Acta Psychologica 164 (2016) 90–95

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

# If you watch it move, you'll recognize it in 3D: Transfer of depth cues between encoding and retrieval

Frank Papenmeier [a,*], Stephan Schwan [b]

[a] University of Tübingen, Germany
[b] Leibniz-Institut für Wissensmedien, Tübingen, Germany

ARTICLE INFO

*Article history:*
Received 9 June 2015
Received in revised form 7 December 2015
Accepted 20 December 2015
Available online 5 January 2016

*Keywords:*
Object recognition
Stereoscopic displays
Memory
Depth perception
Depth from motion
Binocular disparity

ABSTRACT

Viewing objects with stereoscopic displays provides additional depth cues through binocular disparity supporting object recognition. So far, it was unknown whether this results from the representation of specific stereoscopic information in memory or a more general representation of an object's depth structure. Therefore, we investigated whether continuous object rotation acting as depth cue during encoding results in a memory representation that can subsequently be accessed by stereoscopic information during retrieval. In Experiment 1, we found such transfer effects from continuous object rotation during encoding to stereoscopic presentations during retrieval. In Experiments 2a and 2b, we found that the continuity of object rotation is important because only continuous rotation and/or stereoscopic depth but not multiple static snapshots presented without stereoscopic information caused the extraction of an object's depth structure into memory. We conclude that an object's depth structure and not specific depth cues are represented in memory.


Imagine you finally get your hands on your new smart phone. What would you do? Certainly as one of the first things, you would turn it around, look at the back of it, its color, the thickness, and so on, thereby building up a mental model of the new phone. Once you put it down on the table, go away and come back, you would still recognize the object on the table as your new phone. Thus, you must have built up a memory representation of it. But what is the nature of this representation and what information does it contain, in particular regarding an object's depth structure? With the present manuscript, we investigate this question by studying transfer effects of depth from motion during encoding, to depth from binocular disparity during retrieval. In particular, we examine whether rotating an object during encoding supports the extraction of an object's depth structure into memory that can subsequently be accessed by stereoscopic information. Anticipating the results of our present experiments, we found evidence for transfer between continuous object rotation as a depth cue during encoding and stereoscopic information as a depth cue during retrieval, thus showing that the memory representation supporting object recognition is not flat but contains depth information and that it represents the depth structure of objects instead of specific depth cues.

The nature of memory representations supporting object recognition has been the subject of a long standing debate, in particular regarding the mechanisms by which memory representations support the recognition of objects from novel viewpoints (e.g. Biederman & Gerhardstein, 1993; Tarr & Bülthoff, 1995). There are two main lines of theories. On the one hand, there are theories considering object memory as a representation of object structure, be it complete three-dimensional models of objects (Marr & Nishihara, 1978) or the composition of viewpoint-independent primitives (Biederman, 1987; Hummel & Biederman, 1992). On the other hand, there are theories that consider object memory as a representation of objects in the form of multiple 2D view-specific representations as seen during encoding (Bülthoff, Edelman, & Tarr, 1995; Ullman & Basri, 1991). Due to the opposing predictions regarding the viewpoint dependence of object recognition, much work was devoted to determining the conditions for either the viewpoint-dependent or viewpoint-independent representation of objects (e.g. Biederman & Bar, 1999; Biederman & Gerhardstein, 1993; Edelman & Bülthoff, 1992; Farah, Rochlin, & Klein, 1994; Hayward & Tarr, 1997; Rock & DiVita, 1987; Tarr & Bülthoff, 1995; Tarr & Pinker, 1990; Tarr, Williams, Hayward, & Gauthier, 1998). As it turns out, both lines of theories found a way to explain the spectrum of viewpoint-dependent and viewpoint-independent findings by either arguing that the degree of viewpoint dependence depends on the amount that the structural description of objects changes by object rotation (Biederman & Gerhardstein, 1993) or the extent to which interpolation and extrapolation processes can operate on the 2D view-specific representations (Bülthoff et al., 1995). Therefore, this line of research provided the important insight that object

* Corresponding author at: University of Tübingen, Schleichstr. 4, D-72076 Tübingen, Germany.
*E-mail address:* frank.papenmeier@uni-tuebingen.de (F. Papenmeier).

http://dx.doi.org/10.1016/j.actpsy.2015.12.010
0001-6918/

recognition is viewpoint-dependent under many conditions; however, it still remains unclear what information observers represent in memory, in particular regarding the depth structure of objects.

Studies investigating the role of depth on object memory often used stereoscopic presentations and therefore binocular disparity as a depth cue (e.g. Bennett & Vuong, 2006; Burke, 2005). When presenting an object stereoscopically, two slightly different images are presented to the left and right eye of the observer, resembling the different views of the eyes when viewing real objects. Binocular disparity resulting from those views acts as a strong depth cue during perception (Dövencioğlu, Ban, Schofield, & Welchman, 2013; Landy, Maloney, Johnston, & Young, 1995; Marr & Poggio, 1979). If observers use this depth information during the formation of memory representations, recognizing objects or detecting changes between successive presentations of objects should be easier with stereoscopic presentations than without stereo information (same image presented to both eyes), which is what was found in multiple studies for object recognition and face recognition (Bennett & Vuong, 2006; Burke, 2005; Burke, Taubert, & Higman, 2007; Edelman & Bülthoff, 1992; Humphrey & Khan, 1992; Lee & Saunders, 2011; Liu & Ward, 2006). Whereas this stereo advantage was specific for new viewpoints in some studies (Bennett & Vuong, 2006; Burke, 2005; Burke et al., 2007) it also occurred without viewpoint changes in others (Edelman & Bülthoff, 1992; Lee & Saunders, 2011). Importantly, this stereo advantage cannot be explained by the additional two-dimensional information present in two views as opposed to one view because only the stereoscopic presentation of the left and right eyes views but not a side-by-side presentation of both views results in an increased performance (Burke, 2005). Taken together, the presence of stereoscopic information supports object recognition suggesting the representation of depth information in memory. However, the nature of this depth representation is still unspecified, in particular its specificity to stereoscopic depth and potential transfer effects between depth cues usually found during visual perception (e.g., Nawrot & Blake, 1989).

Besides binocular disparity, monocular depth cues are available to the visual system such as shading, texture, or depth from motion (Landy et al., 1995). For example, viewing a rotating object with one eye only would also provide depth information, as is known from studies investigating the kinetic depth effect or structure from motion (Braunstein, 1962; Ullman, 1979; Wallach & O'Connell, 1953). During visual perception, these depth cues are integrated into a combined depth percept (Dövencioğlu et al., 2013; Ichikawa & Saida, 1996; Landy et al., 1995), however, there is some evidence to suggest that this is not necessarily the case (Tittle & Perotti, 1997). The integration of depth cues into a combined depth percept is particularly true for depth from motion and depth from binocular disparity as has been shown by many studies showing that visual adaptation to one of the two depth cues causes transfer effects on the other depth cue (Bradshaw & Rogers, 1996; Nawrot & Blake, 1989, 1991, 1993; Poom & Börjesson, 1999). Furthermore, in perceptual tasks observers can adjust the depth structure perceived through motion to match a successively presented stereoscopic display (Perotti, Todd, Lappin, & Phillips, 1998; Todd & Perotti, 1999). If the representation of depth in object memory resembles the visual perception of depth, we should observe similar transfer effects when presenting depth from motion during object encoding and depth from binocular disparity during object retrieval.

On the other hand, however, mental imagery, and therefore possibly also object memory, has been shown to differ from the visual perception of 3D objects under some circumstances (Lobmaier, Mast, & Hecht, 2010). In their study, Lobmaier et al. (2010) asked participants to view a real 3D object from different viewpoints and to judge the direction where the object pointed to in space. They also had a mental imagery condition in which participants first memorized the object from one perspective and then performed the same direction judgment task with the exception that no physical object was present but participants vividly imagined the object being placed at its original location. In a second experiment, they replaced the real 3D object with a 2D photograph of the object. They found that the pointing errors of participants in the mental imagery condition were akin to the 2D photographs condition and differed from the 3D objects condition, suggesting that the mental imagery of 3D objects differed from the visual perception of 3D objects. Therefore, we might find no transfer between depth cues in memory because the representation of depth in object memory might differ from the visual perception of depth.

With the present set of three experiments, we investigated the nature of depth representations in object memory by studying transfer effects of depth from motion during encoding to depth from binocular disparity during retrieval. After establishing such transfer effects in Experiment 1, we further investigated the contribution of continuous object motion and binocular disparity during encoding on the representation of a depth structure in object memory (Experiments 2a and 2b).

## 1. Experiment 1

Experiment 1 had two objectives: First, we investigated whether object rotation alone results in the formation of three-dimensional memory representation that can subsequently be accessed by stereoscopic information during retrieval. Second, we were interested in whether adding stereo information to object rotation during learning provides additional depth information used in the formation of object memory representations, thus causing a larger stereo effect during retrieval.

### 1.1. Method

#### 1.1.1. Participants

Thirty-two students (24 female; age: 19–30 years; mean age: 23.47 years) participated in exchange for monetary compensation. Fourteen participants reported normal vision, fifteen participants reported corrected-to-normal vision and three participants reported near-sightedness without correction. Our experiments were approved by the institutional review board and we gained informed consent from the participants in all experiments.

#### 1.1.2. Apparatus and stimuli

Stimuli were presented on a 55-in passive polarized stereoscopic display (horizontal interleaved) using the software Blender 2.63 and custom code written in Python 3.2. Participants were standing and placed at a marking line in order to ensure that they kept a viewing distance of about 190 cm to the screen. We presented molecular-like objects consisting of seven spheres connected by six edges (see Fig. 1). We created 176 such objects using an iterative algorithm that started with one sphere and then connected a random number of spheres to each newly created sphere until a total of seven connected spheres was reached. Sphere positions were restricted to prevent overlaps with other spheres. Spheres had a radius of 0.1 units within our virtual

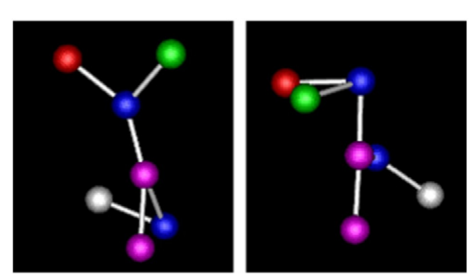

**Fig. 1.** Example of the stimuli (left) used in our experiments. Change trials (right) were constructed by rotating all but one edge by at least 30 degrees.