A priming study on naming real tools versus pictures of tools

Mutindi C. Kithu¹, Elizabeth J. Saccone¹, Sheila G. Crewther¹, Melvyn A. Goodale², Philippe A. Chouinard¹

¹ Department of Psychology and Counselling, School of Psychology and Public Health, La Trobe University, Melbourne, Victoria, Australia.
² The Brain and Mind Institute and the Department of Psychology, The University of Western Ontario, London, Ontario, Canada.

*Corresponding author:

Philippe A. Chouinard, Ph.D.
Senior Lecturer of Psychology
Applied Science 2 Building, Room 3.15
La Trobe University, Bendigo Campus
Bendigo, Victoria, 3550, Australia
Telephone: +61 3 5444 7028
E-mail: p.chouinard@latrobe.edu.au

Keywords: Priming, action representations, real objects, pictures, naming, functional use action, tools, embodied cognition.
Abstract

There is a growing body of literature demonstrating the relationship between the activation of sensorimotor processes in object recognition. It is unclear, however, if these processes are influenced by the differences in how real (3D) tools and 2-dimensional (2D) images of tools are processed by the brain. Here, we examined if these differences could influence the naming of tools. Participants were presented with a prime stimulus that was either a picture of a tool, or a real tool, followed by a target stimulus that was always a real tool. They were then required to name each tool as they appeared. The functional use action required by the target tool was either the same (i.e., squeegee-paint roller) or different (i.e., knife-whisk) to the prime. We found that the format in which the prime tool was presented (i.e., a picture or real tool) had no influence on the participants’ response times to naming the target tool. Furthermore, participants were faster at naming target tools relative to prime tools when the exact same tool was presented as both prime and target. There was no difference in response times to naming the target tool relative to the prime when they were different tools, regardless of whether the tools’ functional actions were the same or different. We also found more errors in naming target tools relative to the primes when different tools had a different functional action compared to when the same tool was presented as both the prime and the target. Taken together, our results highlight that the functional actions associated with tools do not facilitate or interfere with the recognition of tools for the purposes of naming. The theoretical implications of these results are discussed.
Introduction

A large body of literature has investigated the visual perception of objects and their unique influence on cognition and action (Gibson, 1979; Rosenbaum, Chapman, Weigelt, Weiss, & van der Wel, 2012; Valyear, Chapman, Gallivan, Mark, & Culham, 2011). Embodied theories of cognition have provided an approach to understanding the link between human perception and action by proposing that the experiences of the body within our environment form the basis of (Gallese, 2008; Lakoff & Johnson, 1999). Some examples of this include the activation of F5 canonical neurones in the macaque monkey when the monkey views an object or performs an action towards an object. Moreover, mirror neurons fire when the monkey both performs an action towards an object and when they observe another monkey perform the same action (Decety & Grèzes, 2006; Grèzes, Armony, Rowe, & Passingham, 2003; Rizzolatti & Arbib, 1998). Another example is the concept of “affordance”, introduced by Gibson (1979), which posits that the physical properties of a tool influences the way we select actions for that tool (Gibson, 1979). These embodied cognition theories suggest that sensorimotor processes play a role in cognitive processing related to the understanding of actions and objects. This notion is supported by neuroimaging evidence that demonstrates that sensorimotor related areas of the brain are activated when people passively view (Grafton, Fadiga, Arbib, & Rizzolatti, 1997), imagine (Grezes & Decety, 2001) and name actions associated with objects (Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995).

Likewise, behavioural evidence suggests the involvement of action representations in cognitive processes such as object recognition (Helbig, Steinwender, Graf, & Kiefer, 2010; McNair & Harris, 2012). We define action representations as the stored information about how we act upon and manipulate objects to perform their function. This involves the observer encoding
information about a range of features such as an object’s grip and/or gesture information to retrieve knowledge of how the tool is manipulated (Osiurak & Badets, 2017). With this in mind, Helbig, Graf, and Kiefer (2006) used a priming paradigm to investigate whether the action representations of a tool can facilitate visual object recognition. They presented participants with two pictures of objects that were either congruent (i.e., pliers - nutcracker) or incongruent (i.e., frying pan - banjo) in terms of how the objects are manipulated to perform their typical function (e.g., the snipping of scissors); hereafter called functional action. Participants were then instructed to name both objects in the order of presentation and the authors measured the participants' accuracy in responses. Reaction times (voice onset) were not recorded. The findings revealed that naming accuracy for object pairs that were congruent in their functional actions was greater than for incongruent pairs. Based on this, the authors concluded that action representations of objects can improve the recognition of other objects that involve similar motor interactions. These results demonstrate that processing of how a tool is functionally used can influence its recognition.

A second line of evidence suggesting that action representations can influence object recognition was provided by McNair and Harris (2012), who used a priming paradigm to investigate whether the grasp properties and functional actions associated with a tool can facilitate the subsequent recognition of a different tool with a similar grasp and/or functional use properties. For each trial, they presented participants with a picture of a prime tool followed by a picture of a target tool that either shared or did not share a similar grasp and/or functional action. Participants were then asked to identify the tools they had seen during the trial from an array of different tools presented throughout the study. The authors found that pairs of tools that had the same grasp were identified more accurately than those that differed in their grasp properties. In contrast, pairs of tools that had the same functional actions did not present a
similar advantage. Based on these findings, the authors concluded that target tools that are preceded by primes affording a similar grasp can facilitate their recognition because the process of selecting grasps is more readily and automatically assessed than functional use properties. The findings from both McNair and Harris (2012) and Helbig et al. (2006) support embodied theories because they demonstrate that action representations of tools can facilitate the recognition of tools.

In contrast to embodied cognition theories, there is another influential model of visual object processing that posits a distinction between perception and action within the visual system (Goodale, Króliczak, & Westwood, 2005; Goodale & Milner, 1992). According to Goodale and Milner’s two visual system hypothesis (TVSH; Goodale & Milner, 1992), the processing of visual information for object recognition (i.e., vision-for-perception) is mediated by the ventral visual stream, which includes visual areas in the occipito-temporal cortex that receives projections from the early visual cortex. On the other hand, the online control of actions such as reaching for and grasping objects (i.e., vision-for-action) is mediated by a functionally and anatomically distinct visual processing stream, the dorsal stream, which includes areas in the posterior parietal cortex that receive projections from the early visual cortex. In recent years, lines of evidence have expanded the two-visual system theory by suggesting an additional division in the posterior parietal cortex (Binkofski & Buxbaum, 2013; Buxbaum & Saffran, 1998; Osiurak & Badets, 2017; Rizzolatti & Matelli, 2003). According to Buxbaum (2001), the ventro-dorsal stream passes through areas in the inferior parietal cortex and mediates the long term retrieval of sensorimotor knowledge about tool manipulation. The more classical dorsal stream or dorso-dorsal stream (Binkofski & Buxbaum, 2013) mediates the online control of actions, which includes areas in the superior parietal cortex (Binkofski & Buxbaum, 2013).
The TVSH is supported by neuropsychological evidence from patients with brain damage to the ventral stream who present deficits in recognising objects but can still reach out and pick up objects of different widths and orientations (Carey, Harvey, & Milner, 1996; Milner et al., 1991). In contrast, patients with damage to the ventro-dorsal stream exhibit deficits in the ability to retrieve the conceptual knowledge about how we act upon and manipulate objects to perform their function, but preserve the ability to name the same objects (Johnson-Frey, 2004). Patients with damage to the dorso-dorsal stream, or the more classical dorsal stream, exhibit deficits in reaching and grasping objects but are still able to recognize and discriminate between these same objects (Goodale et al., 1994; Goodale, Milner, Jakobson, & Carey, 1991). Importantly, evidence for the TVSH challenges theories of embodied cognition because it demonstrates a dissociation between perception and action in the human visual system. The processing of visual information for retrieving stored sensorimotor knowledge about how to grasp and use a tool is distinct from recognising it.

The present study concerns the process of recognising an object for the purposes of naming it. In the case of naming an object, an individual has to first recognise the tool from its visual properties, such as its shape and visual texture and then retrieve the semantic representation of the tool in order to recall the lexical label of the tool for it to be named (Rothi, Raymer, Maher, Greenwald, & Morris, 1991). According to the TVSH, this process involves a heavy reliance on the ventral stream, whereas action representations, which are mediated by the dorsal stream, are not essential for this task. Conversely, theories of embodied cognition promote the involvement of action representations in tool identification.

Thus, the two theories make opposing predictions about whether or not action representations of tools are involved in the process of identifying and naming a tool. The TVSH predicts that
a tool’s action properties do not influence its identification and naming. The TVSH purports that naming a tool is purely a ventral stream process and identifying a tool is not aided by its action related information. Alternatively, embodied cognition theories predict that the action representations of the tools will aid in recognition for the purposes of naming. The current study aimed to test these predictions.

An important consideration is that the majority of studies within this area have employed 2-dimensional (2D) images as experimental stimuli. Yet, in the real world, tools are 3-dimensional that provide differential depth cues and structure. It is relevant that studies within this area also use real tool stimuli to draw inferences about how we process action representations of tools. The processing of real tools differ from 2D pictures of tools in various aspects (Gerhard, Culham, & Schwarzer, 2016; Snow et al., 2011; Snow, Skiba, Coleman, & Berryhill, 2014; Squires, Macdonald, Culham, & Snow, 2015). For instance, real 3D tools have depth information that provides volumetric cues to the viewer (Riddoch & Humphreys, 2001). Processing this kind of information leads to an understanding of the nature of the object and its position relative to our bodies to afford actions such as grasping and manipulating. Instead, pictures of tools are simply representations of tools displayed on a flat surface. They cannot be manually interacted with in the same way as a real tool. Research has also demonstrated that the differences in how we process real tools compared to pictures of tools can influence the recall of objects (Snow et al., 2014) and performance in the recognition of objects in patients with visual form agnosia (Chainay & Humphreys, 2001).

To further explore this area, Squires et al. (2015) used a priming paradigm to investigate how images of tools primed future actions compared to real tools. They presented participants with a prime tool that was either the same (i.e., spatula-spatula) or different (i.e., spatula-whisk) to
the target tool. The prime was either a photograph or real a tool. The target was always a real tool, which participants physically grasped either to transport it to a nearby pad (i.e., grasp to move) or to demonstrate its functional use (i.e., grasp to use). They then measured reaction (i.e., the time to initiate the action) and movement (i.e., duration time of movements) of the manipulation of the target tool. The authors found that participants had faster reaction times to initiate actions in the grasp-to-move compared to the grasp-to-use condition when the same tool was presented as both the prime and target. Importantly, this difference was the same regardless of whether the prime was a picture or a real tool. Based on these findings, the authors concluded that target tools that are preceded by real tool primes that have volumetric cues elicit the same priming effects as pictures of tools. Their results demonstrate that the representations of action related information in pictures versus real tools can facilitate the manipulation of real tools in the same manner. However, it is still unclear as to whether or not pictures and real tools can lead to facilitation in a similar manner when having to name the tools.

The current study investigated whether the action representations of tools can facilitate object recognition for the purposes of naming. And if so, can this effect of facilitation be influenced by whether the preceding tool is a real tool or a picture of a tool? To answer these questions, our study adopted a similar priming paradigm to Squires et al. (2015) except we had participants name both the prime and target tools rather than having them perform an action towards the target tool. We also asked participants to name both the prime and target tool as soon as they appeared – a task that does not require participants to hold information in working memory. We also incorporated a variety of pairs of tools that either shared or did not share the same functional use actions. The tool pairs we used were matched along several variables as determined by a separate pilot study (for more details, see “Pilot study”). The pairs either had the same functional action (SA), a different functional action (DA), or were identical (SAME).
We also had the prime tool in either a 2D pictorial format or 3D real format to allow us to determine how earlier encounters of the functional action properties of tools (i.e. same vs different functional actions) and also real (3D) vs pictorial (2D picture) format can influence naming responses to subsequent tools.

We had two competing hypotheses with respect to how the representations of action-related information offered by presenting real tools as opposed to pictures of tools had on the naming responses to those tools. The first hypothesis was that if we assumed a more embodied approach to how sensorimotor processes play a role in object processing for the purposes of facilitation in object recognition, then the characteristics of tools would facilitate the naming of the target 3D tool when the prime and target tool shared similar functional actions but were different tools. Namely, this facilitation, or priming, would be indexed by a decrease in reaction times and an increase in accuracy. We should also observe additional influences if the prime was the same format (i.e., a real 3D tool compared to when it was a picture of the same tool). Such a finding, however, would not imply that the dorsal stream is critical but rather that it can aid in object recognition, which is known to be mediated by ventral stream processes. The second hypothesis, in line with the TVSH assumed a more hermetically sealed role for the ventral stream whereby the similarity of functional actions between the prime and target tool stimuli would have no influence on tool naming. Moreover, this latter hypothesis would predict that the characteristics and volumetric cues presented by real tools would have no influence on tool naming. In other words, it would predict no effect of the format of the prime tool on the naming responses of the target tool.

**Method**
Overview

Our main experiment employed a priming paradigm where participants named each tool of the pair that was presented. The first tool (prime) was either a picture of a tool, or a real tool, while the second tool (target) was always a real tool. Our choice of real tools as the target stimulus was motivated by our interest in repeating similar measures applied by Squires et al. (2015). The pairs consisted of tools that either had the same functional action (SA) or a different functional action (DA). Lastly, there was a third condition in which the same tool stimulus was presented as both the prime and target (SAME). A pilot study was also performed for the purposes of determining pair stimuli in the main experiment in such a way that would ensure that the different conditions were matched along several extraneous variables (see Pilot study).

Pilot study

We conducted a pilot study to collect information about the tool stimuli to be included in the main experiment. This included familiarity (i.e., how familiar participants were with the use of the tool), functional use similarity between stimuli pairs (i.e., how similar the functional use actions were between two tools) and frequency of use in everyday life (i.e., how often the participant reported using the tool). The expected lexical labels for the tools in the main experiment were also based on the outcomes of this pilot study.

Pilot study: Participants

Ten right-handed individuals with reported normal or corrected-to-normal vision participated in the study (6 females, mean age: 24 years, age range: 18 – 27 years). All participants who
participated in the pilot study were excluded from the main experiment. All participants provided written informed consent and all procedures were approved by the Human Research Ethics Committee of La Trobe University in accordance with the Declaration of Helsinki. Participants were compensated for their time with gift vouchers.

*Pilot study: Stimuli*

The stimuli consisted of 16 colored photographs of tools that all shared the same handle (see Fig. 1). The photographs were processed in Adobe Photoshop software (Adobe Systems Incorporated, San Jose, CA, USA). The images were resized to 200×200 pixels at 150 dpi on a 23” Dell monitor with a resolution of 1024 x 768. The picture size was 18.67cm x 10.57cm. Participants were seated 40 cm away from the computer screen with the stimuli subtending a visual angle 15°. E-Prime 3.0 (Psychology Software Tools, Pittsburg, PA, USA) was used to present the stimuli and to collect the data. A Shure X2u XLR to USB interface microphone (Shure Distribution UK, Essex, United Kingdom) was used to record participants’ responses to naming the tools. Adobe Premiere Pro software (Adobe Systems Incorporated, San Jose, CA, USA) was used to analyze the recordings offline. The stimuli were presented under unlimited viewing conditions that enabled participants to initiate the next trial after each response.

*Pilot study: Procedure*

Our first objective was to come up with pairs of tools that had similar or dissimilar functional use actions. All possible pair combinations of tools from our stimulus set were presented. Participants viewed photographs of two tools side-by-side and were then asked to rate the
degree of similarity between the actions associated with each stimulus in a pair (i.e. “how similar are the actions associated with these two tools?”). They were then instructed to respond using a Likert rating scale between 1 and 5 by pressing numbers on a keyboard (1= not at all, 2= slightly, 3= somewhat, 4= moderately and 5= extremely). The tools were then rank ordered into pairs based on the lowest and highest mean scores across participants. The rank order was used to help define the ‘prime’ and ‘target’ tool pairs for the same functional action (SA) and different functional action (DA) conditions in the main experiment, with the former consisting of object pairs with lower scores and the latter consisting of object pairs with higher scores.

Our second objective was to ensure that the prime and target tools matched within and between the SA and DA conditions for, familiarity, frequency of use in everyday life and frequency in which they are used in language. For familiarity ratings, participants were presented with a photograph of each individual tool and asked to rate its familiarity (i.e., “How familiar are you with the use of the tool?”). Namely, they were instructed to respond using a rating scale between 1 and 5 by pressing numbers on a keyboard (1= not at all, 2= slightly, 3= somewhat, 4= moderately and 5= extremely). For frequency of use ratings, participants were presented with a photograph of each tool and asked how frequently they used it (i.e., “How often do you use this tool?”). They were asked to respond using a Likert rating scale of 1 and 5 (1= never, 2= rarely, 3= occasionally, 4= frequently and 5= very frequently). For naming use in everyday language ratings, participants were presented with a photograph of each tool and instructed to name the tool. The lexicon for a particular stimulus was determined by identifying the tool name that was most consistently produced by the participants. These were then matched to names of tools in a database of frequency used in language based on TV and film subtitles (Subtlex-US; American English; subtitle frequencies). After scoring the stimuli in this manner, stimuli comprising the SA and DA conditions for the main experiment were adjusted as needed.
to ensure the above nuisance variables were matched (see Table 1 for means and standard deviations for controlled factors).

Pilot study: Statistical analysis

The data were analyzed using the Statistical Package for Social Sciences (SPSS) version 23 (IBM Corporation; Armonk, NY, USA) and JASP software version 0.8 (University of Amsterdam, Amsterdam, Netherlands). A 2 x 2 analyses of variance (ANOVA) was performed for each dependent variable (familiarity, frequency of use, naming use in everyday language and naming frequency in a Subtlex database) and Presentation (Prime vs. Target) and Action condition (SA vs. DA) as within-subject factors. Significance was established when the corrected $p$ value was below .05.

Bayesian ANOVAs were also performed using the same 2 (Presentation) x 2 (Action Condition) analysis. The Bayesian analyses allowed us to determine if a different statistical approach might converge with the more traditional ANOVA, which would provide more confidence in the findings, and also draw more definite inferences from null results. The Bayes factor we report ($BF_{10}$) quantified the likelihood that the data support the alternative relative to the null hypothesis as a ratio between the two. We considered a $BF_{10}$ value of 3 or above as substantial evidence in favour of the alternative hypothesis and values of 0.33 or less as substantial evidence in favour of the null hypothesis (Jeffreys, 1998). As per the default specifications for priors used in JASP, the r scales for the Bayesian ANOVA were set to 0.5 and 1.0 for fixed and random effects respectively. Bayesian $t$-tests using a Cauchy prior set at 0.707 was used to evaluate significant main effects and interactions.
Pilot study: Results

In brief, the pairs of tools in both the SA and DA conditions matched in familiarity, frequency of use, naming use in everyday language and naming frequency in the Subtlex database. All means and standard deviations for controlled factors are shown in Table 1.

The ANOVA revealed no main effect of Action Condition for familiarity of use ($F_{(1,24)} = 0.000$, $p > .99$, $\eta_p^2 < .001$, $BF_{10} = 0.353$), frequency of use ($F_{(1,24)} = 0.298$, $p = 0.590$, $\eta_p^2 = 0.012$, $BF_{10} = 0.398$), naming use in everyday language ($F_{(1,24)} = 0.018$, $p = 0.894$, $\eta_p^2 < .001$, $BF_{10} = 0.356$) and naming frequency in the Subtlex database ($F_{(1,24)} = 0.000$, $p > .99$, $\eta_p^2 < .001$, $BF_{10} = 0.353$). There was also no main effect of Presentation for familiarity of use ($F_{(1,24)} = 0.009$, $p = 0.927$, $\eta_p^2 = 0.005$, $BF_{10} = 0.354$), frequency of use ($F_{(1,24)} = 0.429$, $p = 0.519$, $\eta_p^2 = 0.018$, $BF_{10} = 0.419$), naming use in everyday language ($F_{(1,24)} = 0.018$, $p = 0.894$, $\eta_p^2 < .001$, $BF_{10} = 0.356$) and naming frequency in the Subtlex database ($F_{(1,24)} = 3.479$, $p = 0.074$, $\eta_p^2 = 0.127$, $BF_{10} = 1.379$). Finally, there was no Action Condition X Presentation interaction for familiarity of use ($F_{(1,24)} = 0.106$, $p = 0.748$, $\eta_p^2 = 0.004$, $BF_{10} = 0.439$), frequency of use ($F_{(1,24)} = 0.107$, $p = 0.746$, $\eta_p^2 = 0.004$, $BF_{10} = 0.433$), naming use in everyday language ($F_{(1,24)} = 0.073$, $p = 0.789$, $\eta_p^2 = 0.003$, $BF_{10} = 0.428$) and naming frequency in the Subtlex database ($F_{(1,24)} = 0.003$, $p = 0.984$, $\eta_p^2 = 0.100$, $BF_{10} = 0.410$).

Main experiment: Participants

Sixteen right-handed individuals (11 females, mean age: 26 years, age range: 18-52 years) with reported normal or corrected-to-normal vision participated in the study. Scores from the
Edinburgh handedness inventory (Oldfield, 1971; scores could range from -100 to +100) indicated that all participants were strongly right-handed (mean score 85, range: 60-100). All participants provided written informed consent and all procedures were approved by the Human Research Ethics Committee of La Trobe University in accordance with the Declaration of Helsinki. Participants were compensated for their time with gift vouchers.

**Main experiment: Stimuli and materials**

The stimuli consisted of 14 real tools and 14 pictures of the same tools. The picture stimuli were created by photographing each individual tool placed on top of the LCD (liquid crystal display) monitor used in the study. We then edited the photographs of the tools using Adobe Photoshop software and set the picture of the tools onto a grey background and matched the tool size to fit the exact dimensions as the real tools. All tools had identical handles so that they required the same grasp aperture (see Figure 1).

The stimuli were presented as pairs with a ‘prime’ tool appearing first followed by a ‘target’ tool. The prime tool stimulus either consisted of a real tool or a picture of a tool. The target tool was always a real tool. The image pairs consisted of tools that either had the same functional action (SA) or a different functional action (DA). Lastly, there was a third condition in which the same tool stimulus was presented as both the prime and target (SAME). The presentation order within the pairs of tools were counterbalanced such that each tool within a pair appeared as often as the prime as it did the target for each condition (e.g., paint roller-squeegee, squeegee-paint roller). For the prime stimuli, each tool appeared the same number of times as a 2D picture as often as it appeared as a real 3D object.
We used E-Prime 3.0 to present the picture stimuli on a 17” Acer LCD monitor screen (Acer Inc., Xizhi, New Taipei, Taiwan) that was positioned horizontally on a turntable apparatus with the screen facing upwards. The monitor had a resolution of 1024 x 768 and the picture stimuli was resized to 2300 x 1300 pixels at 314 dpi. The picture size was approximately 18.75 x 10.56 cm and subtended with visual angles of 15° horizontally and 26° vertically. The real tool stimuli were placed on top of the same LCD monitor.

Participants wore PLATO visual occlusion goggles (Translucent Technologies Inc., Toronto, Ontario, Canada) that allowed us to have millisecond precision timing over the visual access to the stimuli. The goggles transitioned from clear (‘shutter open’) to occluding (‘shutter closed’) throughout the experiment. Participants also wore Bose QuietComfort QC35 Noise Cancelling Headphones (Bose Corporation, Framingham, MA, USA) that played white noise throughout the experiment. This masked external noise during stimulus changeovers. Lastly, participants wore a Shure X2u XLR to USB interface microphone that recorded vocal responses.

Main Experiment: Training session

The main experiment began with a training session that ensured that participants learned the correct names of each tool used in the study. Correct responses to tool names were defined and justified by the names of tools we established in the pilot study. Evidence from the training data revealed that 86% of participants named all the tools correctly in their first attempt, 92% in their second attempt, and 98% in their third attempt. The experimenter began the training session by reciting the names of each tool from a booklet that had pictures of the tools and their corresponding names printed below it. Thereafter, participants sat in front of a table that had
the same real tools laid out in front of them (see Figure 3). They were then instructed to name each tool at which the experimenter pointed. Each tool had to be named correctly at least 3 times before the participant could proceed to the main experiment.

Main experiment: Procedure

The main experiment procedure included 95 trials. There were 27 trials for each condition (SA, DA and SAME), as well as 14 catch trials, in which participants were instructed to remain silent when a blank screen appeared instead of the target stimulus. We included catch trials to discourage the perseveration of repeating the same tool name twice. Participants sat with their chin resting on a chin rest that was positioned 40 cm above the monitor screen.

As shown in Figure 4, each trial began with participants sitting comfortably wearing occluded PLATO goggles. The goggles opened and they viewed the prime for 500 ms and then named the tool as quickly and accurately as possible. The goggles then closed for a 3-s interval. The goggles then reopened and participants viewed the target tool for 500 ms and named it. Each trial was then followed by a 5-s inter-trial interval (ITI) in which the goggles remained closed, whilst the experimenter prepared the tools for the following trial. Each stimulus was presented in one of 3 different orientations (50°, 90° and 120°) relative to the centre of the monitor, and the primes and targets were always in the same orientation in a given pair. We presented the stimuli in three different orientations to avoid the predictability that would develop if all the tools were always presented in the same orientation. Figure 4 outlines the experimental layout and design employed in the experiment.

Main experiment: Statistical analyses
The Statistical Package for Social Sciences (SPSS) version 23 (IBM Corporation; Armonk, NY, USA), JASP software version 0.8 (University of Amsterdam, Amsterdam, Netherlands), and MATLAB (MathWorks Inc. Natick, MA, USA) were used to analyse the data. Two dependent variables were analysed. The first was mean reaction time (RT) for accurate naming and the second was number of naming errors. RT was defined as the time between when participants began viewing the tool stimulus to when they began to name the tool. RT was measured using an in-house program written in MATLAB. The naming of each tool was plotted over time in milliseconds from the output of a voice recorder. The beginning of each tool name was selected on the plots manually as the onset of the vocal response occurred, a method that allowed us to exclude any type of vocalisations that were not the name of the tool (i.e., “umm”). The values were identified on a trial-by-trial basis by the same rater (M.C.K). Naming errors were expressed as a percentage. This was calculated as the total number of trials minus the total number of correct responses, divided by the total number of trials multiplied by 100 [((total number of trials- total number of errors) / the total number of trials * 100]. The means and standard deviations for RT and naming errors are also reported.

A 2 x 2 x 3 analyses of variance (ANOVA) were performed on each dependent variable with Presentation (Prime vs. Target), Prime Format (Picture vs. Real) and Action Condition (DA vs. SA vs. SAME) as within-subject factors. To further evaluate significant main effects and interactions, t-tests corrected for multiple comparisons using the Bonferroni method were performed. Significance was established when the corrected p value was below .05. All reported p values represent corrected values unless specified otherwise. Bayesian ANOVAs were also performed with the same 2 (Presentation) x 2 (Prime Format) x 3 (Action Condition) model as the classical ANOVAs and Bayesian t-tests were used to further evaluate any
significant main effects or interactions. The default prior settings in JASP were used for these Bayesian tests.

Results

Main Experiment: Results

Naming RT

In summary, participants were faster at naming target tools relative to prime tools when the exact same tool was presented twice (SAME condition) compared to all the other conditions. In addition, we observed that the format in which the prime tool was presented (i.e., a picture or real tool) had no influence on participants’ naming RT of the target tool. Furthermore, RTs for the target tool in the DA and SA conditions did not differ. The results are shown in Figure 5 and Table 2 provides descriptive statistics.

The ANOVAs demonstrated a main effect of Action Condition \((F_{(2,30)} = 66.113, p = .001, \eta^2_p = 0.815, BF_{10} > 1000)\). Pairwise comparisons revealed that naming RT in the SAME condition differed from the other two conditions (all \(p < .001, BF_{10} > 1,000\)) and all other pairwise comparisons did not differ (all \(p > .500, BF_{10} = 0.140\)). There was also a main effect of Presentation \((F_{(1,15)} = 66.725, p < .001, \eta^2_p = 0.816, BF_{10} > 1,000)\). Pairwise comparisons revealed faster naming RT to targets than primes \((p < .001, BF_{10} > 1,000)\). An interaction between Action Condition and Presentation was also found \((F_{(2,30)} = 66.26, p < .001, \eta^2_p = 0.815, BF_{10} > 1,000)\). This interaction was driven by faster naming RT to the target tool relative to the prime in the SAME condition compared to the other two conditions (all \(p < .001, BF_{10} > 1,000\)). All other
pairwise comparisons did not differ (all $p > .500$, $BF_{10} = 0.140$). No statistical differences in RT were found between action conditions for the prime stimuli, regardless of the format of the prime tool (i.e., a real or pictorial format: all $p > .958$, $BF_{10} = 1.293$).

There was no main effect of Prime Format ($F_{(1,15)} = 2.765, p = 0.117, \eta^2_p = 0.156, BF_{10} = 0.293$), nor was there an interaction between Prime Format and Action Condition ($F_{(2,30)} = 0.333, p = 0.719, \eta^2_p = 0.022, BF_{10} = 0.108$) or between Prime Format and Presentation ($F_{(1,15)} = 1.776, p = 0.202, \eta^2_p = 0.106, BF_{10} = 0.197$). There was also no three way interaction between Prime Format, Presentation, and Condition ($F_{(2,30)} = 0.021, p = 0.979, \eta^2_p = 0.001, BF_{10} = 0.165$).

**Naming errors**

Figure 6 shows the mean percentage of naming errors of the prime and target tools while Table 2 provides all other descriptive statistics. In brief, we observed that individuals made more errors in naming target tools relative to primes. Furthermore, participants made more naming errors in the DA condition compared to the SAME condition for target tools only.

The ANOVAs demonstrated main effects of Presentation ($F_{(1,15)} = 7.275, p = 0.017, \eta^2_p = 0.327$) and Action Condition ($F_{(2,30)} = 3.833, p = 0.033, \eta^2_p = 0.204, BF_{10} = 3.216$). Pairwise comparisons indicated that there were more naming errors to the target tools overall compared to primes ($p = 0.017, BF_{10} = 1.038$) and more errors in the SAME condition compared to the DA condition ($p = 0.011, BF_{10} = 4.781$). All other pairwise comparisons did not differ (all $p > 0.1$). An interaction between Presentation and Action Condition was also found ($F_{(2,30)} = 4.157, p = 0.025, \eta^2_p = 0.217$). This interaction was driven by more naming errors to the target tools relative to the prime ($p = 0.017, BF_{10} = 1.038$), with naming errors in the SAME condition
differing to all other conditions (DA vs. SAME: $p=0.005$, $BF_{10}=4.781$ and SA vs. SAME: $p=0.022$, $BF_{10}=2.792$). All other pairwise comparisons did not differ (all $p>0.99$). There was no main effect of Prime Format ($F_{(1,15)}=3.182$, $p=0.095\;\eta^2_p=0.175$) nor was there an interaction between Prime Format and Action Condition ($F_{(2,30)}=0.427$, $p=0.656\;\eta^2_p=0.028\;BF_{10}=0.141$) and Prime Format and Presentation ($F_{(1,15)}=0.082$, $p=0.779\;\eta^2_p=0.005$, $BF_{10}=0.259$). There was also no three way interaction found between Prime Format, Presentation and Condition ($F_{(2,30)}=1.267$, $p=0.296\;\eta^2_p=0.078$, $BF_{10}=0.397$).

Contrary to the outcomes of the classical ANOVA, the Bayesian ANOVA revealed inconclusive support for or against the main effects of Presentation ($BF_{10}=0.953$) and Prime Format ($BF_{10}=0.472$) and the interaction between Presentation and Action Condition ($BF_{10}=2.615$).

**Discussion**

The present study investigated whether the action representations between tools can facilitate object recognition for the purposes of naming. Specifically, we examined if this effect of facilitation could be influenced by whether the preceding tool was a real tool or a picture of a tool. We aimed to test two competing hypotheses. The first was that if we assumed a more embodied approach to how sensorimotor processes play a role in object processing for the purposes of facilitation in object recognition, then the characteristics of tools would facilitate the naming of the target 3D tool when the prime and target tool shared similar functional actions but are different tools. Furthermore, we should also observe additional influences if the prime was the same format (i.e., a real 3D tool compared to when it was a picture of the same tool). The alternative hypothesis in line with the TVSH, assumed a more hermetrical role
for the ventral stream whereby the similarity of functional actions between the prime and target tool stimuli would have no influence on tool naming. This meant that we would observe no influence of the characteristics and volumetric cues presented by real tools on tool naming. Our results supported the second hypothesis.

Contrary to our first hypothesis, our results showed that the functional action similarity between different tools did not affect the naming of the target tool relative to the prime. The reason for this could be that understanding how objects are manipulated requires more conceptual processing than what is required for their identification. This suggests that the perceptual processing of the tool’s physical properties that enable its identification is dissociable from the processes involved in accessing the knowledge of how that tool is used – as is predicted by the TVSH. This dissociation in processing mechanisms has been demonstrated in neuropsychological studies of some patients with ideomotor apraxia arising from brain damage. Patients with damage to areas within parietal areas surrounding the intraparietal sulcus have shown preservation of their ability to recognise and name objects but have deficits in the recognition of object-related actions such as the knowledge about an object’s function (Buxbaum, 2001; Heilman, Rothi, & Valenstein, 1982).

This dissociation has also been demonstrated in neurologically healthy subjects when investigating the role of functional-use gestures in object identification. Bub and his colleagues (Bub, Masson, & Bukach, 2003) had participants learn to perform different tool-related gestures in response to different colours. Then, in each trial, participants viewed a picture of a tool as well as a colour cue. They performed the colour-cued gesture, and responses were faster when the tool’s functional use action was congruent rather than incongruent with that gesture. That is, the objects’ functional use properties influenced performance on the gesturing task –
but this effect was not demonstrated when participants had to name the objects instead. The authors concluded that the retrieval of functional knowledge about the tool was not present when having to name it, and that the motor representations associated with the objects did not play a role in object identification. Our study, together with Bub et al.’s findings, suggest that processes involved in the functional knowledge of tools are not implicated in object recognition for the purposes of naming. Note that these findings are in line with a recent paper by Saccone, Thomas, and Nicholls (2020), which also showed that performance for tool naming was not influenced by the tools’ action-relevant properties.

In addition, patients with damage to areas within parietal areas have also shown their preserved ability to manipulate objects but have differential deficits in the knowledge of their function (Buxbaum, 2001; Goldenberg, 2009; Heilman et al., 1982). This dissociation has also been demonstrated in neurologically healthy subjects when investigating the role of processing motor information for retrieving knowledge about an object’s function compared to an object’s manipulation. Garcea and Mahon (2012) presented participants with object stimuli that were either in a word or pictorial format. They then examined how quickly participants could name which object out of a pair matched a third object on either manipulation (i.e., the action deployed upon the object; referred to in our study as functional action) or function. They defined function as the purpose and intended goal of using the object (e.g., the goal of cutting a piece of paper can be done with a pair of scissors or a knife). They found that participants were slower at naming matching objects that shared manipulation attributes in word stimuli compared to pictorial stimuli. Conversely, participants were faster at naming responses to matching objects that shared function attributes in word stimuli compared to pictorial stimuli. The authors concluded that the differences in functional compared to manipulation matching with word compared to picture stimuli show that these types of object knowledge are
dissociable. Furthermore, this dissociation is highlighted by how the processes involved in retrieving manipulation knowledge about an object is not necessary for retrieving information about an object’s function, especially when processing tools in a word format (that provide less representations of action related information than viewing a picture of a tool). Our study, together with Garcea et al.’s findings suggest that processes involved in retrieving knowledge about how an object is manipulated does not aid in object recognition for the purposes of naming. More importantly, these results challenge theories of embodied cognition as they demonstrate that the processing of visual information for retrieving stored sensorimotor knowledge about how a tool is manipulated does not aid in object recognition, specifically for the successful naming of tools. In line with our second hypothesis, our findings instead support the TVSH, which purports that identifying a tool is driven by ventral-stream processes and is not aided by the processing of the action related information of the tool.

We note that these ideas and our findings contrast the study by Helbig et al. (2006) and McNair and Harris (2012). Their findings support theories of embodied cognition by showing that action representations of objects can improve the recognition of other objects that involve similar motor interactions. One thing to note is that the differences in the Helbig et al. (2006) and the current study’s findings may be due to differences in measurements and task requirements. For example, the experiment implemented by Helbig et al. (2006) also involved participants naming the prime and target stimuli at the end of each trial. This task required the use of working memory, and therefore the retrieval of the stored motor information of the target tools may have been used as part of a recall strategy to recognise attributes that were similar to the prime tool. Alternatively, we measured participants errors and response times to naming tools in quick succession. We propose that the task demands involved in our study did not require participants to maintain any information in working memory of what they saw as a
target to complete the task. Instead, they responded to both the prime and the target as soon as they each appeared. In this view, the differences between our findings and Helbig et al.’s suggest that working memory mechanisms can be a factor that can aid in the perceptual analysis of conceptual (i.e., motor-related) knowledge between tool pairs. Another methodological difference worth considering is that all the tools in our study had identical handles so that they required the same grasp aperture. We did not want participants responses to be influenced by characteristics based on grasping styles rather than their functional actions.

It should also be noted that Helbig et al. had a shorter ISI (167 ms) than the ISI used in our study. We included a longer ISI as this was the length of time required for the experimenter to change over the prime stimuli to the target stimuli, when real tools rather than pictures are used. It could be argued that this duration is too long to induce priming effects; however, a number of studies including those from co-authors Chouinard and Goodale have used an ISI of 2 to 3 seconds and have demonstrated strong behavioral priming and fMRI adaptation (Huettel & McCarthy, 2000; Chouinard & Goodale, 2009; Chouinard & Goodale, 2012; Chouinard, Morrissey, Köhler, & Goodale, 2008; de Groot, Thomassen, & Hudson, 1986). Thus, the longer ISI in our experiment cannot account for the lack of facilitation in the condition where the prime and target tool shared the same functional action but were different tools. We are left to conclude that the differences in the results between studies may have been due to differential task demands.

Aside from the functional action conditions, our participants’ naming performance for target tools did not differ according to the 2D or 3D format of the prime. These findings are surprising in the context of other studies that have shown differences in processing real versus pictures of objects. For example, real tools have elicited stronger viewing preferences than pictorial
versions of the same objects in 7-9 month old infants (Gerhard et al., 2016). Real tools have produced better recall and recognition performance for real tools than matched colour photos of the same items (Snow et al., 2014) and can elicit differential repetition-related changes in haemodynamic responses (Snow et al., 2011). On the other hand, our findings agree with Squires et al. (2015) who also demonstrated a similar effect in relation to how the realness of a prime tool compared to a picture of a tool provides no difference in performing a grasp-to-use action on a given tool. One consideration is that in their study, participants were merely viewing the prime tool, and the volumetric properties associated with passively viewing the preceding tool was not sufficient to invoke the action representations of the subsequent tool. Our findings suggest that even though the realness of tools must provide depth cues that may influence the way we act upon them, the processing required to identify the tool is not influenced by these characteristics. We propose that the computations required for the retrieval of the semantic representations of a tool for it to be named are not influenced by the format in which the tool is presented. This means that the visual processing of the volumetric cues provided by real tools compared to pictures of tools do not differ in how they contribute to object identification. Our results support the notion that the visual processing of tools for their identification is mediated by ventral stream processes, as the higher-level knowledge about the realness of the tool does not influence how quickly the tool is named.

Given the above, our study challenges theories of embodied cognition that propose that cognitive operations, such as identifying objects, depend on sensorimotor processes. Our results demonstrate that the representation of action related information of tools does not influence naming, regardless of whether subsequently named tools share functional use properties or if they are presented in a real or pictorial format. According to the TVSH (Goodale & Milner, 1992), the findings from this investigation highlight that the perceptual processing
involved in naming tools is predominantly driven by ventral stream processes, regardless of the objects’ affordances. Needless to say, sensorimotor processes have been shown to be involved in the visual processing of objects for other purposes (Almeida, Mahon, & Caramazza, 2010; Chen, Snow, Culham, & Goodale, 2018).

Conflicts of interest

All authors declare no conflicts of interest.

Acknowledgements

This work was supported by the Australian Research Council (DP170103189).
Figure 1. The tool stimuli. A. All tool stimuli used in the main experiment with the expected tool names as determined by the pilot study. B. Tools used in the same functional action condition (SA). C. Tools used in the different functional action condition (DA). Tools that were the same exemplar presented as a pair (SAME). Each tool shown in TOOL 1 was only paired with the tool in TOOL 2. This meant that the tool pairs were never matched with a pair that was not pre-determined as sharing or not sharing a functional use action (as determined by...
the pilot study). All tools had identical handles and similar grasp aperture to ensure that any differences between conditions would be attributable to the functional action of the tool, and not the grasp.
Figure 2. Experiment set up. Participants sat facing the screen with their head held stationary on a chin rest. The visual occlusion goggles were coded to transition from transparent to occluded during the interstimulus interval (ISI) and inter-trial interval (ITI). The prime and target tools were either displayed on the LCD monitor or placed on top of the monitor. Participants wore noise-cancelling headphones that played white noise throughout the experiment and an interface microphone that recorded their verbal responses.
Figure 3. Training session set up. Participants sat facing the tools whilst the experimenter pointed to each tool. Thereafter, participant responded by naming each tool.
Figure 4. Experimental layout and design. The goggles opened for 500 ms to display the prime stimulus, which was either a picture of a tool displayed on the LCD monitor (in the same position as where the real tool would be placed) or a real tool placed on top of the monitor. The participant then named the prime tool from the onset of when the stimulus appeared. The goggles closed during the ISI for 3 s whilst the experimenter substituted the prime stimulus with the target tool. The goggles then re-opened and displayed the target stimulus to which the participants responded to by naming the tool. The trial was completed by the goggles closing again for an ITI of 5 s.
Figure 5. Naming reaction time (RT) results. The graphs display the mean ± SEM RT naming scores. A. Naming RTs for trials in the DA, SA and SAME condition with pictures of tools as the prime tool. B. Naming RTs for trials in the DA, SA and SAME condition with real tools as the prime tool. An interaction between Action Condition, and Presentation was found. Asterisks (*) denotes a significant difference with the SAME condition compared to all the other conditions after correcting for multiple comparisons (p < 0.01).
Figure 6. Naming error results. The graphs display the mean ± SEM percentage of naming errors. A. Naming errors for trials in the DA, SA and SAME condition with pictures of tools as the prime tool. B. Naming errors for trials in the DA, SA and SAME condition with real tools as the prime tool. A main effect of Action Condition was found. Asterisks (*) denote a significant difference with the SAME condition compared to the DA condition after correcting for multiple comparisons (p=0.011).
Table 1. Pilot study means and standard deviations for controlled factors

<table>
<thead>
<tr>
<th>Condition</th>
<th>Similarity scores (rating scale between 1 and 5)</th>
<th>Familiarity scores (rating scale between 1 and 5)</th>
<th>Frequency of use in everyday life (rating scale between 1 and 5)</th>
<th>Naming use in everyday language (binary score of 1=correct, 0=incorrect)</th>
<th>Naming frequency in the Subtlex database</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(A) Same function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paint roller</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Spatula</td>
<td>4.95</td>
<td>0.00</td>
<td>4.4</td>
<td>1.07</td>
<td>4.4</td>
</tr>
<tr>
<td>Trowel</td>
<td>4.25</td>
<td>1.32</td>
<td>4.9</td>
<td>0.32</td>
<td>4.6</td>
</tr>
<tr>
<td>Screwdriver</td>
<td>4.2</td>
<td>1.45</td>
<td>2.4</td>
<td>1.17</td>
<td>4.5</td>
</tr>
<tr>
<td>Pizza cutter</td>
<td>3.35</td>
<td>1.14</td>
<td>4.5</td>
<td>1.27</td>
<td>4.7</td>
</tr>
<tr>
<td>Cheese slicer</td>
<td>3.15</td>
<td>1.45</td>
<td>3.8</td>
<td>1.23</td>
<td>2.9</td>
</tr>
<tr>
<td>Mixing spoon</td>
<td>2.4</td>
<td>1.16</td>
<td>5</td>
<td>0.00</td>
<td>4.9</td>
</tr>
<tr>
<td>Dust pan brush</td>
<td>2.35</td>
<td>1.43</td>
<td>4.9</td>
<td>0.32</td>
<td>4.8</td>
</tr>
<tr>
<td><strong>(B) Different function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knife</td>
<td>1.45</td>
<td>0.71</td>
<td>4.5</td>
<td>1.27</td>
<td>4.9</td>
</tr>
<tr>
<td>Hand rake</td>
<td>2.75</td>
<td>1.37</td>
<td>2.9</td>
<td>1.29</td>
<td>4.9</td>
</tr>
<tr>
<td>Cheese slicer</td>
<td>1.3</td>
<td>0.42</td>
<td>3.8</td>
<td>1.23</td>
<td>2.4</td>
</tr>
<tr>
<td>Ping pong racket</td>
<td>1.6</td>
<td>0.42</td>
<td>4.8</td>
<td>0.42</td>
<td>4.4</td>
</tr>
<tr>
<td>Dust pan brush</td>
<td>Paint roller</td>
<td>Pizza cutter</td>
<td>Trowel</td>
<td>Mixing spoon</td>
<td>Screw driver</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>--------------</td>
<td>--------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>2.9 1.66 4.9 0.32 4.4 1.07 3.9 1.37 1.9 0.99 0.9 0.43 0.8 0.42 722</td>
<td>722</td>
<td>293</td>
<td>59</td>
<td>349</td>
<td>388</td>
</tr>
</tbody>
</table>
Table 2. Reactions times and errors for primes and targets.

<table>
<thead>
<tr>
<th>Condition</th>
<th>RT (ms)</th>
<th></th>
<th></th>
<th>Naming error (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prime M</td>
<td>SD</td>
<td>Prime M</td>
<td>SD</td>
<td>Target M</td>
<td>SD</td>
<td>Target M</td>
</tr>
<tr>
<td>(A) Picture Primes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>858.49</td>
<td>120.95</td>
<td>813.43</td>
<td>150.01</td>
<td>4.46</td>
<td>7.32</td>
<td>4.91</td>
</tr>
<tr>
<td>SA</td>
<td>864.25</td>
<td>124.01</td>
<td>828.55</td>
<td>122.66</td>
<td>3.13</td>
<td>5.81</td>
<td>6.25</td>
</tr>
<tr>
<td>SAME</td>
<td>875.77</td>
<td>141.11</td>
<td>588.60</td>
<td>82.29</td>
<td>1.79</td>
<td>3.19</td>
<td>1.79</td>
</tr>
<tr>
<td>(B) Real Primes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>840.93</td>
<td>156.75</td>
<td>809.84</td>
<td>131.20</td>
<td>3.57</td>
<td>3.69</td>
<td>6.70</td>
</tr>
<tr>
<td>SA</td>
<td>825.86</td>
<td>144.59</td>
<td>813.16</td>
<td>113.37</td>
<td>3.57</td>
<td>6.39</td>
<td>7.14</td>
</tr>
<tr>
<td>SAME</td>
<td>841.41</td>
<td>150.54</td>
<td>572.34</td>
<td>76.22</td>
<td>4.91</td>
<td>5.03</td>
<td>2.68</td>
</tr>
</tbody>
</table>
References


