Contents

Method

Participants

A total of 374 subjects completed the online experiment (see Table 1 for a breakdown of the gender and age of the sample). Subjects were recruited from both voluntary participation websites such as Prolific Academic (where they received payment at the rate of £5/hr), and via the in-school research participation system (where they received course participation credits).

Table 1: Gender and age (*sd*) of the current sample.

Gender	N	Age	
Female	*197*	33.18	*(11.26)*
Male	*170*	33.23	*(10.28)*
Non-binary	*2*	23.50	*(-)*
Unspecified	*5*	29.40	*(6.11)*
Total	***374***	**33.10**	***(10.76)***

To meet our YA requirements, all participants were required to be aged between 18-59 years (actual obtained range: 18-59 years). As our experiment involved typing the English labels for a range of image stimuli, we also asked subjects whether English was their first language; all but one participant indicated that English was indeed their first language (99.73%).

Materials

Items from the previous experiment were brought forward into the current study; these consisted of a random sample of line-drawn, everyday object stimuli (Rossion & Pourtois, 2004), along with their written-word counterparts. Before sourcing photographs to match these items, some were excluded due to unique restrictions that meant they could not easily be translated into photographic format (for example, the shapes "arrow" and "star" can not be represented similarly as photographs).

Photo stimuli were obtained by searching open-source, copyright-free image websites (e.g. Unsplash; Pexels) for photographs similarly depicting the line-drawn everyday objects (the full list of image references can be found in Appendix A). For each item, three unique photographs were selected. An emphasis was placed on variety when selecting the three images, with the aim of obtaining photographs that very closely resembled the line drawings and those offering more modern depictions. All photographs were imported into Adobe Photoshop (20.0.04 Release), where the background was removed in order to isolate the objects of interest from other potentially distracting visual details (this was completed manually using the magnetic lasso and polygonal lasso tools; the edges were either feathered by 1px or left un-feathered). Next, the orientation of isolated objects was adjusted to ensure they matched as closely as possible to their line-drawn counterparts (e.g. photographs of the 'boot' item were adjusted so the toe was facing left and the heel facing right, as in the line drawing); this was often achieved by flipping or mirroring the object to 'correct' the direction.

Despite isolating objects from their background, a small number of photographs still contained irrelevant and potentially distracting details (e.g. in one photograph of a 'piano', there was a sign atop the object that may have impacted how the item is encoded when ultimately used in recognition tasks). Such details were removed using the clone stamp and content-aware fill tools (some examples of these photo manipulations can be found in Appendix B). Any obvious text (e.g. brand names) and numbers were also removed from photographs using the same method. The primary aim of the current study was to obtain photographic stimuli that could be clearly separated from words and line drawings; if some of our photograph stimuli contained text, whilst others did not, it is conceivable that they may not be directly comparable in terms of recognition. There were a couple of exceptions to this rule, when such details happened to be integral to the depiction of an object (e.g. the numbers found on a ruler or clock).

The photographs were exported from Photoshop in ".png" format in both their original colour and in greyscale (by setting saturation levels to 0). Final edits were completed in Adobe Lightroom (Classic, 8.2 Release): exposure (brightness) adjustments were made on images that appeared too light or too dark; highlights were decreased if some areas were too bright compared to the rest of the photograph; shadows were raised if some areas were too dark compared to the rest of the photograph; noise reduction was applied to some items after isolating the subject had inadvertently made unwanted noise/grain more visible. The changes made to each image were systematically applied to both the colour and greyscale versions (e.g. if one variation of "shoe" had an exposure increase of .010 for the colour version, the greyscale version also received an exposure increase of .010). Some colour-specific adjustments were made to the colour photographs only, however; common photo artefacts such as chromatic aberration (purple fringing) were corrected, along with white balance normalisation. Finally, all photographs were placed on a 600x600 pixel white background, and made to fill this frame as much as possible (i.e. some items were restrained by height, whilst others were restrained by width).

Design

Across three blocks, all participants provided five types of response toward photo stimuli: i) Naming; ii) Familiarity; iii) Visual Complexity, iv) Colour Diagnosticity; and v) Mental Imagery Agreement. Excluding the Naming task (consisting of a typed single-word answer), all responses were provided on a 5-point Likert scale. Within participants, the maximum number of response type provided for any one item was two; Naming and Familiarity responses were paired in one block, Visual Complexity and Colour Diagnosticity responses were paired in another, and Mental Imagery Agreement responses were always presented in a separate block. The order of these three blocks was counterbalanced across participants. Toward each individual photograph, participants made only one or two types of response before moving on to the next item, and the same items were not repeated to participants. For each photograph, we obtained the five types of data needed by counterbalancing between participants (e.g. for the first variation of the "cat" photograph, the Naming and Familiarity data was obtained from one participant, the Visual Complexity and Colour Diagnosticity data was obtained from another, and the Mental Imagery Agreement data was obtained from another).

Procedure

Data collection was conducted via two online platforms; i) Qualtrics - a survey platform that allowed for straightforward collection of consent, demographics, and computer compatibility data, and ii) Pavlovia - an open-source experiment hosting platform for studies programmed in Javascript (Peirce et al., 2019).

In the Naming and Familiarity block, participants were first asked "What is the name of the item depicted?". Subjects were instructed to name each photograph as briefly and unambiguously as possible, with one name only, and respond by typing their answer into the response box. If they did not know the name of an item, or had a tip-of-the-tongue experience, participants were instructed to type "no" for their answer (the term "don't know" was avoided so as not to encourage subjects to deviate from single-word responses, as instructed). Following the naming judgement, with the same photograph still present on-screen, participants were next asked "How familiar is the item depicted?". Subjects were instructed to judge each photo according to how usual or unusual the item was in their realm of experience; specifically, we defined familiarity as "the degree to which you come in contact with, or think about, the concept", and encouraged participants to rate the concept itself rather than the particular way it was currently shown. Participants selected one value from the 5-point scale, ranging from very unfamiliar (1) to very familiar (5), and were encouraged to use the full range of the scale throughout the set of photographs.

In the Visual Complexity and Colour Diagnosticity block, participants were first instructed to respond to the question "How visually complex is this picture?" using a 5-point scale that ranged from "very simple" (1) to "very complex" (5). Complexity was defined to subjects as "the amount of detail in the picture"; in contrast to the familiarity ratings, participants were encouraged here to rate the complexity of the picture itself, rather than the real-life item. If the photograph shown was greyscale, subjects would simply move on to the next item. If the item shown was in colour, however, participants were also required to make a colour diagnosticity judgement. We defined this concept as “how typical / normal the colour of the item is”, instructing subjects to rate on a 5-point scale ranging from “Not at all diagnostic (i.e. this item could be in any other colour equally well)” (1) to ”Highly diagnostic (i.e. this item appears only in this colour in real life). Participants were instructed to utilise the full range of options on the scale when making visual complexity and colour diagnosticity judgements. After making these ratings, a fixation cross was presented during a 1s interstimulus interval.

Due to the slight change in procedure and increased task complexity, Mental Imagery Agreement ratings were always acquired in an individual block (i.e. not alongside any other response types). First, participants were presented with a written label for 3s (e.g. “cat”) and told to focus their attention on the word. Once the written word disappeared, a beep tone was played alongside the instruction “close your eyes and imagine this item” (subjects were encouraged to close their eyes and begin imagining the item as soon as they heard the tone, but the written instruction were included as a further prompt). After 3s a second beep tone sounded to alert subjects to open their eyes, where they were presented with a photograph of the item they had been instructed to imagine. On a 5-point scale, participants were asked to “rate the agreement between your mental image and the picture”, from “low agreement” (1) to “high agreement” (5). The degree of agreement was defined as “how similar your mental image of the item is to the picture shown”. A fixation cross was displayed for 1s before the next word item was shown.

All responses were self-paced; the timing was only controlled during the study/imagine section of the Mental Imagery Agreement block.

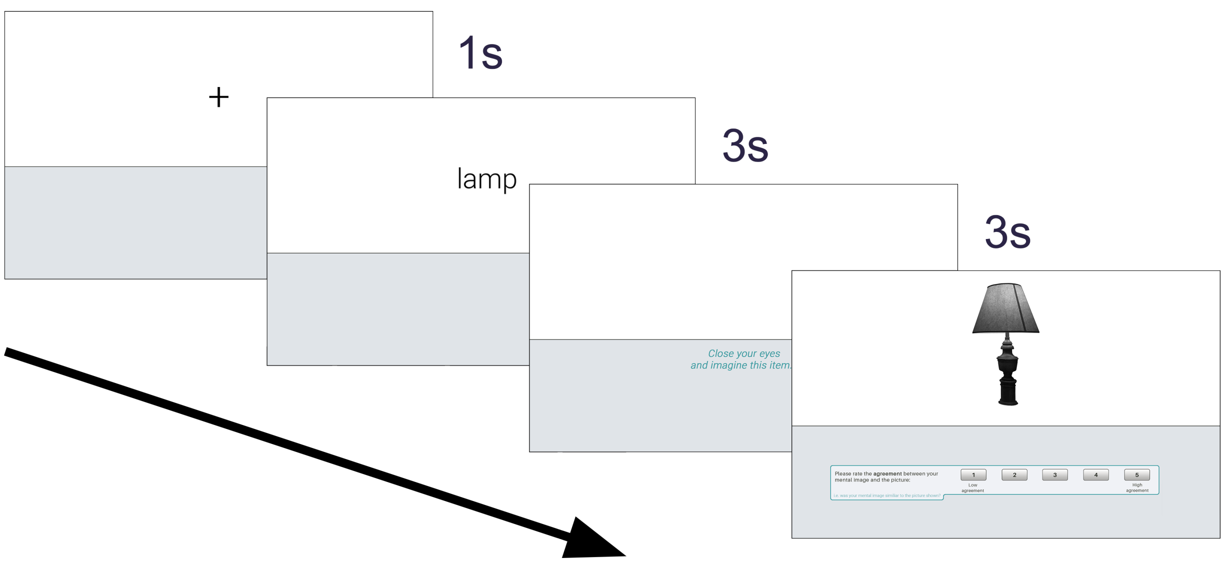

Figure 1: Figure 1: Data collection procedure for Mental Imagery Agreement responses.

References

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's Object Pictorial Set: The Role of Surface Detail in Basic-Level Object Recognition. *Perception*, *33*(2), 217–236. https://doi.org/10.1068/p5117

Appendices