S., Butterworth, B., & Rees, G. (2010). Unconscious numerical priming despite interocular suppression. *Psychological Science*, *21*, 224–233.
- Buxbaum, L.J., & Saffran, E.M. (2002). Knowledge of object manipulation and object function: Dissociations in apraxic and non-apraxic subjects. *Brain and Language*, *82*, 179–199.
- Buxbaum, L.J., Schwartz, M.F., & Carew, T.G. (1997). The role of semantic memory in object use. *Cognitive Neuropsychology*, *14*, 219–254.
- Buxbaum, L.J., Sirigu, A., Schwartz, M.F., & Klatzky, R. (2003). Cognitive representations of hand posture in ideomotor apraxia. *Neuropsychologia*, *41*, 1091–1113.
- Carey, D.P., Harvey, M., & Milner, A.D. (1996). Visuomotor sensitivity for shape and orientation in a patient with visual form agnosia. *Neuropsychologia*, *34*, 329–337.
- Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.F., Poline, J.B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*, 752–758.
- Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, *8*, 1380–1385.
- Forster, K.I., & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124.
- Goodale, M.A., & Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*, 20–25.
- Heilman, K.M., & Rothi, L.J.G. (1997). Limb apraxia: A look back. In L.J.G. Rothi & K.M. Heilman (Eds.), *Apraxia: The neuropsychology of action* (pp. 7–18). Hove, England: Psychology Press.
- Helbig, H.B., Graf, M., & Kiefer, M. (2006). The role of action representations in visual object recognition. *Experimental Brain Research*, *174*, 221–228.
- Hodges, J., Bozeat, S., Lambon-Ralph, M., Patterson, K., & Spatt, J. (2000). The role of conceptual knowledge in object use: Evidence from semantic dementia. *Brain: A Journal of Neurology*, *123*, 1913–1925.
- Hodges, J., Spatt, J., & Patterson, K. (1999). “What” and “how”: Evidence for the dissociation of object knowledge and mechanical problem-solving skills in the human brain. *Proceedings of the National Academy of Sciences, USA*, *96*, 9444–9448.
- Jeannerod, M., Decety, J., & Michel, F. (1994). Impairment of grasping movements following a bilateral posterior parietal lesion. *Neuropsychologia*, *32*, 369–380.

- Johnson, S.H., & Grafton, S.T. (2003). From 'acting on' to 'acting with': The functional anatomy of object-oriented action schemata. In D.P.C. Prablanc & Y. Rossetti (Eds.), *Progress in brain research* (Vol. 142, pp. 127–139). New York: Elsevier.
- Johnson-Frey, S.H. (2004). The neural bases of complex tool use in humans. *Trends in Cognitive Sciences*, 8, 71–78.
- Mahon, B.Z., Milleville, S.C., Negri, G.A.L., Rumiati, R.I., Caramazza, A., & Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron*, 55, 507–520.
- Negri, G.A., Lunardelli, A., Reverberi, C., Gigli, G.L., & Rumiati, R.I. (2007). Degraded semantic knowledge and accurate object use. *Cortex*, 43, 376–388.
- Ochipa, C., Rothi, L.J., & Heilman, K.M. (1989). Ideational apraxia: A deficit in tool selection and use. *Annals of Neurology*, 25, 190–193.
- Rizzolatti, G., & Matelli, M. (2003). Two different streams form the dorsal visual system: Anatomy and functions. *Experimental Brain Research*, 153, 146–157.
- Rosenthal, R., Rosnow, R.L., & Rubin, D.B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach*. New York: Cambridge University Press.
- Sheinberg, D.L., & Logothetis, N.K. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences, USA*, 94, 3408–3413.
- Shmuelof, L., & Zohary, E. (2005). Dissociation between ventral and dorsal fMRI activation during object and action recognition. *Neuron*, 47, 457–470.
- Tong, F., Nakayama, K., Vaughan, J.T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, 21, 753–759.
- Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, 8, 1096–1101.
- Valyear, K.F., & Culham, J.C. (2010). Observing learned object-specific functional grasps preferentially activates the ventral stream. *Journal of Cognitive Neuroscience*, 22, 970–984.