A role for action knowledge in visual object identification

GENEVIÈVE DESMARAIS, MIKE J. DIXON, AND ERIC A. ROY
University of Waterloo, Waterloo, Ontario, Canada

We evaluated the impact of visual similarity and action similarity on visual object identification. We taught participants to associate novel objects with nonword labels and verified that in memory visually similar objects were confused more often than visually dissimilar objects. We then taught participants to associate novel actions with nonword labels and verified that similar actions were confused more often than dissimilar actions. We then paired specific objects with specific actions. Visually similar objects paired with similar actions were confused more often in memory than when these same objects were paired with dissimilar actions. Hence the actions associated with objects served to increase or decrease their separation in memory space, and influenced the ease with which these objects could be identified. These experiments ultimately demonstrated that when identifying stationary objects, the memory of how these object were used dramatically influenced the ability to identify these objects.

Which aspects of objects are important in visual identification? An object that is long, cylindrical, and tapered at one end is likely to be recognized as a "pen." Is it only the visual attributes of the pen that are involved in recognizing this object, or do certain nonvisual attributes stored within semantics (i.e., that a pen is held a certain way, and is used for writing) also play a role? We propose that in addition to the visual features of objects, other nonvisual features of these objects, specifically the actions habitually associated with these objects, play a key role in the visual identification and the naming of visually presented objects. We will briefly review the impact of visual and semantic similarity on visual identification before discussing the importance of action information and presenting the current studies.

Visual Similarity Influences Visual Identification

In healthy participants, visual identification is influenced by visual similarity: Items that come from visually dissimilar categories are identified faster than items that come from visually similar categories (Dickerson & Humphreys, 1999; Humphreys, Riddoch, & Quinlan, 1988; Lloyd-Jones & Humphreys, 1997a, 1997b). Visual similarity also impacts the identification performance of patients with visual agnosia. In these patients, only the identification of certain categories of objects is impaired, a condition labeled category-specific visual agnosia (CSVA). In typical patients with CSVA, the visual identification of items like animals, fruits, vegetables, and musical instruments is impaired while the visual identification of items from other categories, such as tools, vehicles, and body parts is spared (Damasio, 1990; Dixon, Bub, & Arguin,

1997; Forde, Francis, Riddoch, Rumiati, & Humphreys, 1997; Gainotti & Silveri, 1996; Tranel, Logan, Frank, & Damasio, 1997). Arguably, the items that are typically impaired in CSVA patients come from categories where the exemplars are more visually similar to one another than the items that are spared. Gaffan and Heywood (1993) have highlighted the importance of visual similarity by showing that even nonhuman primates have more difficulty identifying drawings of animals than drawings of man-made objects, presumably because the shapes of animals are more visually similar than the shapes of man-made objects.

In contrast to animals, objects such as tools share few visual attributes, and patients with CSVA show relatively few problems identifying them. However, for humans who know about animals and about tools, visual similarity may not be the only factor influencing visual object identification; semantic similarity may also play a role. Consider a patient who confuses depictions of lions and tigers but not depictions of hammers and saws. The felines are more visually similar than the tools, but also share more semantic features. Using depictions of real objects (e.g., comparing identification times of different felines to those of different articles of tools) it is difficult to disentangle the effects of visual similarity on identification from the effects of nonvisual similarity; for many categories of real-world objects visually similar objects are more semantically related than dissimilar ones. In order to circumvent this problem some researchers have favored the use of novel objects.

Desmarais and Dixon (2005) trained healthy undergraduates to identify eight novel shapes (see Figure 1) using nonword labels (e.g., *baiv*, *jorl*, *yoot*). When participants

G. Desmarais, gdesmara@ahsmail.uwaterloo.ca

1712
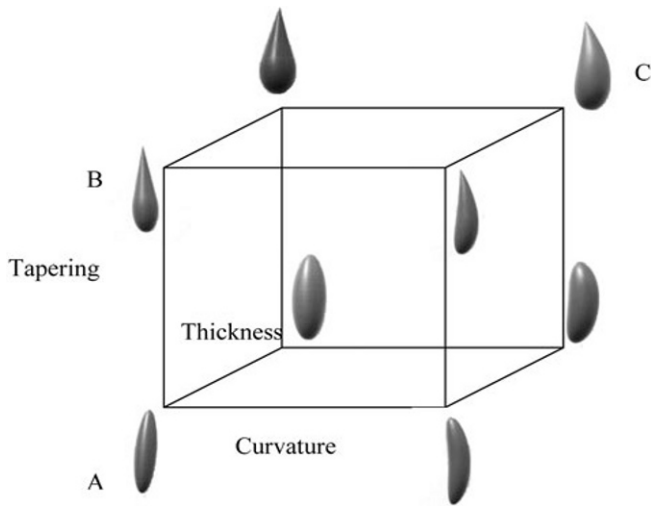
Figure 1. 3-D object space. Desmarais and Dixon (2005) used two-dimensional representations of these objects.

made identification errors, Desmarais and Dixon noted which shapes were confused with which other shapes, and observed that participants made more confusion errors between shapes that were visually similar (for example between shapes A and B) than between shapes that were visually dissimilar (for example between shapes B and C).

This finding demonstrates that with novel objects, visually similarity influences identification in healthy observers. These novel objects were also used to demonstrate that semantic similarity could influence object identification.

Semantic Similarity Influences Visual Object Identification

Dixon et al. (1997) explored the role played by nonvisual semantic information on visual object identification using a simpler variant of this shape-labeling paradigm. Dixon et al. tested the CSVA patient ELM using quadruplets of novel shapes that were paired with the sounds of objects. On test trials, ELM was presented with each shape alone, and asked to recall which shape went with which sound. Semantic similarity was varied by pairing shapes with sounds made by objects from the same category or pairing shapes with sounds made by objects from different categories. By using the same set of shapes to stand for either similar or dissimilar concepts, the authors could hold visual form constant and assess the impact of semantic similarity. They found that ELM's identification abilities were markedly influenced by the semantic proximity of the concepts he attempted to pair with the shapes; if the concepts were unrelated, he performed well, but if the concepts were similar his performance suffered dramatically. Similar patterns of results were also shown in patients with Alzheimer's disease (Dixon, Bub, Chertkow, & Arguin, 1999).

Dixon, Desmarais, Gojmerac, Schweizer, and Bub (2002), showed that object identification depended not only on what objects looked like, but also what patients know about these objects. Part of what one knows about an object often involves how an object is used. De Renzi and Luchelli (1994) posited that it was the *functions* of objects like tools that differentiated them from objects like animals. They proposed that objects like hammers were designed with a specific function in mind ("used to hit nails"), while objects like lions do not have a specific function. Instead animals are associated with nonfunctional attributes ("is a hunter"). According to De Renzi and Luchelli (1994), seeing the form of a hammer would directly activate its function and functional knowledge would then aid in the object's identification. A number of researchers have posited that it is this functional knowledge that serves to protect items like tools against identification impairments and serves as the basis for the most common form of category-specific deficits (e.g., Farah & McClelland, 1991; Funnel & Sheridan, 1992; Temple, 1986).

A recent study by Gauthier, James, Curby, and Tarr (2003) provides converging evidence that nonvisual information can influence visual object identification not only in patients but also in healthy participants. Using methods similar to Dixon and colleagues, these authors imbued novel objects with meaning through training. Healthy participants were trained to associate *yufos* (novel, creature-like objects) with specific nonvisual attributes that were either similar or distinct. The authors demonstrated, using a perceptual matching task, that when *yufos* possessed, via training, *dissimilar* nonvisual attributes participants found them to be more *visually* discriminable. Taken together these experiments suggest that nonvisual attributes can influence visual object identification even in a task that can be performed purely by making same/different judgments based on structural features of the objects alone.

Action Knowledge

The studies of Dixon et al. (1997, 2002) and Gauthier et al. (2003), all suggest that nonvisual knowledge can influence visual identification. An important type of nonvisual knowledge concerns "action attributes", or in other words knowledge of how one typically interacts with an object. Could actions help make objects semantically dissimilar, and hence influence visual object identification? James and Gauthier (2003) imbued creature-like objects (greebles) with action attributes or auditory attributes using a training regimen similar to the ELM paradigm. After an initial training phase, participants were scanned using fMRI while performing a perceptual matching task with the greebles from either the action or auditory quadruplets. When participants completed the matching task with the "action" greebles, scans revealed activity in the posterior temporal sulcus—a brain area normally recruited when processing specific actions. In contrast, when the matching task involved "auditory" greebles, the scans revealed activity in the superior temporal gyrus—a brain area normally recruited when processing sounds. Because the greebles that were imbued with action or auditory attributes had highly similar visual structures, these different brain activation patterns must have been attributable to the action or auditory attributes conveyed by the labels. As such, this study provides strong evidence that nonvisual action representations do indeed play a role in object processing, even

in a task that could be completed purely by evaluating the subtle visual differences between the presented stimuli.

Converging evidence that actions may play a role in visual identification comes from the literature on CSVA. Recall that in the typical pattern of CSVA, tool identification is spared relative to other categories of objects that habitually do not have any specific actions associated with them (e.g., animals). Thus it is tempting to propose that the actions associated with tools contribute to their preservation in CSVA. However, it does not suggest *how* action knowledge, which is nonvisual and not readily available to the senses during static object identification, could serve to visually identify objects.

Humphreys and Forde (2001) proposed a cascade processing framework for understanding visual object recognition that can explain how action knowledge and other nonvisual object information could influence visual object identification. This hierarchical interactive theory (HIT) contains three types of stored knowledge: (1) visual descriptions, (2) stored semantic knowledge that contains action knowledge, and (3) name representations. When an object is visually presented, the structural description corresponding to the presented object would be most activated, but there would also be some activation of structural descriptions corresponding to objects that are similar in shape to the presented object. Before identification has been achieved, activation is fed forward to the semantic level, where there would be competition between the activation patterns associated with the stored semantic knowledge of the portrayed object (which would receive the most activation), and activation patterns associated with other entities that are semantically similar to the portrayed object (which would receive less activation). Semantic level activation then feeds back to the structural description level. This feedback is crucial, for the activation patterns associated with the stored semantic knowledge of the portrayed object would reinforce the activation of the correct structural description, and drive down the activation levels of competing structural descriptions. In the hierarchical interactive theory, therefore, similarity at the structural description level *and* at the semantic level plays a key role in visual identification. More importantly, this model predicts that nonvisual information can influence visual identification.

Humphreys and Forde's (2001) model groups all types of nonvisual information together under the heading of semantic knowledge. However, it has often been suggested that actions possess a privileged relationship with the objects they are associated to (Chainay & Humphreys, 2002a, 2002b; Rumiati & Humphreys, 1998). In a series of experiments where they asked neurologically healthy participants to make either an action decision (pour vs. twist?) or a semantic decision (found in the kitchen?) to objects or words (names of objects), Chainay and Humphreys (2002b) reported that participants were consistently faster and more accurate when asked to make an action decision to objects compared to when they were asked to make a semantic decision. They also reported that action decisions were faster and more accurate when made in response to an object than to a word. These findings led

the authors to conclude that objects had privileged access to action knowledge compared to other forms of semantic knowledge, and that this advantage was contingent on learned associations between objects and actions.

Yoon, Heinke, and Humphreys (2002) proposed a connectionist model of action selection and name selection, the naming and action model (NAM), which builds on the HIT model and explains how objects could have preferential access to actions via a direct route to action in addition to an indirect route via semantics. According to this model, and similar to the HIT model, the visual presentation of an object will activate the structural description system, first through input units that encode the visual attributes of the object presented, then through a network that captures the structural similarity between objects. The output of this network then feeds into the semantic system, which contains superordinate knowledge as well as item-specific knowledge. The semantic system then feeds into both a phonological name output and an action output. However, the output of the objects' network also feeds directly to action outputs (bypassing the semantic system), indicating that when objects are presented, actions can be selected without accessing semantics but that names cannot be selected without accessing semantics. When objects are presented to the model, the authors demonstrated that action selection was reached faster than name selection, indicating that action selection was completed before naming would occur.

The authors then lesioned the model by adding noise to the route mapping visual input descriptions of objects into the semantic system—a lesion that left the direct route from objects to actions intact (as in optic aphasia). The lesion successfully created naming impairments: The lesion model took longer to select a name, and produced more errors than the unlesioned model. Furthermore, this lesion also affected action selection by increasing reaction time and errors, albeit not to the extent of the effect observed for name selection. However, and more importantly, the model demonstrated how the correct action selection could facilitate name selection: Activity from the correctly selected action fed backward to the semantic system, helping the correct semantic representation, and consequently the correct name, reach threshold.

In sum, evidence from patients and healthy observers implicates that action knowledge may influence the manner in which objects are visually identified, and the HIT and NAM models provide an explanation as to how action knowledge can influence visual object identification. However, with confrontation naming of either line drawings or real objects, one never knows whether tools are simply better identified because they have more distinct shapes than other categories of objects. In studies of healthy persons, where visual properties of objects has been more adequately controlled for, the objects have been imbued with action properties using verbal labels—a situation that is somewhat removed from the "watch and learn" scenarios of tool use learning in real life. Furthermore, in the James and Gauthier (2003) study, the types of actions referenced by these labels were primarily movements associated with creatures (e.g., crawls, walks, hops, bites, burrows, etc.). Arguably, what James and Gauthier
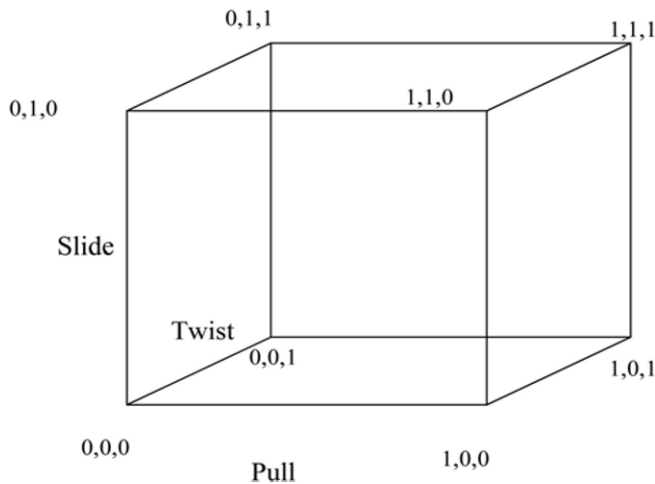
Figure 2. 3-D action space.

call "action semantic features" are often more akin to biological motion features, and may therefore play a different role in object identification than the action features associated with the knowledge of how we use real tools.

The Present Experiments

In order to effectively demonstrate that action knowledge can impact visual object identification, our approach was to create a visual space and an action space in which the visual attributes are known, and the action attributes ascribed to the shapes are also known, and under control of the experimenter. The visual shape space can be formed by creating shapes that possess different combinations of curvature, tapering and thickness. In order to specifically demonstrate the effect of action knowledge on visual object identification, we created a (de novo) semantic space consisting solely of action attributes. Importantly, participants could learn these actions by watching the experimenter interact with objects in realistic ways.

Manipulating and controlling visual and nonvisual (action) similarity. We manipulated similarity in this study as follows.

Controlling visual similarity. Desmarais and Dixon (2005) have developed a set of psychophysically scaled shapes. This set consists of eight novel shapes formed using various combinations of curvature, tapering and thickness (see Figure 1). The visual dimensions of these shapes form a 3-D space where the position of each shape can be described by a set of Cartesian coordinates. In this space the shape in position 0,0,0 is called the "origin shape." The values refer to the fact that it is not curved, is not tapered and has (for ease of labeling) minimal thickness. All other shapes involve perturbations of this origin shape. In this shape space, some objects are close together and some are further apart. The distance between two shapes can be construed as city-block (CB) distance, where the distance between two shapes differing on a given number of dimensions is defined as the sum of the distances obtained separately on each of the dimensions (Shepard, 1987). Note that the number of distinguishing attributes equals

the number of CBs that separate two shapes. Shapes 0,0,0 and 0,0,1 are one CB apart, whereas shapes 0,0,0 and 1,0,1 are two CBs apart, and shapes 0,0,0 and 1,1,1 are three CBs apart. Desmarais and Dixon (2005) psychophysically scaled these shapes to ensure that changes on each of the three dimensions were equally salient. By using this shape set, one knows, and can therefore control the visual similarity relations among the exemplars that make up this collection of objects.

Controlling action similarity. An action space similar to the visual space described above can be created using the action dimensions of "pulling," "sliding," and "twisting" (see Figure 2).

Here the origin action (0,0,0) that anchors the action space would essentially be simply grasping an object such as a cylinder. The other points in the space would involve adding the different actions to the grasp. For example the 1,0,0 action would involve a pulling but not sliding and twisting, whereas the 1,1,0 action would involve pulling and sliding the object. Like the visual shape space, this action space should be governed by the similarity between the actions—actions that are closer together in action space have fewer distinguishing attributes than actions that are further apart. As in the shape space previously described, the number of distinguishing attributes equals the number of CBs that separate two actions. Hence, as in the shape set, in this set of actions, one knows, and can therefore control the similarity relations among the actions that comprise this multidimensional action space.

Demonstrating that action knowledge can influence visual identification will be achieved by combining the multidimensional visual shape space described in Figure 1 with the action space described in Figure 2. This goal was achieved in three steps. First, in Experiment 1, we showed that visual similarity influenced memory confusions for 3-D objects. Second we showed that the visual similarity of actions influenced memory confusions among these actions. Finally, in Experiments 2 and 3 we showed that visually similar objects could be rendered more or less prone to visual identification errors by imbuing these objects with either similar actions or dissimilar actions. That is, in a completely controlled object space where the visual properties of the objects were known, and the action properties of these objects were also known, we were able to show that the visual identification of these objects depended on the knowledge of "what these objects do."

EXPERIMENT 1

The goal of Experiment 1 was to look individually at action space and 3-D object space to demonstrate that the similarity relationship demonstrated by Desmarais and Dixon (2005) also exists for the visual identification of 3-D objects and actions. Participants learned to identify novel objects and novel actions with pronounceable nonwords. If visual similarity drives visual object identification, similar objects will be confused more often than dissimilar objects. Also, if action similarity drives action identification, similar actions will be confused more often than dissimilar actions.

Method

Participants

Participants were 30 undergraduate students from the University of Waterloo, who received $8 for their participation in the experiment.

Stimuli

3-D object space. A 3-D object space was created with eight 3-D, graspable objects having different combinations of curvature, tapering and thickness. Each object was 81 mm long (see Figure 1). Objects were milled out of hard PVC plastic, mounted on a white piece of Styrofoam, and displayed through a window at approximately 12.24º of visual angle.

3-D action space. A 3-D action space was created using perturbations of pulling, sliding and twisting (see Figure 2 and Table 1). Actions were performed on an 81-mm-long cylinder that was 48 mm in diameter. The cylinder was distinct from the eight objects used to construct the 3-D object space, and was mounted on a manipulandum (depicted in Table 1). One manipulandum was placed in front of the experimenter, and a second manipulandum was placed in front of participants. The manipulandi allowed users to move the (mounted) cylinder by sliding it 15 cm, pulling it out by 15 cm, or rotating in such a way that the top portion of the hand moved a distance of 15 cm.

Eight nonwords—*baiv*, *fint*, *grov*, *hong*, *jorl*, *malg*, *verp*, and *yoot*—were used to label the eight objects. Another eight nonwords—*bryf*, *cauv*, *demb*, *gyte*, *lafe*, *tarb*, *vook*, and *weff*—were used to label the eight actions. All nonwords were equated in bigram frequency.

Procedure

Action naming. For this part of the experiment, the participants performed learning and test trials.

Action learning trials. A hollow cylinder was mounted on each manipulandum. On learning trials, the experimenter placed an 8.9 × 6.3 cm card on the manipulandum in front of the participants. The card bore a nonword that was to be associated with an action. The experimenter demonstrated the action by manipulating the cylinder mounted on the manipulandum to the right of the participant. Participants were instructed to look at each action, as well as its name; no response was made. Each action took approximately 2 sec to perform, and actions were presented in random order within a block of (eight) learning trials.

Action test trials. On test trials, the experimenter performed one of the eight actions on the cylinder mounted on the manipulandum that was to the right of participants. A 25.5 × 18 cm card containing all eight possible nonwords was placed in front of participants. Participants were instructed to "name" the action with its correct nonword label or to guess by selecting one of the eight alternatives if they could not remember the name of the action that was just performed. Test trials continued until all eight actions had been presented. Within a block of test trials, actions were presented in random order. Participants did not receive feedback on their performance.

Object naming. This part was also divided into learning and test trials.

Object learning trials. Participants sat in front of a window with shutters that allowed participants to view the 3-D objects for 2 sec.

On learning trials, each of the eight objects was presented for two seconds accompanied by its nonword label typed on a 8.9 × 6.3 cm card. Participants were instructed to look at each object, as well as its name; no response was made. The shutters were then closed for two seconds between object presentations. The eight objects were presented in random order within a block of (eight) learning trials.

Object test trials. On object test trials, participants were presented with one of the eight objects, along with a 25.5 × 18 cm card containing all eight possible nonword labels. Participants were instructed to "name" the object with its correct label, and to make their best guess if they could not remember using one of the eight alternatives. The object remained in front of the participant until the experimenter recorded the participant's answers. Test trials continued until all eight objects had been presented. Within a block of test trials, objects were presented in random order. Participants did not receive immediate feedback on their performance.

The experimenter interleaved the action-naming and object-naming tasks. Half of the participants were first presented with action naming (learning and test trials) followed by object naming (learning and test trials), and the other half performed the tasks in the reversed order. Testing continued until all eight objects and all eight actions had been correctly named by the participant for 16 consecutive trials. Testing time was approximately 45 min.

Results and Discussion

The number of confusions for each pair of actions and each pair of objects was collected for all participants, and data points that were three standard deviations beyond the mean were removed. This resulted in the removal of data for two participants for action identification; no data was removed for object identification. The numbers of confusions were analyzed separately for action naming and object naming.

Action Naming

Each error was classified according to which other action the presented action had been confused with. For example, if action 0,0,1 was presented but participants used the label associated with action 0,1,1, the error was classified as a confusion between actions that are 1 CB apart (1ACB). Because there were 12 pairs of actions that are 1 ACB apart, 12 pairs of actions that are 2 ACBs apart, and 4 pairs of actions that are 3 ACBs apart the number of confusion errors in each category was divided by the number of pairs of actions that could generate such errors. These averages were entered in a one-way ANOVA.

The results are shown in Table 2. The analysis revealed a main effect of similarity [$F(2,54) = 3.69$, $p < .05$]. Planned directional comparisons revealed that actions that were 1 ACB apart were confused more often than actions that were 3 ACBs apart [$t(27) = 3.06$, $p < .005$]. Other differences, although nominally in the predicted direction, were not significant.

Object Naming

Again, three averages were calculated for each participant, based on the number of object confusion errors between the 12 pairs of objects that were 1 visual CB apart (1 VCB), the 12 pairs of objects that were 2 visual CBs apart (2 VCBs), and the 4 pairs of objects that were 3 visual CBs apart (3 VCBs). These average values were entered in a one-way ANOVA.
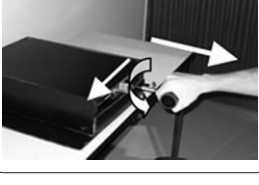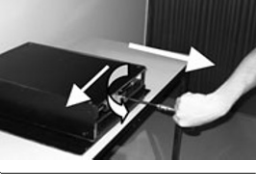
The results are shown in Table 2. The analysis revealed a main effect of similarity [$F(2,58) = 11.32$, $p < .001$]. Planned directional comparisons showed that pairs of objects that were 1 VCB apart were confused more often than pairs of objects that were 2 VCBs apart [$t(29) = 1.59$, $p = .06$] and also more often than pairs of objects that were 3 VCBs apart [$t(29) = 4.26$, $p < .001$]. Pairs of objects that were 2 VCBs apart also generated a mean number of

Table 1
Still Depictions of the Eight Actions

Action	Start Position	Halfway Position	End Position
Hold			
Pull			
Slide			
Twist			
Pull + Slide			
Pull + Twist			
Slide + Twist			
Pull + Slide + Twist			

Table 2
Mean Numbers of Object Identification Errors and Action Identification Errors (With Standard Errors) to Naming Objects and Actions, As a Function of Similarity (City Block [CB])

		1 CB		2 CBs		3 CBs	
		M	SE	M	SE	M	SE
Experiment 1	Actions	.53	.05*	.44	.05	.35	.06*
	Objects	.48	.06*	.38	.04*	.19	.05*
Experiment 2	Actions	.54	.04*	.45	.03*	.29	.03*

*Conditions significantly different from each other within a given row.

objects confusions that was greater than pairs of objects that were 3 VCBs apart [$t(29) = 3.33$, $p = .001$].

Both 3-D action identification and object identification followed the expected general pattern: Actions that were similar were confused more often than actions that were dissimilar and 3-D objects that were similar were confused more often than 3-D objects that were dissimilar. Thus, in the representational space used by memory, both object and action identification were constrained by similarity. Because actions were performed on a cylinder and were therefore not linked to any specific objects, it was not yet possible to assess the impact of action similarity on visual object identification.

EXPERIMENT 2

The goal of Experiment 2 was to assess the impact of action similarity on visual object identification. The same objects were paired either with similar actions or with dissimilar actions. If, as predicted by the HIT model (Humphreys & Forde, 2001), action information can serve to disambiguate between object representations in memory, specific objects will be confused more often when they are associated with similar actions than when they are associated with dissimilar actions.

Method

Participants
Participants were 60 undergraduate students from the University of Waterloo, who received eight dollars for their participation in the experiment.

Materials
The materials were the eight objects and the hollow cylinder, the eight actions, and the manipulandi used in Experiment 1. The specific pairings between objects and actions are depicted in Table 3 and Table 4. In Pairing 1 (used for half the participants), the twelve pairs of objects that were the most visually similar (1 VCB apart) were associated with similar actions (1 ACB apart), and twelve pairs of dissimilar objects (eight pairs that were 2 VCBs apart and 4 pairs that were 3 VCBs apart) were associated with dissimilar actions (actions that were either 2 ACBs or 3 ACBs apart). In Pairing 2 (used for the other half of participants) this setup was reversed.[1] The eight nonwords used to label objects in Experiment 1 were used in Experiments 2 and 3.

Procedure
Unlike Experiment 1, a given nonword label was used to describe a particular object *and its action* (akin to the noun and verb form of the word "hammer"). Participants were administered both action naming (learning and test trials) and object naming (learning and test) trials. Since an action and the object associated with this action

were both identified by the same label, care had to be taken that participants didn't just learn to associate the nonword label with the object and ignore that object's action. To prevent participants from doing this, all action naming trials were administered first, followed by the object naming trials.

Action naming. The learning and test trials of this phase were performed as follows.

Action learning trials. Participants were seated at arm's length from objects that were presented one at a time, mounted on the manipulandum directly in front of them. At the beginning of each learning trial, participants were instructed to keep their eyes closed. The experimenter signaled the beginning of a trial by saying "this is how you use this object." Upon hearing this phrase, participants opened their eyes and saw an object along with its written nonword label placed on the manipulandum in front of them. A copy of the same object was placed on the manipulandum in front of the experimenter, and the experimenter used this copy to demonstrate how the object was used by performing the action associated with the object. The experimenter then signaled the end of the learning trial by saying "now close your eyes." This sequence continued until all eight objects had been presented and all eight actions had been performed (one per object). Objects were presented in random order within a block of learning trials.

Action test trials. On test trials, participants were again instructed to keep their eyes closed until the beginning of the trial. The experimenter signaled the beginning of a test trial by saying "what was the name of this action?" The experimenter then performed one of the

Table 3
Specific Pairings Between Actions and Objects for Pairing 1, in Which All Similar Objects Are Associated With Similar Actions

Object	Curvature	Tapering	Thickness	Pull	Slide	Twist
	0	0	0	1	1	1
	0	0	1	1	1	0
	0	1	0	1	0	1
	0	1	1	1	0	0
	1	0	0	0	1	1
	1	0	1	0	1	0
	1	1	0	0	0	1
	1	1	1	0	0	0

Note—A "1" indicates that an attribute is present, and a "0" indicates that it is not. For simplicity, a "0" on thickness represented the minimum thickness within this shape set.

Table 4
**Specific Pairings Between Actions and Objects for Pairing 2,
in Which All Similar Objects Are Associated
With Dissimilar Actions**

Object	Curvature	Tapering	Thickness	Pull	Slide	Twist
	0	0	0	0	0	0
	0	0	1	0	1	1
	0	1	0	1	1	0
	0	1	1	1	0	1
	1	0	0	1	1	1
	1	0	1	1	0	0
	1	1	0	0	0	1
	1	1	1	0	1	0

Note—A "1" indicates that an attribute is present, and a "0" indicates that it is not.

eight actions using the 81-mm hollow cylinder. Participants were instructed to "name" the action by generating that action's nonword label and, if they could not remember, to guess (by selecting one of eight alternatives presented on a card). Once they generated a name for the action, participants then closed their eyes until the beginning of the next test trial. Test trials continued until all eight actions had been performed. Within a block of test trials, actions were presented in random order. Participants did not receive immediate feedback on their performance. A pattern of 8 action learning trials followed by 8 action test trials continued until all eight actions had been correctly identified for 16 consecutive trials. At this point, the action naming stage of the Experiment was terminated and the object naming stage began.

Object naming. The learning and test trials here were performed as follows.

Object learning trials. The learning trials for object naming were identical to the learning trials for action naming.

Object test trials. On test trials, participants were instructed to keep their eyes closed while the experimenter mounted a copy of the objects on the manipulandum in front of participants. The experimenter then signaled the beginning of a trial by saying "what was the name of this object?" Participants were instructed to "name" the object mounted with its associated label or to guess by choosing one of the eight labels form the card if they could not remember. On test trials each of the eight objects were presented twice using two, eight trial, random sequences (i.e., 16 object test trials followed the 8 object learning trials). This interleaving of 8 object learning trials followed by 16 object test trials continued until participants

were able to name the presented objects correctly on 16 consecutive trials (i.e., name each object correctly twice). Testing time took approximately 45 min.

Results and Discussion

The number of confusions for each of the pairs of actions and each pair of objects was collected for all participants, and data points that were more than three standard deviations beyond the mean were removed. This resulted in the removal of data for six participants for action naming (three for each Pairing), and for four participants for object naming (three in Pairing 1, and one in Pairing 2). The numbers of confusions were analyzed separately for action naming and object naming.

Action Naming

Errors were classified according to the distance between the confused actions (1ACB, 2ACB, 3ACB). The average numbers of confusions were entered in a two-way ANOVA in which similarity (1ACB, 2ACB, 3ACB) and pairing (Pairing 1, Pairing 2) were the factors.

The results are presented in Table 2. The analysis revealed a main effect of action similarity [$F(2,104) = 20.32$, $p < .001$]; no other effects were significant. Planned directional comparisons showed that pairs of actions that were 1 ACB apart were confused more often than pairs of actions that were 2 ACBs apart [$t(54) = 2.1, p < .05$], and more often than pairs of actions that were 3 ACBs apart [$t(54) = 6.42, p < .001$]. Also, pairs of actions that were 2 ACBs apart were confused more often than pairs of actions that were 3 ACBs apart [$t(56) = 4.04, p < .001$].

The findings of Experiment 2 replicated those of Experiment 1; similar actions were confused more often than dissimilar actions.

Object Naming

The purpose of this Experiment was to see whether the actions paired to objects influenced how these objects were visually identified. We collected the total number of errors produced by each participant on the 12 pairs of visually similar objects (1 VCB apart) and the 12 pairs of dissimilar objects (2 or 3 VCBs apart). For visually similar and dissimilar objects we contrasted the number of errors made by participants for whom these objects were paired with similar or dissimilar actions.

The results are presented in Table 5. For the visually similar objects, planned directional comparisons showed that participants produced more confusion errors when these objects were associated with similar actions than when the same objects were associated with dissimilar actions [$t(54) = 2.80, p < .005$]. In contrast, for dissimilar objects, associating the objects with similar actions did not result in more confusions than when the same objects were associated with dissimilar actions [$t(56) = 0.95$, n.s.].

The important aspect of this finding is that both groups of participants were asked to learn to identify the same objects—only the actions associated with these objects changed. Associating novel, visually similar objects with dissimilar actions served to make them dissimilar in mem-

ory, and consequently reduced visual identification errors. Associating dissimilar actions to dissimilar objects did not however reduce confusion errors in visual identification. This could possibly arise because confusion errors for these objects were already very low, resulting in a floor effect. In Experiment 3, we sought to increase the number of confusion errors by increasing the number of objects that participants had to name.

EXPERIMENT 3

Experiment 3 sought to replicate the finding that the actions paired with objects influenced how these objects were identified, and to extend the findings of Experiment 2 by using a learning regime that more adequately mimics the manner in which we learn to use tools in real life. Rather than learning the names of actions first, then learning the names of the objects that the actions were performed on at a later time period, in Experiment 3 we taught people the names of objects, and how they were used all at the same time. As in the previous experiment, we predict that objects will be confused more often when they are associated with similar actions than when they are associated with dissimilar actions.

Method

Participants
Participants were 60 undergraduate students registered at the University of Waterloo who received twelve dollars for their participation in the experiment.

Materials
The materials were the 8 gray objects and the two manipulandi used in Experiment 2, plus four new objects. These were copies of the 4 untapered objects used in Experiment 2. These objects were painted black, creating 4 extra stimuli. The 4 new objects were associated with the same actions as their gray counterparts, except that instead of being gripped with the whole hand, they were gripped only with the thumb and the index finger of the right hand, creating 4 new actions. The 8 nonwords used in Experiment 2 were used, as well as 4 new nonwords (*bryf*, *cauv*, *lafe*, and *tarb*). Thus, in total there were 12 objects (8 gray, and 4 black), 12 actions and 12 nonword labels.

Procedure
To more adequately mimic how tool names and their actions are learned in real life, in this Experiment the action learning phase and the object learning phase were merged.
Learning trials. Participants were shown an object accompanied by its nonword label, and an action was performed on the object. This was repeated for all 12 objects.
Test trials. During test trials, one of the 12 objects was placed on the manipulandum in front of participants. Participants were instructed to "name" the object or if they could not remember to guess by selecting a label from a card containing the 12 alternatives. Once participants named the object, they were then asked to perform the action associated with that object using the object mounted on the manipulandum in front of them. Testing continued until all 12 objects were presented. Within a block of test trials, objects were presented in random order. Twelve learning trials were followed by 12 test trials, and this interleaving of learning and test trials continued until all 12 objects had been identified correctly and all 12 actions had been performed correctly for 24 consecutive trials. Testing time took approximately 90 min.

Table 5
Mean Numbers of Errors (With Standard Errors) to Naming Similar and Dissimilar 3-D Objects, As a Function of the Similarity of the Actions Associated With Them

| | Similar Objects | | | | Dissimilar Objects | | | |
| | Similar Actions | | Dissimilar Actions | | Similar Actions | | Dissimilar Actions | |
	M	SE	M	SE	M	SE	M	SE
Experiment 2	7.93	0.66*	4.90	0.64*	3.17	0.62	3.21	0.65
Experiment 3	9.41	0.67*	7.77	0.68*	4.89	0.70	6.13	0.65

*Conditions significantly different from each other within a given row.

Results and Discussion

The four new objects and actions were used to increase the overall task difficulty and were not analyzed. The key data were the identification errors for the gray shapes. For each participant, the total number of object confusions for similar objects and dissimilar objects was collected and data points falling three standard deviations beyond the mean were removed. This resulted in the removal of data for five participants (one in Pairing 1, and four in Pairing 2).

The results are shown in Table 5. As in Experiment 2, for visually similar objects, assigning similar actions to these objects led to poorer performance than assigning dissimilar actions to these same objects [$t(53) = 1.62, p = .055$]. For dissimilar objects it did not appear to matter whether objects were assigned to similar or dissimilar actions [$t(56) = 1.39$, n.s.].

Importantly, in both Experiments 2 and 3 the similarity of the actions that were mapped to the visually similar objects determined how well these objects could be identified. As such the actions associated with objects played a meaningful role in visual object identification in both an artificial and a more ecologically valid version of the object-labeling task.

GENERAL DISCUSSION

The aim of this paper was to demonstrate that nonvisual information such as action knowledge can impact visual object identification. This goal was completed in three experiments. Experiment 1 showed that similarity governed a 3-D shape space, and an action space composed of combinations of three actions. Experiments 2 and 3 provided direct evidence that pairing actions with objects could influence visual identification performance. In this Experiment the exact same objects were paired with either similar actions or with dissimilar actions. Pairing objects with dissimilar actions afforded a significant reduction in object identification error rates for the visually similar objects.

There are two main findings from the present studies. The first main finding is that action identification follows the same pattern as object identification: similar actions were confused more often than dissimilar actions. The second main finding is that the similarity between the actions associated with objects can influence visual object identification. Specifically, similar objects were confused

more often when they were paired with similar actions than when they were paired with dissimilar actions.

Implications for Visual Object Identification

Visual similarity and action similarity. The findings concerning visual similarity observed throughout Experiments 1, 2, and 3 support the notion that visual similarity is an important component of visual object identification. Indeed, throughout these experiments, objects that were visually similar were always confused more often than objects that were visually dissimilar. The finding is reminiscent of the pattern of performance produced by CSVA patients. If a CSVA patient misidentifies an object, the patient is more likely to confuse the object presented with another object that is visually similar to it (Forde et al., 1997; Hillis & Caramazza, 1991; Vitkovitch & Humphreys, 1991). This relation between visual similarity and object confusions appears to be quite general. Similarity appears to account for the confusion errors among the healthy elderly and even nonhuman primates, who misidentify visually similar objects more often than dissimilar objects (Gaffan & Heywood, 1993).

An important way in which these experiments improve upon other studies of object identification is that in the current series of experiments the visual similarity of the objects that were employed could be empirically specified. In many studies investigating the relationship between object similarity and performance (whether it be naming, categorizing, etc.), similarity was either intuitively determined or determined by participants' subjective similarity ratings (Dixon et al., 1997; Dixon et al., 1999; Dixon et al., 2002; Schweizer, Dixon, Desmarais, & Smith, 2002). Dixon et al. (1997) and Desmarais and Dixon (2005) have shown that participants' subjective ratings of visual similarity do not adequately reflect similarity in memory space. In the current experiments, the consistent main effect of similarity attests to the importance of visual similarity in visual object identification.

The finding concerning action similarity in Experiment 1 and Experiment 2 nicely demonstrates that similarity governs memory confusions in multiple domains. In these experiments, actions that were similar were confused more often than actions that were dissimilar. Whereas we have shown that action similarity governed the errors made within this highly constrained action space, Ska and Croisile (1998) have shown that action similarity also governs the types of errors people make identifying actions linked to real world objects. They reported that healthy aging participants made visual errors when visually identifying actions. By showing a similar pattern in healthy young participants, our findings suggest that the influence of action similarity is a component of normal action identification, and not solely a consequence of aging. Furthermore, because all of the actions were combinations of sliding, pulling and twisting a cylinder or a *novel* object mounted on a manipulandum, the actions were decoupled from semantic knowledge (the twisting motion of a key) and the distance between actions was known and could be empirically specified in city-block distance. Thus, error rates in these experiments reflect purely the role of action

similarity decoupled from the objects on which these actions were performed.

Action similarity in visual object identification: The influence of nonvisual attributes in visual identification. In Experiment 2 and Experiment 3, since visual identification performance was tested simply by showing participants stationary objects and asking them to "name" them, the observed identification improvements when these objects were associated with dissimilar (as opposed to similar) actions means that nonvisual information (the actions associated with the objects) influenced participants' visual identification performance.

These findings extend those of Bub, Masson, and Bukach (2003). Bub et al. trained participants on gesture–color associations (if red then make an open grasp gesture). They then presented participants with colored objects that were either congruent or incongruent with the gesture associated with the color. Participants alternated between making the gesture that corresponded to the color or making the gesture that corresponded to the object and observed that participants' reaction times were faster for congruent trials than incongruent trials. Thus, gestures associated with the objects interfered with making color-driven gestures, and color-driven gestures interfered with making object-driven gestures.

Next, participants alternated between gesturing to the color and naming the object. Here they reasoned that if newly learned color–gesture associations could interfere with making gestures to the object, and *if* naming involved the automatic recruitment of functional knowledge, then there should be congruency effects when participants attempted to name the object. Bub et al. (2003) found strong congruency effects for the "gesture to color" condition but no congruency effects when participants named the colored objects. Interestingly though, when the authors looked at congruency effects during naming for each object separately, two of the objects did show significant congruency effects (needle and match). This led the authors to suggest that the volumetric properties of the objects relating to how objects were grasped (rather than the functional knowledge of how the objects are used) may have been automatically recruited during object naming.

Recent findings from Yoon and Humphreys (2005) also suggest that how objects are grasped can influence visual object identification. They asked participants to verify the names of hand-held implements and manipulated the congruency of the handgrip (in still photos) or the congruency of the object's movement (in motion clips). They reported an effect of congruency on name verification for an object's handgrip, but not for its associated movement—for positive responses, when presented objects were grasped incorrectly, response times were slower than when presented objects were grasped correctly. This did not occur when objects were presented with an incongruent motion. This also suggests that how objects are grasped may be automatically recruited during name verification.

In our experiments, the volumetric properties of the object were held constant. Therefore, our findings show that it is not just *how objects are picked up* that can influence object identification. Our findings show that it is *how the*

objects are used that significantly influenced object identification. One must be cautious to recognize that with real objects (as opposed to our novel ones) how an object is picked up is often related to how it is used. Our findings do, however, implicate that even when volumetric properties are controlled, how an object is used can influence how that object is identified.

Clearly, the actions associated with the objects influenced object identification. Exactly which aspect of these actions aided object identification is a matter of debate. The actions associated with these objects were learned visually; hence it is reasonable to assume that the effects of action on object identification were mediated by vision. However, work by Stefan et al. (2005) has demonstrated that observing actions is sufficient to form a *motor* memory. The authors recorded the baseline movements of the thumb produced under transcranial magnetic stimulation (TMS) of the primary motor cortex. They then either asked participants to practice a movement opposite that of the baseline or to observe the same opposite movement displayed on a computer screen, and reported that both physical practice and observational practice disturbed subsequent TMS-evoked movements of the thumb toward the "practiced" movement. This finding suggests that the observation of an action is sufficient to create a motor memory. For our purposes, however, the important point is that during visual identification of a stationary object, the actions associated with that object (whether visually based or motor engrams) influence identification. Hence identification is more than meets the eye.

Summary and Conclusion

The goal of this paper was to demonstrate the impact of nonvisual information on visual object identification. To achieve this goal, we first demonstrated that similar 3-D objects were confused more often than dissimilar 3-D objects, and that similar actions were confused more often than dissimilar actions. In Experiments 2 and 3, we demonstrated the effect of action knowledge on visual object identification by showing that similar objects that are paired to similar actions are confused more often than similar objects that are paired to dissimilar actions. These results specifically demonstrate that nonvisual information such as action knowledge influences visual object identification.

REFERENCES

Bub, D. N., Masson, M. E. J., & Bukach, C. M. (2003). Gesturing and naming: The use of functional knowledge in object identification. *Psychological Science*, **14**, 467-472.

Chainay, H., & Humphreys, G. W. (2002a). Neuropsychological evidence for a convergent route model for action. *Cognitive Neuropsychology*, **19**, 67-93.

Chainay, H., & Humphreys, G. W. (2002b). Privileged access to action for objects relative to words. *Psychonomic Bulletin & Review*, **9**, 348-355.

Damasio, A. R. (1990). Category-related recognition defects as a clue to the neural substrate of knowledge. *Trends in Neurosciences*, **13**, 95-98.

De Renzi, E., & Luchelli, F. (1994). Are semantic systems separately represented in the brain? The case of living category impairment. *Cortex*, **30**, 3-25.

Desmarais, G., & Dixon, M. J. (2005). Understanding the structural determinants of object confusion in memory: An assessment of psychophysical approaches to estimating visual similarity. *Perception & Psychophysics*, **67**, 980-996.

Dickerson, J., & Humphreys, G. W. (1999). On the identification of misoriented objects: Effects of task and level of stimulus description. *European Journal of Cognitive Psychology*, **11**, 145-166.

Dixon, M. J., Bub, D. N., & Arguin, M. (1997). The interaction of object form and object meaning in the identification performance of a patient with category-specific visual agnosia. *Cognitive Neuropsychology*, **14**, 1085-1130.

Dixon, M. J., Bub, D. N., Chertkow, H., & Arguin, M. (1999). Object identification deficits in dementia of the Alzheimer type: Combined effects of semantic and visual proximity. *Journal of the International Neuroscience Society*, **5**, 330-345.

Dixon, M. J., Desmarais, G., Gojmerac, C., Schweizer, T. A., & Bub, D. N. (2002). The role of premorbid expertise on object identification in a patient with category-specific visual agnosia. *Cognitive Neuropsychology*, **19**, 401-419.

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, **120**, 339-357.

Forde, E. M. E., Francis, D., Riddoch, M. J., Rumiati, R. I., & Humphreys, G. W. (1997). On the links between visual knowledge and naming: A single case study of a patient with a category-specific impairment for living things. *Cognitive Neuropsychology*, **14**, 403-458.

Funnel, E., & Sheridan, J. (1992). Categories of knowledge? Unfamiliar aspects of living and nonliving things. *Cognitive Neuropsychology*, **9**, 135-153.

Gaffan, D., & Heywood, C. A. (1993). A spurious category-specific visual agnosia for living things in normal human and nonhuman primates. *Journal of Cognitive Neuroscience*, **5**, 118-128.

Gainotti, G., & Silveri, M. C. (1996). Cognitive and anatomical locus of lesion in a patient with a category-specific semantic impairment for living beings. *Cognitive Neuropsychology*, **13**, 357-389.

Gauthier, I., James, T. W., Curby, K. M., & Tarr, M. J. (2003). The influence of conceptual knowledge on visual discrimination. *Cognitive Neuropsychology*, **20**, 507-523.

Hillis, A. E., & Caramazza, A. (1991). Category-specific naming and comprehension impairment: A double dissociation. *Brain*, **114**, 2081-2094.

Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioural & Brain Sciences*, **24**, 453-509.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, **5**, 67-103

James, T. W., & Gauthier, I. (2003). Auditory and action semantic features activate sensory-specific perceptual brain regions. *Current Biology*, **13**, 1792-1796.

Lloyd-Jones, T. B., & Humphreys, G. W. (1997a). Categorizing chairs and naming pears: Category differences in object processing as a function of task and priming. *Memory & Cognition*, **25**, 606-624.

Lloyd-Jones, T. B., & Humphreys, G. W. (1997b). Perceptual differentiation as a source of category effects in object processing: Evidence from naming and object decision. *Memory & Cognition*, **25**, 18-35.

Rumiati, R. I., & Humphreys, G. W. (1998). Recognition by action:

Dissociating visual and semantic routes to action in normal observers. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 631-647.

Schweizer, T. A., Dixon, M. J., Desmarais, G., & Smith, S. D. (2002). Not all triads are created equal: Further support for the importance of visual and semantic proximity in object identification. *Brain & Cognition*, **48**, 537-541.

Shepard, R. N. (1987). Towards a universal law of generalization for psychological science. *Science*, **237**, 1317-1323.

Ska, B., & Croisile, B. (1998). Effects of normal aging on the recognition of gestures. *Brain & Cognition*, **37**, 136-138.

Stefan, K., Cohen, L. G., Duque, J., Mazzocchio, R., Celnik, P., Sawaki, L., et al. (2005). Formation of a motor memory by action observation. *Journal of Neuroscience*, **25**, 9339-9346.

Temple, C. (1986). Anomia for animals in a child. *Brain*, **109**, 1225-1242.

Tranel, D., Logan, C. G., Frank, R. J., & Damasio, A. R. (1997). Explaining category-related effects in the retrieval of conceptual and lexical knowledge of concrete entities: Operationalization and analysis of factor. *Neuropsychologia*, **35**, 1329-1339.

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 664-680.

Yoon, E. Y., Heinke, D., & Humphreys, G. W. (2002). Modelling direct perceptual constraints on action selection: The naming and action model (NAM). *Visual Cognition*, **9**, 615-661.

Yoon, E. Y., & Humphreys, G. W. (2005). Direct and indirect effects of action on object classification. *Memory & Cognition*, **33**, 1131-1146.

NOTE

1. Four pairs of objects that were two VCBs apart were associated with actions that were two ACBs apart for both Pairing 1 and Pairing 2. These four pairs were not analyzed.