

Contents

Chapter 1	5
Lit Review	5
Chapter 2	6
Experiment: Proof of concept	10
Method	10
Participants	10
Materials	11
Design	12
Procedure	13
Data processing	13
Results	14
Picture superiority	14
Response option availability	16
Recollection and Familiarity	17
Discussion	21
Chapter 3	23
Experiment: Development of a new set of standardised photographic stimuli	26
Method	27
Participants	27
Materials	27

Design	30
Procedure	31
Data processing	33
Analysis preparation	35
Results	36
<i>Naming</i>	36
<i>Mental imagery agreement</i>	37
<i>Familiarity</i>	38
<i>Visual complexity</i>	38
<i>Selection of final items</i>	38
Discussion	40
<i>The role of colour</i>	40
<i>Establishing a new set of stimuli</i>	42
Experiment: Effect of stimuli format (greyscale) and response option on recognition memory judgements.	43
Method	44
Participants	44
Materials	45
Design	45
Procedure	46
Data processing	47
Results	48
Overall hits, false alarms (FAs), and recognition	48

Discrimination (d') and response bias (c)	50
Hits assigned Recollection, Familiarity, and Guessing	52
FAs assigned Recollection, Familiarity, and Guessing	56
Discussion	60
Chapter 4 (The role of colour)	61
Experiment: Effect of stimuli format (colour) and response option on recognition memory judgements.	61
Method	61
Participants	61
Materials	61
Design	62
Procedure	63
Data processing	64
Results	65
Distinctiveness	65
Response option	66
Role of colour	67
Discussion	68
Chapter 5 (Distinctiveness of word stimuli)	69
Experiment: Manipulating word distinctiveness	69
Method	69
Results	69

Proportion of total hits	69
Proportion of total FA	71
Overall recognition:	72
d' scores	73
C scores:	74
Prop Rec hit:	75
Prop FAM hit:	76
Prop GUESS hit:	77
Prop REC FA:	78
Prop FAM FA:	79
Prop GUESS FA:	80
Results	81
Manipulation check	81
References	83
Appendices	84
Appendix A: <i>d'</i> calculations	84
Appendix B: Spelling corrections / manipulations to naming responses.	85
Appendix C: Normative data for all photograph items.	88
[@tidyverse]	

Chapter 1

Lit Review

Hamilton and Geraci (2006)

IMPLICIT MEMORY: PSE results from conceptual processing of a picture's distinctive features (rather than semantic information). General semantic task: "What is a used car sometimes called?" No PSE. Distinctive conceptual information task: "What fruit is egg shaped?" PSE.

EXPLICIT RECOGNITION: PSE always evident?

aMCI Show larger PSE effects than controls.

Impaired REC, so this PSE must rely on FAM?

Mixed findings whether fam is intact in aMCI. Intact - generally use picture stim. Impaired - generally use verbal stim.

Is PSE in aMCI driven by intact FAM for pictures, but impaired FAM for words? Yes (Embree, Budson, & Ally, 2012): aMCI - Picture FAM - same as healthy OAs aMCI - Word FAM - impaired compared to healthy OAs

Ally, McKeever, 2009: Examined early frontal old/new effect (FAM) in aMCI: Intact for pictures. Impaired for words. BUT, P did not provide subjective Rec/Fam reports.

Embree, Budson, & Ally, 2012: Deep encoding (verbal like/dislike response). Modified Old/New (6-point rating scale): 6. Certain the item is old - to - 1. Certain the item is new.

Both used the same picture stim - colour photos.

###-----

Chapter 2

Dual-process theories of recognition memory suggest that two independent processes - recollection and familiarity - are implicated in the successful recognition of previously encountered material [@paivio1971; @paivio1972]. Recollection typically refers to the conscious recall of encoded information, whereby contextual details (usually obtained by mentally re-experiencing a previous encounter with the stimulus) facilitate successful recognition. Familiarity, on the other hand, describes the unsubstantiated *feeling* of having encountered the stimulus before, and despite the inability to retrieve any associated diagnostic information, is still able to produce accurate recognition [@schoemaker2014]. While single-process accounts of recognition memory have been proposed, with the view that such experiences can be understood simply as varying levels of memory strength [@squire2007; @dunn2008], the majority of memory researchers agree that multiple processes are necessary to account for a range of dissociable experimental findings [@yonelinas2002]. Evidence from studies utilising event related potentials [ERPs; @curran2011], functional magnetic resonance imaging [fMRI; @scalici2017] and comparisons between healthy and clinical subject groups [e.g. Mild Cognitive Impairment; @belleville2011] all implicate the existence of two functionally distinct processes. Despite this consensus, disagreement persists in the literature regarding the extent to which recollection and familiarity are independent, and the methods that should be used to measure them most effectively.

Experiments into recognition memory often focus on obtaining separate estimates of recollection and familiarity, with experimental paradigms generally categorized into two broad groups: task-dissociation and process-estimation methods.

Task-dissociation methods combine two different task formats (e.g. old/new recognition paired with forced-choice recognition), and make comparisons in an attempt to isolate either recollection or familiarity [@yonelinas2002].

Old/new recognition tests are thought to isolate recollection; stimuli are presented individually and participants attempt to make an accurate judgement about items they previously studied (targets) when presented amongst novel items with highly similar characteristics (lures). Forced-

choice recognition paradigms, on the other hand, can rely on recollection or familiarity; discrimination between previously studied targets and highly similar lures is again required, however, recognition trials present target-lure pairs at the same time and require that participants select between them.

The wide range of studies utilising process-estimation methods generally use one of three main procedures. First, the process-dissociation procedure [PDP; @jacoby1991] is commonly used in recognition memory to separately estimate the contributions of recollection and familiarity by assessing whether participants are able to determine when and where an item was previously studied [yonelinas2002].

Third, and arguably the most commonly used process-estimation method is the Remember/Know (RK) paradigm - a task endorsed by a wide body of literature [@tulving1985; @yonelinas1995; @gardiner2000; @jacoby1991; @jacoby1997]. In a typical RK procedure, participants are generally tasked with making an 'old' vs. 'new' recognition decision, before introspectively categorising any recognised items (old) as either 'Remember' when recognition is accompanied by recollection, or as 'Know' when recognition is not accompanied by recollection, but instead is accompanied by a feeling of familiarity or certainty. Often a 'Guess' option is also provided to allow for uncertainty (RKG), in an effort to avoid biasing the total number of R and K responses (@belleville2011; @gardiner2008; @gardiner2002).

A large body of literature also reports that the RK procedure produces reliable estimations of recollection and familiarity in clinical populations [@lombardi2016]; for example, those with Mild Cognitive Impairment (MCI) typically show recollection impairments but intact familiarity compared to healthy older adults [@belleville2011; @wang2013; @lombardi2016; @hudon2009; @serra2010].

A second method of measurement is to ask participants to make confidence judgements or ratings and interpret high confidence as indicating recollection and low confidence as indicating familiarity. Instead of confidence, some researchers ask participants to make specific ratings of the amount of recollection and familiarity they have for each item they recognise on the memory

test; here, rather than a binary Remember-Know judgement which means that the processes of recollection and familiarity are seen as mutually exclusive, in this method recollection and familiarity are viewed as processes that can occur conjointly.

Finally, a forth method which uses judgements but permits participants to report that they have both recollection and familiarity for an item is the Recollect-Familiar-Both judgement task.

The objective of my research is to better understand what these different methodologies can tell us about the underlying processes of recollection and familiarity. In order to achieve this objective my proposed research will compare different methodologies against each other in different ways:

1. How to define the response options to participants: The way the ‘Know’ response is defined to participants can differ across researchers – some emphasise a feeling of ‘just knowing’ or ‘certainty’, whereas other labs use definitions that emphasise a ‘feeling of familiarity’. I want to know whether how we are defining ‘Knowing’ to participants when we use Remember-Know recognition tests changes how they use the response options. 2. Are ratings tasks better than judgement tasks: Recollection and familiarity or confidence ratings allow for a more fine-grained examination of memory processes but perhaps they are more complicated for participants to use and perhaps they are no more ‘accurate’ than more standard RK judgement tasks? I want to examine how patterns of responding differ across these methodologies.

The RK procedure has been criticized for its reliance on participants’ subjective understanding of the provided instructions [@schoemaker2014], and the introspective nature of recognition judgements make it difficult to confirm whether all participants have understood the definitions (and thus responded) similarly [@lombardi2016]. Indeed, slight differences in the experimental definition of “knowing” across studies and the way in which the subjective memory states are defined may affect the achieved pattern of results [@geraci2009; @williams2015].

In addition to the availability of different response options, there is evidence to suggest that the format of to-be-remembered stimuli

The picture superiority effect (PSE) – stimuli presented as pictures are markedly better remembered on tests of recall or recognition than stimuli presented as words (Shepard, 1967).

The magnitude of the PSE is greater in aMCI and AD patients when compared to healthy older (Embree, Budson, & Ally, 2012). Understanding this phenomenon helps to conceptualise how memory breaks down in AD.

Studies agree that pictures enhance recollection compared to words in healthy younger and older adults: Ally & Budson (2007); Ally et al. (2008); Curran & Doyle (2011); Rajaram (1996). However, it has been demonstrated that certain methodological factors can mediate the PSE, at least in certain populations; findings generally show intact familiarity for individuals with aMCI when *pictures* are shown during recognition memory paradigm, but impaired familiarity when word stimuli are utilised [@embree2012; @ally2009; @ally2009a; @wolk2008; @algarabel2009; @anderson2008; @hudon2009; @oconnor2010; @serra2010; @westerberg2006].

Is the magnitude of this enhanced recollection for pictures consistent across RFG, RFBG, and RF Ratings in healthy Ps?

May further demonstrate how the different response options are affecting how people respond (e.g. may reveal if certain conditions over- or under-estimate recollection?)

Are familiarity ratings also better for pics across the 3 conditions?

Is the achieved pattern also evident in healthy older?

What actually is magnitude? Difference between rec for words and rec for pics taken for each group (healthy vs. MCI), and then the difference of those compared??

Illustrative

The objective of Experiment 1 was to establish baseline PSE response patterns in younger adults, and examine how PSE magnitudes differ across the three response option conditions (RF-judgements / RFB-judgements / RF-Ratings). A 2x3 mixed factorial design was utilised, consisting of a within-subjects variable of stimulus type (words / greyscale line-drawings) and a between-subjects variable of response option (RF-judgements / RFBG-judgements / RF-Ratings).

To determine whether a PSE is evident in the current paradigm, d' (d-prime) scores will be calculated for each participant. d' is a signal detection statistic, calculated by taking the standardised difference between the signal (i.e. correct hits) and *lack of signal* (i.e. false alarms); in other words, d' offers a representation of participants' ability to distinguish target items from lures [@doi:<https://doi.org/10.1002/9781118445112.stat06743>]. A d' score of

[see Appendix A]

Within the currently proposed paradigm, the following results are hypothesised:

1. A PSE will be evident, manifesting as higher overall d' scores for pictures compared to words.
 - A PSE is also expected to manifest as i) a higher proportion of correct hits, ii) lower proportion of false alarms, and iii) better overall recognition.
2. Point 2.
3. Item 3
 - Item 3a
 - Item 3b

Experiment: Proof of concept

Method

Participants A total of 186 subjects completed the online experiment ($M = 26.7$ years ($SD = 10.36$ years; see Table 1 for a comprehensive breakdown of the sample). The current sample was primarily comprised of participants sourced from voluntary participation websites such as Prolific Academic¹ (52.15%) (where they received payment at the rate of £5/hr) and via the in-school research participation system² (where they received course participation credits; 41.4%). A small number of participants were also recruited from social media and other online sources

¹<https://www.prolific.co/>

²<https://keelepsychology.sona-systems.com/>

(Facebook: 3.76%; Call For Participants: 1.61%; Reddit: 0.54%; unspecified: 0.54%). To meet our YA requirements, all participants were required to be between 18-59 years of age (actual range: 18-59). As our experiment involved English word stimuli, we also asked subjects whether English was their first language; the vast majority (93.01%) reported that English was indeed their first language.

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

Gender	N	Age	
Female	122	26.02	(10.04)
Male	60	28.10	(10.98)
Non-binary	2	19.50	(2.12)
Unspecified	2	39.00	(0)
Total	186	26.70	(10.36)

Materials Pictures of innocuous, everyday objects (e.g. clock, rabbit, shoe) and their written-word names were sourced from @rossion2004. The picture stimuli consisted of greyscale line-drawn illustrations (containing shaded surface details), while word stimuli were simply the written-word names of each object presented in a clear Sans-serif typeface. A total of 136 unique items were randomly selected for use in the current experiment, from a pool consisting of: i) items with a written name between 4 and 7 letters; ii) items that would conjure the same intended concept in our UK-based sample (e.g. “ladder” should be universally understood across English-speaking cultures, whereas “wagon” or “pants” can be interpreted differently); iii) items that were not unknown, or uncommon, for our sample (e.g. Americanisms such as “wrench”); and iv) non-specific concepts such as “bird” (since the pool of items already contained specific exemplars of birds, such as “peacock” and “penguin”). As the current experiment involved memorising word stimuli, a single item (“glass”) was also removed as it shared too many letters with another item (“glasses”). Selected items were split into four separate lists for counterbalancing purposes; using the normative data provided by @rossion2004, each list was balanced based on the length of the written name, as well as scores of naming accuracy, familiarity, visual complexity, and mental imagery agreement. A series of independent samples t-tests confirmed that

no list was significantly different from another on any of the aforementioned criteria.

The picture stimuli utilised in the current study were created in Photoshop CC (20.0.04 Release), by importing the greyscale, surface-shaded, line-drawings onto a plain 250x250px white canvas. Written word stimuli were created using the Calibri sans-serif typeface on the same size canvas (see Figure 1 for example stimuli). All items were exported as .pngs files for presentation by the online survey platform.

bottle	ladder	orange	shirt
			

Figure 1: Example word and picture stimuli from the current study.

Design The current study utilised a mixed design, with a 2-level within-subjects factor of stimuli format (words, drawings), and a 3-level between-subjects factor of response option (RFG, RFBG, RF-Ratings). Subjects completed two study blocks - one consisting only of word stimuli, the other consisting only of picture stimuli - before completing a single mixed format recognition test, where previously studied word and picture items were randomly shown among new, unseen items. Subjects passed through 2 levels of blocked randomization during the experiment (equally sized, predetermined blocks). First, subjects were randomly allocated into one of two study block orders, which determined the order in which they were presented with the picture and word blocks at study. Second, subjects were assigned into one of three possible recognition tests (identical aside from the response options available when categorising recognition experiences): 1) RFG: “Recollection”, “Familiarity”, “Guessing”; 2) RFBG: “Recollection”, “Familiarity”, “Guessing”, “Both”, or 3) RF-Ratings: two independent 0-5 rating scales to separately report the

contribution of Recollection and Familiarity. These randomisation processes were completed automatically by the experiment software using balanced methods.

Procedure Data collection was conducted via the online survey platform Qualtrics³. Subjects initially completed an encoding block, where target words and pictures were randomly presented one-at-a-time on-screen. To ensure attention was directed to the presented stimuli, participants were required to respond to a simple encoding question toward each item at study: “Is this a picture or a word?”. This question allowed for the assessment of performance during the study block (to determine whether participants were concentrating at study), whilst also avoiding potential levels-of-processing effects that can accompany deeper encoding judgments (e.g. pleasantness ratings). The encoding phase was followed by a short distractor task comprised of 20 multiplication sums. Finally, subjects completed the recognition task, where they were again randomly presented with word and picture items one-at-a-time on-screen, and were required to respond “Old”/“New” depending on whether they recognised the item or not. “Old” responses were succeeded by a follow-up screen whereby participants were asked to report their recognition experience for the current item; the response options available during this follow-up response page differed between participants, with random allocation into either the RFG, RFBG, or RF-Ratings response option conditions. Recollection and Familiarity were defined identically across conditions, and the only deviations in instructions were: i) to define the additional “Both” response option in the RFBG condition; and ii) explain how certain responses should be reported in the RF-Ratings condition (i.e. subjects could still report a “Guess” in this condition by providing a 0-rating on both of the scales).

Data processing Measured variables included the total number of hits and FAs, and the total number of hits and FAs assigned to each of the available response options (RFG, RFBG, and RF Ratings). In order to create a common dependant variable, proportions were calculated from these variables in slightly different ways depending on the response option group. In the RFG-

³<https://www.qualtrics.com/uk/>

judgement group, simple proportions were created from the total number of R responses and the total number of F responses. In the RFBG condition, similar proportions were calculated by separately adding the proportion of Both responses to the proportion of R and proportion of F responses. In the RF-Ratings group, proportions of R and F were calculated based on the number of responses scoring >3 ; a response was classified R when subjects rated between 3-5 on the “Recollection” scale (regardless of the Familiarity rating), and a response was classified F when subjects rated between 3-5 on the “Familiarity” scale (regardless of the Recollection rating). The scales therefore allowed for pure R responses ($R=3-5 + F=0-2$), pure F responses ($F=3-5 + R=0-2$), both responses ($R=3-5 + F=3-5$) and Guessing responses ($R=0 + F=0$). Additional DVs included: i) d' (d-prime, a signal detection measure of sensitivity); ii) c-value (a measure of response bias); iii) overall accuracy (hits / (hits + FAs)); iv) reaction times for all responses.

d' scores were calculated by @Article{, title = {The Psycho Package: An Efficient and Publishing-Oriented Workflow for Psychological Science.}, author = {Dominique Makowski}, journal = {Journal of Open Source Software}, volume = {3}, number = {22}, pages = {470}, doi = {10.21105/joss.00470}, year = {2018}, note = {R package}, url = {https://github.com/neuropsychology/psycho.R}, }

A series of exclusion criteria were defined before analysis. First, subjects were to be excluded from analysis if they showed poor performance during the encoding task; the relative ease of reporting whether each item was shown as a word or picture prompted a performance cut off of 90% accuracy. This would allow for some accidental clicks, though subjects scoring less than 90% were to be excluded on the assumption they did not dedicate their full attention to the task. Second, subjects would be considered outliers (and thus excluded from analysis) if they presented extreme z-scores of ± 3 for total hits, total FAs, or overall recognition (hits minus FAs). However, no subjects were found to meet any of these criteria.

Results

Picture superiority To establish baseline picture superiority effects in the current paradigm, and assess whether there were any interactions with the availability of different response op-

tions at test, a series of 2 (stimuli format: words, pictures) x 3 (response option condition: RFG-judgements, RFBG-judgements, RF-ratings) mixed ANOVAs were conducted on a number of outcome variables. Namely, the proportion of overall hits, false alarms (FAs), and overall recognition (hits - FAs) [see Table 2]. To further examine response patterns between pictures and words, ANOVAs were also run on the signal detection measures of d' (measure of sensitivity) and c (decision criterion). Significant main effects and interaction effects were followed-up with Bonferroni-adjusted pairwise comparisons.

Table 2: Mean proportion of hits and FAs by stimuli-format and response-option condition.

	Hits	FAs	Overall recognition
Stimuli-format			
Words	0.47	0.21	0.27
Pictures	0.62	0.12	0.50
Response-option			
RFG	0.62	0.19	0.43
RFBG	0.54	0.16	0.38
RF Ratings	0.48	0.14	0.34

The ANOVA on the proportion of hits demonstrated a significant main effect of stimuli-format, $F(1, 183) = 131.77$, $MSE = 0.01$, $p < .001$, $\hat{\eta}_G^2 = .092$; a PSE was evident, with pictures ($M = 0.62$) showing a higher number of overall hits compared to words ($M = 0.47$), $t(183) = -11.48$, $p < .001$. Similarly, the ANOVA on the proportion of FAs also supported a PSE; a significant main effect of stimuli-format $F(1, 183) = 61.18$, $MSE = 0.01$, $p < .001$, $\hat{\eta}_G^2 = .084$ showed that words ($M = 0.21$) produced more FAs than pictures ($M = 0.12$), $t(183) = 7.82$, $p < .001$. Overall recognition performance (a measure that takes into account both hits and FAs) offered further support for a PSE in the current paradigm; a significant main effect of stimuli-format $F(1, 183) = 409.20$, $MSE = 0.01$, $p < .001$, $\hat{\eta}_G^2 = .236$ showed pictures ($M = 0.50$) produced better overall recognition on the task compared to words ($M = 0.27$), $t(183) = -20.23$, $p < .001$. No interaction effects were found between stimuli format and response option for any of these variables. Taken together, the findings demonstrate a replication of the PSE in the current memory paradigm, and suggest stimuli format plays a key

role in memorability that is independent of response option.

The ANOVA on d' scores produced findings consistent with those found for hits, FAs, and overall recognition; there was a significant main effect of stimuli-format, $F(1, 183) = 295.80$, $MSE = 0.18$, $p < .001$, $\hat{\eta}_G^2 = .223$, whereby pictures ($M = 1.62$) facilitated better discrimination between hits and FAs than words ($M = 0.86$), $t(183) = -17.20$, $p < .001$. C scores, however, showed no significant main effect of stimuli-format, $F(1, 183) = 2.31$, $MSE = 0.11$, $p = .130$, $\hat{\eta}_G^2 = .002$, suggesting response biases were similarly conservative between pictures and words. Neither the ANOVA on d' or C scores showed any significant interaction effects.

Response option availability In the ANOVAs on the proportion of hits, FAs, and overall recognition, the main effects of response option were also examined to determine whether the availability of different response option had an impact on task performance. There was a significant main effect of response option only in the ANOVA for the proportion of hits, $F(2, 183) = 6.46$, $MSE = 0.09$, $p = .002$, $\hat{\eta}_G^2 = .057$, with the RFG group ($M = 0.62$) showing more hits than the RF-Ratings group ($M = 0.48$), $t(183) = 3.60$, $p = .001$. The ANOVAs on the proportion of FAs and overall recognition showed no significant main effect of response option. This finding might indicate fewer available response options facilitate accurate recognition compared to more response options - however, the absence of a main effect for overall recognition (a measure that takes into account both hits and FAs) suggests this interpretation should be treated with caution.

Discriminability (d') between hits and FAs showed no significant main effect of response option, however, scores of response bias (C) did, $F(2, 183) = 6.44$, $MSE = 0.51$, $p = .002$, $\hat{\eta}_G^2 = .054$; those in the RF-Ratings condition ($M = 0.67$) showed higher c-scores (and thus a more conservative response bias) than those in the RFG condition ($M = 0.34$), $t(183) = -3.59$, $p = .001$. This indicates subjects were less likely to respond “Old” when they were required to provide more detailed follow-up recognition judgements (i.e., using separate 0-5 scales for R

and F), compared to simply selecting one of three options (R,F, or G).

Recollection and Familiarity To determine whether stimuli format, or the availability of different response options, had an impact on rates of R and F, additional a series of 2 (words, pictures) \times 3 (RFG-judgements, RFBG-judgements, RF-ratings) mixed ANOVAs were conducted on the mean proportion of hits and FAs assigned R, F, and G.

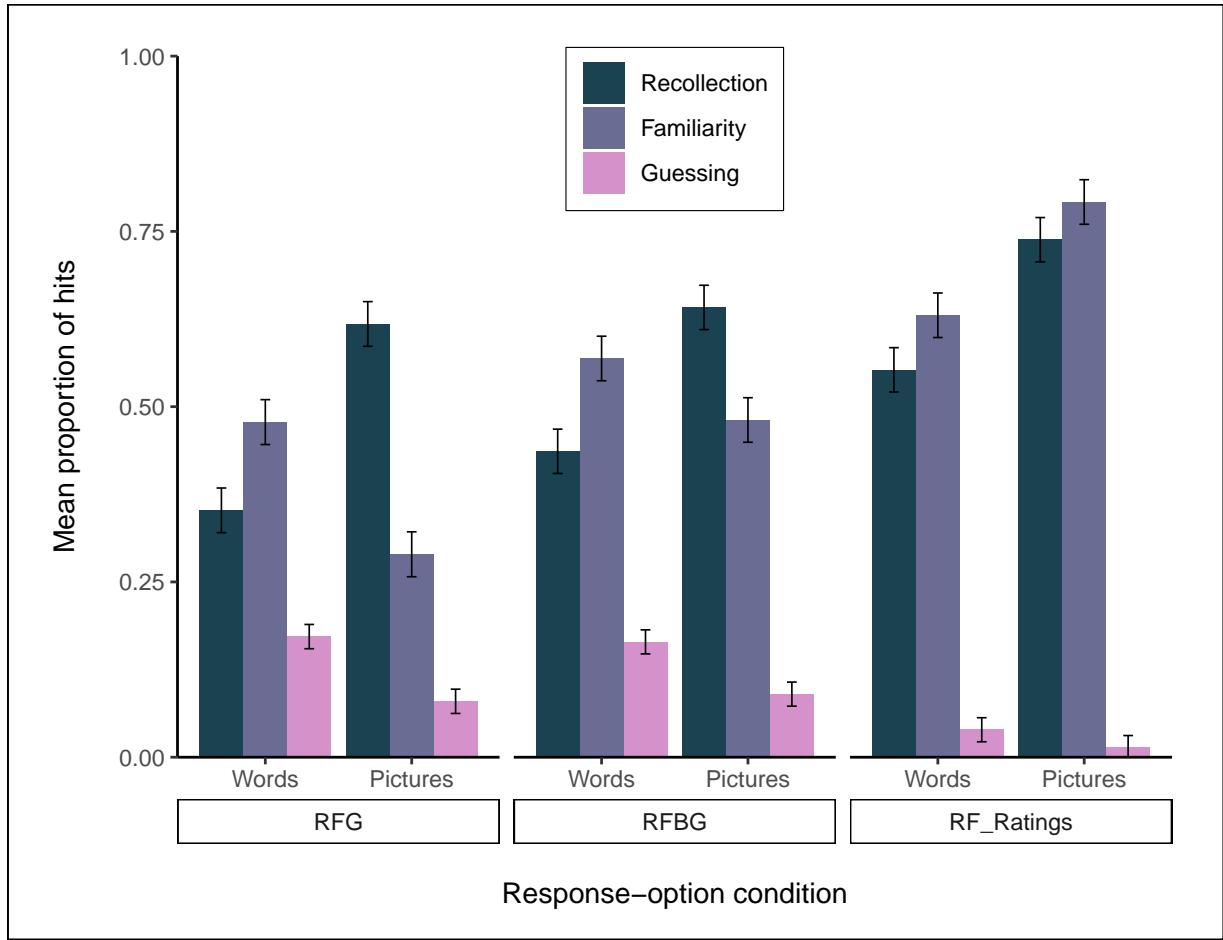


Figure 2: Proportion of hits assigned Recollection, Familiarity, and Guessing, by stimuli-format and response-option condition.

Recollection: For R hits, there was a significant main effect of stimuli-format, $F(1, 178) = 158.42$, $MSE = 0.03$, $p < .001$, $\hat{\eta}_G^2 = .167$, with pictures ($M = 0.67$) showing a higher proportion of Recollected hits than words ($M = 0.45$). There was also a significant main effect of response-option, $F(2, 178) = 8.55$, $MSE = 0.09$, $p < .001$, $\hat{\eta}_G^2 = .069$; the

RF-Ratings group ($M = 0.65$) showed significantly more Recollected hits compared to both the RFG-group ($M = 0.49$), , and the RFBG-group ($M = 0.54$), . There were no significant interaction effects, $F(2, 178) = 1.91$, $MSE = 0.03$, $p = .151$, $\hat{\eta}_G^2 = .005$. In the ANOVA on Recollection FAs, there was no significant main effect of stimuli-format, $F(1, 136) = 2.78$, $MSE = 0.04$, $p = .098$, $\hat{\eta}_G^2 = .005$, but there was for response-option, $F(2, 136) = 10.70$, $MSE = 0.12$, $p < .001$, $\hat{\eta}_G^2 = .106$, with those in the RF-ratings group ($M = 0.38$) showing more Recollection FAs than both the RFG-group ($M = 0.14$), , and RFBG-group ($M = 0.24$), . This could indicate that when subjects were required to individually report Recollection and Familiarity on two 0-5 rating scales, they were more likely to experience (or report) false recognition (accompanied by non-existent contextual details) than when only three (RFG) or four (RFBG) response-options were provided. There were no significant interaction effects, $F(2, 136) = 2.08$, $MSE = 0.04$, $p = .129$, $\hat{\eta}_G^2 = .008$.

Familiarity: F hits demonstrated a significant interaction between stimuli format and response option, $F(2, 178) = 34.42$, $MSE = 0.03$, $p < .001$, $\hat{\eta}_G^2 = .083$ (see Figure 3). Within response-option conditions, words resulted in more Familiarity hits than pictures in both the RFG group (words: $M = 0.48$; pictures: $M = 0.29$, $t(178) = 6.07$, $p < .001$) and RFBG group (words: $M = 0.57$; pictures: $M = 0.48$, $t(178) = 2.87$, $p = .005$). Conversely, the RF-Ratings group showed the opposite pattern, with more Familiarity hits produced for pictures ($M = 0.79$) than words ($M = 0.63$), $t(178) = -5.29$, $p < .001$. The ANOVA on Familiarity FAs did not yield any significant results; with no significant main effect of stimuli-format, $F(1, 136) = 1.12$, $MSE = 0.04$, $p = .292$, $\hat{\eta}_G^2 = .002$, no significant main effect of response-option, $F(2, 136) = 0.62$, $MSE = 0.15$, $p = .539$, $\hat{\eta}_G^2 = .007$, and no significant interaction effects, $F(2, 136) = 1.12$, $MSE = 0.04$, $p = .331$, $\hat{\eta}_G^2 = .004$.

Between response-option conditions, word stimuli produced significantly more Familiarity hits in the RF-Ratings group ($M = 0.63$) compared to the RFG group ($M = 0.48$), $t(276.78) = -3.37$, $p = .002$. For pictures, a higher number of Familiarity hits was evident in the RF-Ratings group ($M = 0.79$) compared to both the RFG group ($M = 0.29$), $t(276.78) = -11.13$, $p < .001$ and RFBG group ($M = 0.48$), $t(276.78) = -6.94$, $p < .001$. The

RFBG group ($M = 0.48$) also showed a significantly higher number of Familiarity hits compared to the RFG group ($M = 0.29$), $t(276.78) = -4.24, p < .001$.

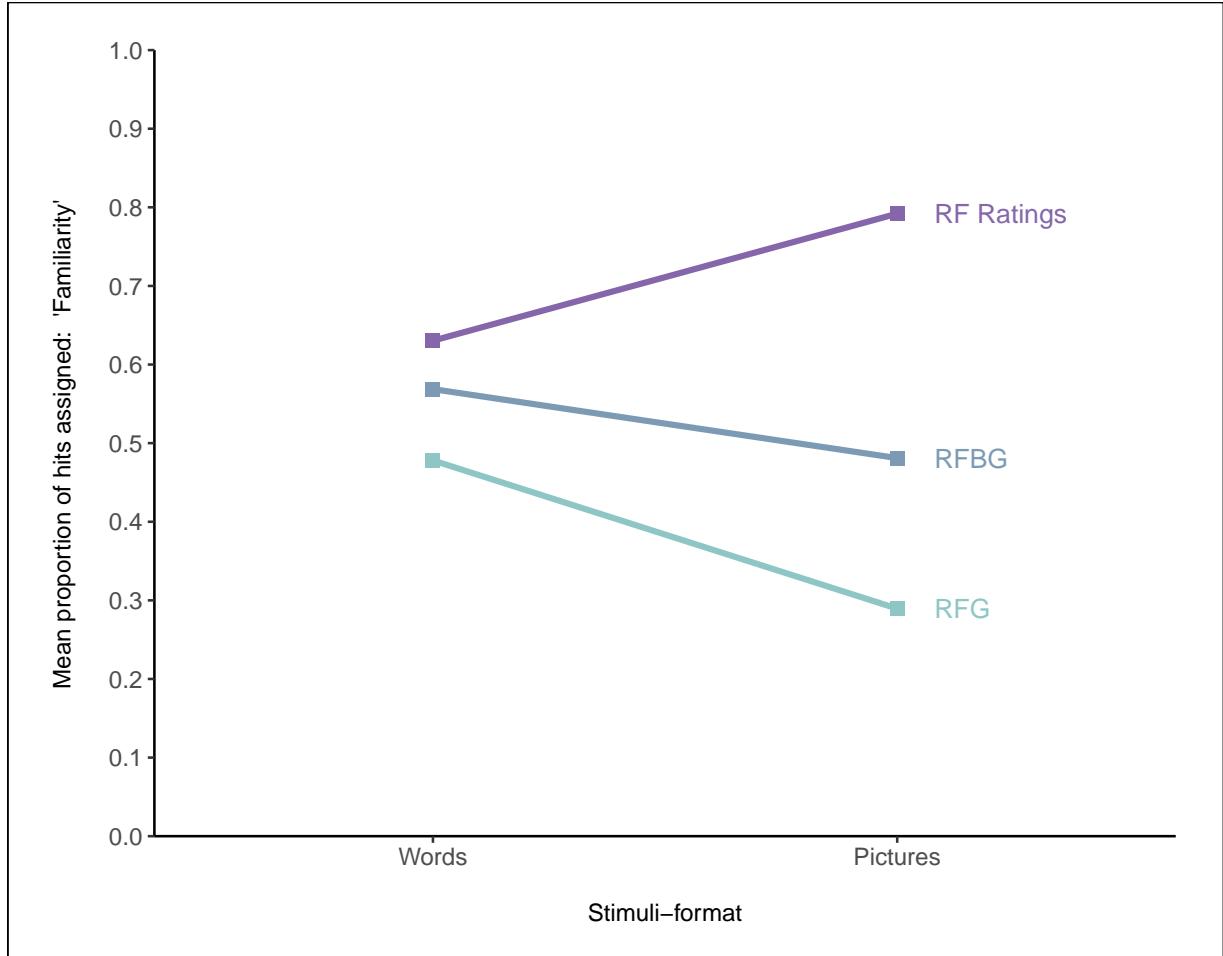


Figure 3: Interaction plot between stimuli-format and response-option condition for the mean proportion of hits assigned 'Familiarity'.

Guessing: The ANOVA on Guessing hits also showed a significant interaction between stimuli format and response option, $F(2, 178) = 4.17, MSE = 0.01, p = .017, \hat{\eta}_G^2 = .011$, (see Figure 4). Within response-option conditions, words resulted in more Guessing hits than pictures in both the RFG group (words: $M = 0.17$; pictures: $M = 0.08$), $t(178) = 5.38, p < .001$; and RFBG group (words: $M = 0.16$; pictures: $M = 0.09$), $t(178) = 4.42, p < .001$.

Between response-option conditions, word stimuli produced significantly fewer Guessing hits in the RF-Ratings group ($M = 0.04$) compared to both the RFG group ($M = 0.17$),

$t(281.42) = 5.44, p < .001$, and the RFBG group ($M = 0.16$), $t(281.42) = 5.18, p < .001$. A similar pattern was also evident for pictures, with significantly fewer of Guessing hits in the RF-Ratings group ($M = 0.01$) compared to both the RFG group ($M = 0.08$), $t(281.42) = 2.70, p = .020$ and RFBG group ($M = 0.09$), $t(281.42) = 3.15, p = .005$.

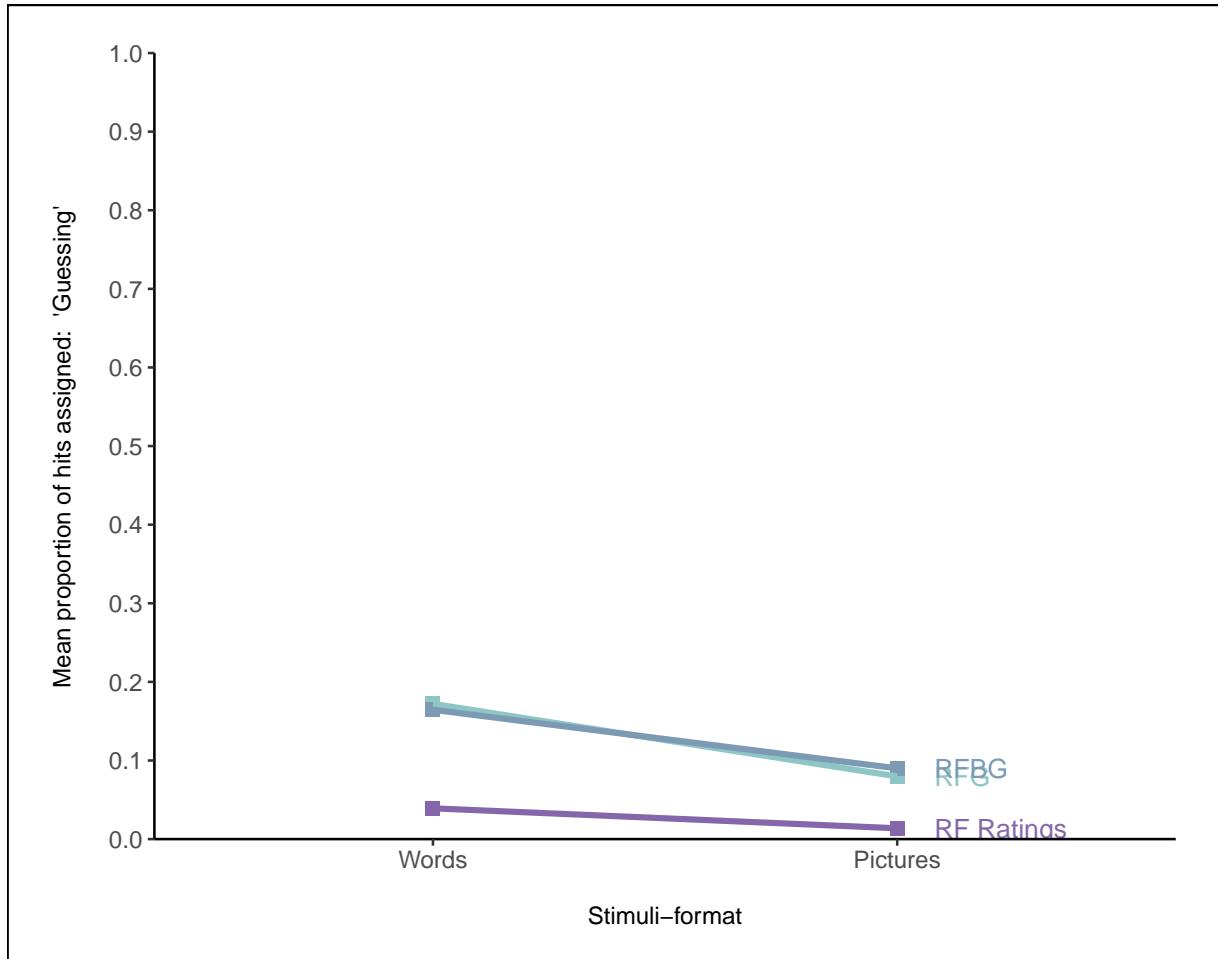


Figure 4: Interaction plot between stimuli-format and response-option condition for the mean proportion of hits assigned ‘Guessing’.

Finally, the ANOVA on Guessing FAs also showed no significant main effect of stimuli-format, $F(1, 136) = 0.51, MSE = 0.04, p = .476, \hat{\eta}_G^2 = .001$, but there was a significant main effect of response-option, $F(2, 136) = 15.69, MSE = 0.11, p < .001, \hat{\eta}_G^2 = .144$, with the RF-ratings group ($M = 0.08$) showing significantly fewer Guessing FAs than both the RFG ($M = 0.35$), , and RFBG groups ($M = 0.30$), . This again aligns with previous results,

suggesting that those in the RF-Ratings group were less likely to report guesses than the other groups, whether accurate or not. There were no significant interaction effects, $F(2, 136) = 0.07$, $MSE = 0.04$, $p = .935$, $\hat{\eta}_G^2 = .000$.

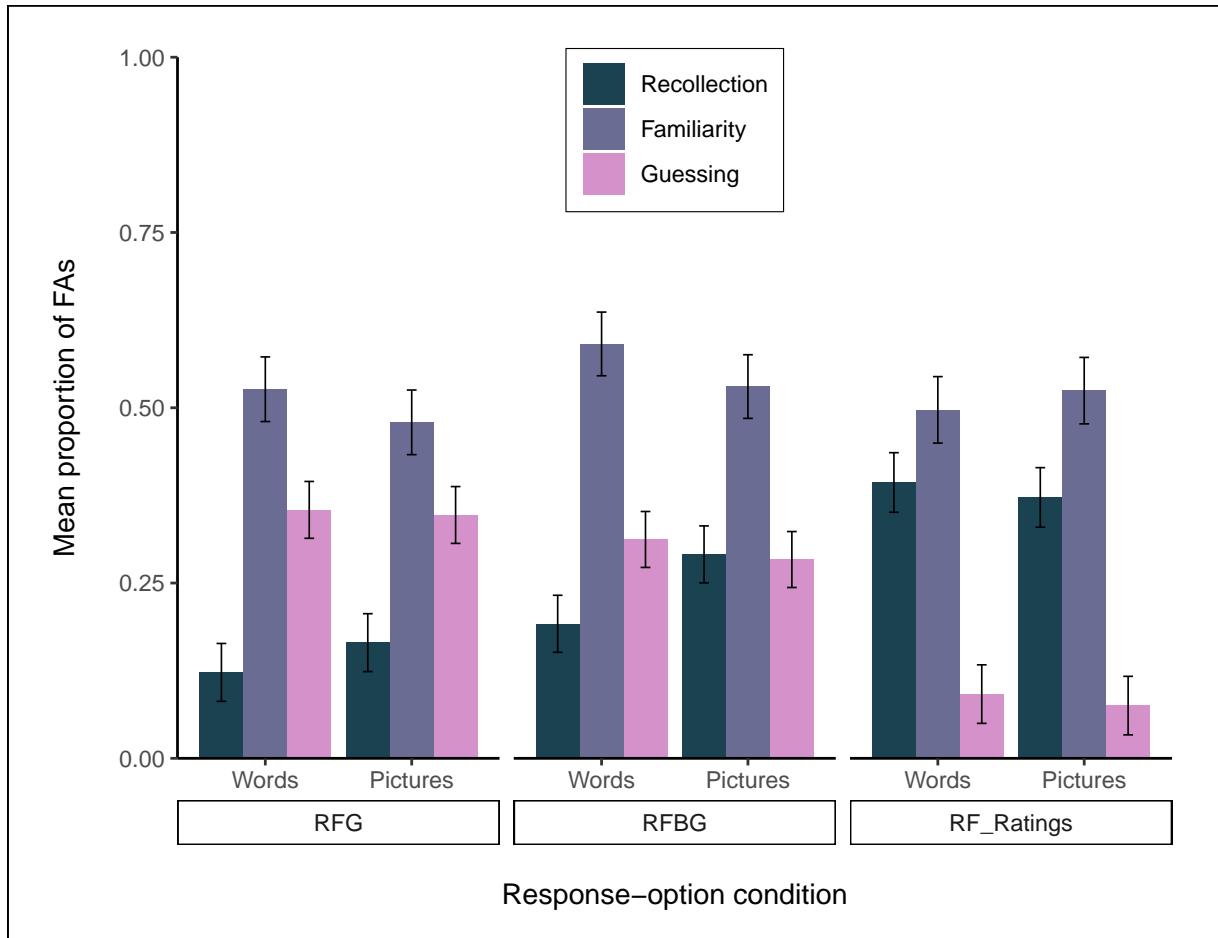


Figure 5: Proportion of FAs assigned Recollection, Familiarity, and Guessing, by stimuli-format and response-option condition.

Discussion

How do colour (vs. greyscale), and illustrations (vs. photographs) differentially affect responses across RFG, RFBG, and RF Ratings?

“perhaps the current study stimulus set of high-resolution colour images may have helped patients overcome visual processing deficits, and use distinctive information to improve memory for pictures over words” - (Ally, Gold, & Budson, 2009) – expected visual processing deficits to

manifest as reduced PSE in MCI, but as the results did not support this, suggest the colour images may have helped to bypass this. “Future studies can perhaps examine whether increasing the amount of detail present in a picture enhances discrimination”

Suzuki & Takahashi (1997): recognition performance was highest when colour pictures were used in both the study and test phases (compared with black and white).

Taken together, these findings with expected PSE are consistent with the notion that pictures offer an enhanced memorability in comparison to words. When they were correctly recognised, words were not recognised in the same context-rich nature as pictures, evidenced by a higher proportion of F responses. Despite ‘Guessing’ responses being permissible in any of the response-option groups, it seems when two independent rating scales were required participants were less likely to report a “Guess”. This could be because those in the RF-Ratings condition selected “New” more often than “Old” when they were having a complete guess, or instead, subjects might have opted to report lower levels of Recollection and Familiarity ratings (i.e., 1-3), rather than responding 0 on both scales.

#####-----

Chapter 3

The Picture Superiority Effect (PSE) is a highly robust and replicable phenomenon. In recognition memory paradigms, the PSE has been shown to manifest as both increased recollection and familiarity [@rajaram1993; @rajaram1996; @dewhurst1994; @wagner1997; @yonelinas2002]. The effect is present in children, adolescents and healthy older adults [@whitehouse2006], though perhaps more striking is the fact that patients with Alzheimer's disease or those presenting early isolated memory impairments, known as amnestic mild cognitive impairment (aMCI), also show memorial benefits toward pictures [@ally2012]. This is supported by ERP studies demonstrating comparable enhancements to recollection-based ERP components between healthy older and aMCI groups when pictures, rather than words, are utilised [@ally2009]. There is debate within the literature attempting to characterise the nature of memory deficits in aMCI, whereby despite general agreement that recollection processes are impaired in such individuals, findings show great inconsistency with regard to familiarity [@algarabel2012; @belleville2011; @pitarque2016; @wolk2013; @wolk2011]. The PSE may have been largely overlooked as an area for further research in an effort to help settle this debate, despite recent reviews highlighting methodological differences across studies as the potential source of inconsistent findings [@koen2014; @migo2012; @schoemaker2014]. The level at which stimuli distinctiveness impacts successful recognition is currently unclear, and there is little consistency across studies with regard to what is considered a 'picture'.

Many experiments utilise illustrations for their picture stimuli [@vandermeulen2012; @westerberg2013; @wolk2011], with a standardised set of items published by @snodgrass1980 among the most-used illustrated picture stimuli within the domain of memory research [@weldon1987; @weldon1989; @wagner1997; @mcbride2002; @martins2006; @whitehouse2006; @hockley2008; @schmitter-edgecombe2009; @vandermeulen2012; @deason2015; @wammes2016; @bermudez-margaretto2018; @meade2019]. The set consists of 260 line drawings of common, everyday objects (in black ink), along with their written word counterpart (e.g. "shoe"). Items were selected on the basis of exemplifying a number of semantic categories, including animals, furniture, fruit, etc., and a range of normative data

was collected for each item; indices of naming agreement, mental imagery agreement, visual complexity, and familiarity were all recorded for each drawing. The normative data for the @snodgrass1980 items has been continually revisited, with a number of studies gathering culturally-appropriate norms (e.g. in Spanish [@sanfeliu1996], Chinese [@yoon2004], and Russian [@tsaparina2011], and additional testing of the relationship between reaction time and naming agreement [@szekely2003]. There are multiple theories of object recognition; the recognition-by-components theory proposed by @biederman1987 identifies shape as the most crucial factor for successful recognition, in which case, the object outlines found in the set by @snodgrass1980 should be more than sufficient for experimental cognitive research. Other theories, however, posit that surface details such as colour and texture are just as crucial in forming object representations [@tanaka2001; @tarr1998]. The wide-ranging applicability of the @snodgrass1980 items throughout a number of cognitive disciplines has led to a more recent revision of the items by @rossion2004. This revision consists of the exact same objects, digitally re-drawn to include surface textures and shading. Additionally, this set provides greyscale and colour versions for all items, as opposed to the greyscale-only items found in the @snodgrass1980 set (see Figure 6 for example items contained in the @snodgrass1980 and @rossion2004 stimuli sets). The @rossion2004 revision now appears to be favoured over the original @snodgrass1980 set among many cognitive researchers [@stenberg2006; @wolk2008; @rollins2018, @ensor2019b], almost certainly attributable to the increased detail and ability to choose whether colour is a necessary condition.

Despite their widespread use, line drawings have been criticised for their relative simplicity and lack of realism (@viggiano2004), with many researchers favouring the use of photographs as experimental stimuli [@embree2012; @pitarque2016; @troyer2012; @troyer2016; @wang2013a]. Photographs of faces are especially useful in research examining emotion and face recognition [@bowen2019; @cui2016; @herzmann2018; @barba1997], though a number of common-object photograph sets have also emerged as ecological alternatives to line-drawn items [@viggiano2004; @moreno-martinez2012; @adlington2009]. While the published sets of photographs are undoubtedly useful in a range of cognitive domains, they

do not allow us to specifically examine stimuli format as a factor on its own, as the concepts depicted are unique to the set they derive from. In order to make such comparisons, and ensure any differences in performance (e.g. recognition memory ability) are indeed attributable to stimuli format, the objects depicted must be consistent across stimuli formats. The current study presents a new set of photographic stimuli that extend the set of words and drawings provided by @rossion2004, wherein each of the concepts depicted has been carefully matched across formats. These new stimuli will be utilised throughout a number of planned recognition experiments that aim to systematically compare measures of recognition against different ‘levels’ of stimuli. The curation of a new set of photographs - carefully matched to other formats - allows investigation into whether picture superiority magnitudes are mediated by the format pictures are presented in. The inconsistent use of different formats across studies has previously made it difficult to reconcile effects obtained in response to drawings with those obtained in response to photographs - an inherent problem when concepts are not matched across format. Normative data for the new set of photographs is also presented, allowing others who also wish to use our photograph stimuli to filter items by measures of naming agreement, mental imagery agreement, familiarity, visual complexity, and colour diagnosticity.

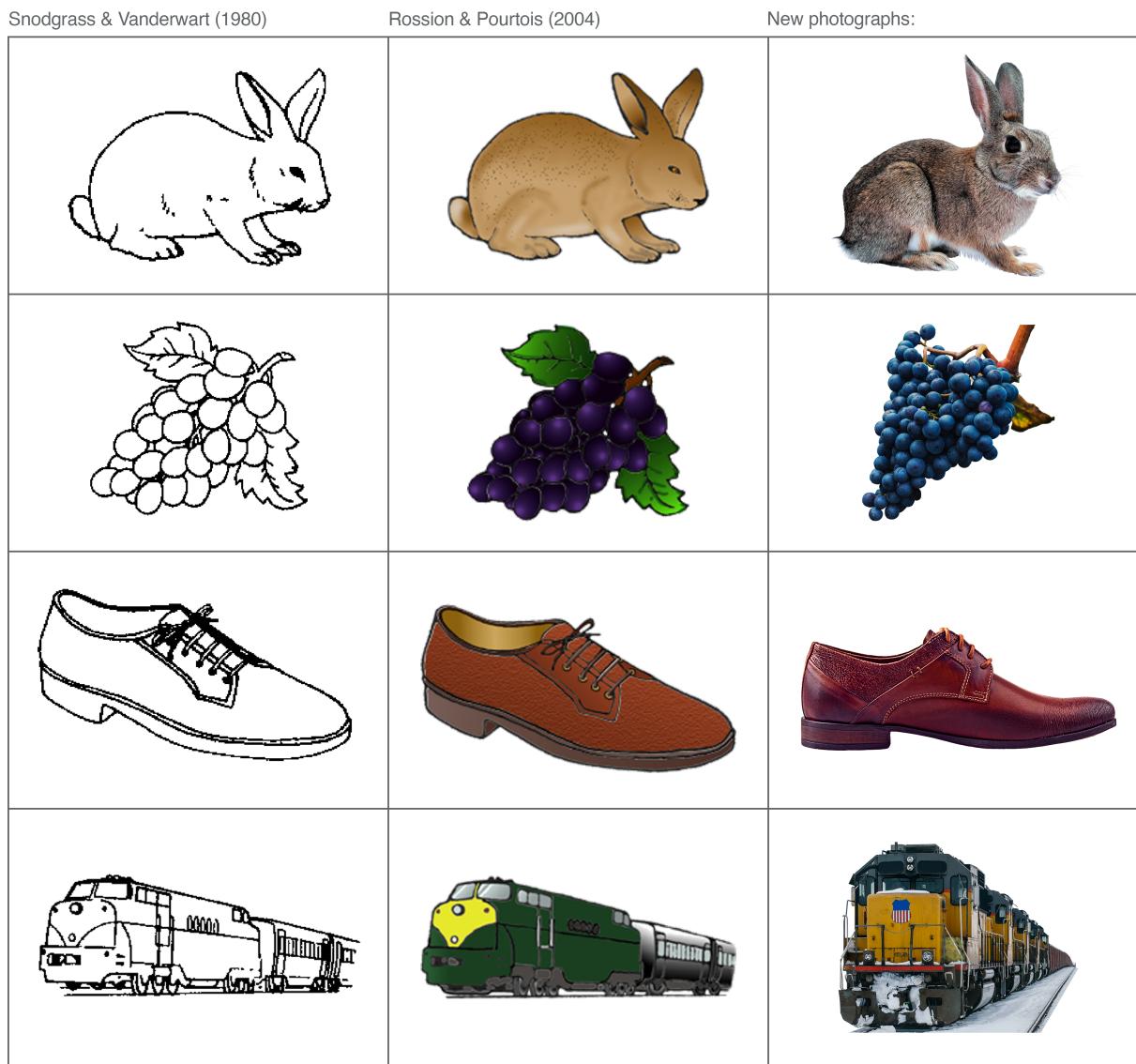


Figure 7: Examples of matching pictures across @snodgrass1980, @rossion2004, and photographs from the current study. Greyscale versions of the drawings and photographs are not presented in this example.

Experiment: Development of a new set of standardised photographic stimuli

Method

Participants A total of 377 subjects completed the online experiment (see Table 3 for a breakdown of the gender and age of the sample). This sample size provided 20 data points for each of the five response types, while also ensuring the experiment did not last too long for participants (approx 25-mins). Subjects were recruited from both voluntary participation websites such as Prolific Academic⁴ (where they received payment at the rate of £5/hr), and via the in-school research participation system⁵ (where they received course participation credits).

Table 3: Gender and age (*SD*) of the current sample.

Gender	N	Age	
Female	196	33.22	(11.28)
Male	171	33.15	(10.3)
Non-binary	2	23.50	(-)
Unspecified	5	29.40	(6.11)
Total	377	NA	NA

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, subjects were also asked whether English was their first language; all but one participant indicated that English was indeed their first language (99.2%).

Materials A pool of 136 line drawings [@rossion2004] - depicting common, everyday objects - were brought forward from the previous experiment. These items (along with their written-word labels) would form two of the unique stimuli formats that would be used in future recognition experiments (words and drawings). In this study, the drawings from @rossion2004 were simply used as a reference in the photograph matching process. Corresponding photographs were obtained online with the aim of depicting the everyday objects in a similar manner to the drawings. The inherent subjectivity involved in this process may have led to images that were not a reliable 'match' to the concepts they were selected to depict (for example, the photograph chosen

⁴<https://www.prolific.co/>

⁵<https://keelepsychology.sona-systems.com/>

to depict the concept “bottle” may inadvertently provoke the majority of participants to give the label “wine”, thus indicating that this particular photograph fails to accurately depict the intended concept). To address this issue, and ensure all photographs more objectively depict the same concepts as the line drawings, three different photograph variations were found for each everyday object, with the aim of taking the best ‘match’ forward. An emphasis was placed on variety across these variations, with the aim of obtaining at least one photograph that very closely resembled the line-drawn depiction, and another offering a more modern depiction. Some items were substituted 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. Unsplash⁶; Pexels⁷) for photographs that depicted the same everyday objects as the line drawings (see Appendix C for the full list of image references).

The matching process produced a total of 408 unique photographs. All were imported into Adobe Photoshop (20.0.04 Release), where the background was removed to isolate the object of interest from other potentially distracting visual details. This was completed manually using the magnetic lasso and polygonal lasso tools (edges were either feathered by 1px or left unfeathered). The orientation of isolated objects was adjusted to ensure they matched as closely as possible with their line-drawn counterpart (e.g. all photograph variations of the item ‘boot’ 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. For example, in one photograph variation of the item ‘piano’, there was a sign on the object that may have impacted how the item was named or rated. Such details were removed as best as possible using the clone stamp and content-aware fill tools. Any obvious text (e.g. brand names) and numbers were also removed from photographs using the same method (see Figure 8). The primary aim of the current study was to obtain photographs that could be clearly distinguished as a unique stimuli format among words and line

⁶<https://unsplash.com/>

⁷<https://www.pexels.com/>

drawings; it is conceivable that combining these formats (i.e. inadvertently including photographs that also contain written words) might affect recognition performance in ways that are not directly comparable to items defined only by a single category. Any text in our photographs was therefore removed, apart from a couple of exceptions whereby such details happened to be integral to the depiction of the object (e.g. the numbers found on a ruler or clock).

All 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).

Original



Manipulated



Figure 8: Examples of background and text removal in photograph items.

Design This was a descriptive study; a mix of qualitative and quantitative data were gathered. Across three blocks, all participants provided five types of response toward photograph 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 ordinal 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, the five types of required data were obtained 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) Qualtrics⁸ - a survey platform that allowed for straightforward collection of consent, demographics, and computer compatibility data, and ii) Pavlovia⁹ - an open-source experiment hosting platform for studies programmed in Javascript [@peirce2019].

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, familiarity was defined 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

⁸<https://www.qualtrics.com/uk/>

⁹<https://pavlovia.org/>

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. This concept was defined 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.

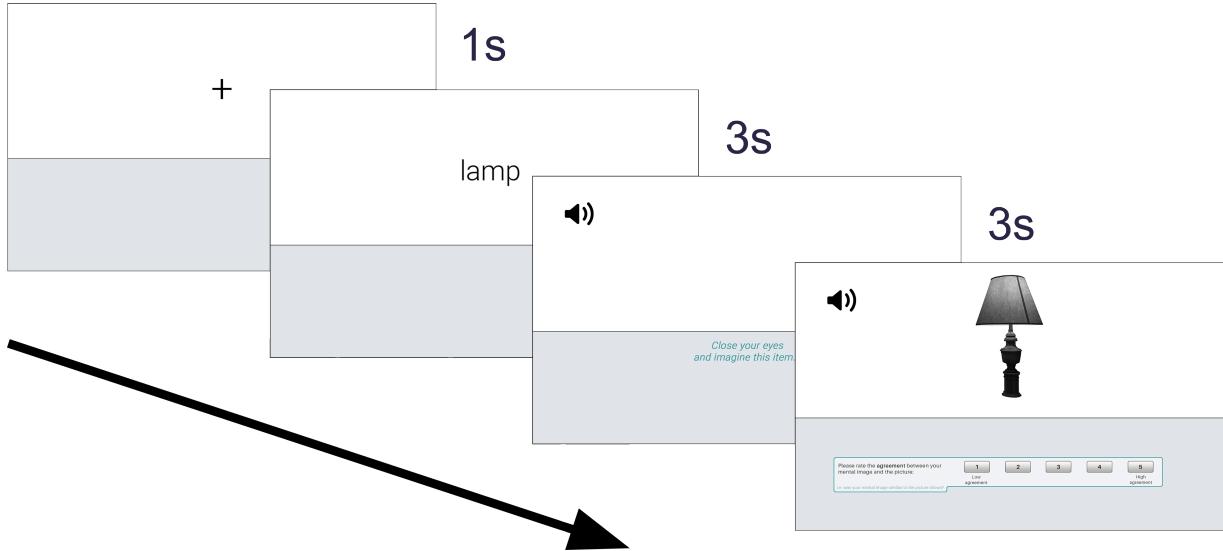


Figure 9: Data collection procedure for Mental Imagery Agreement responses.

Data processing The naming responses for each photograph item were manually assessed for spelling and typing errors. Automatic spell checking software was avoided in an effort to avoid inadvertently introducing unique names that were not actually given by participants. The vast majority of errors were unambiguous and easy to correct (e.g. “anker” = “anchor”, “peguin” = “penguin”, “ssnowman” = “snowman”), or consisted of transforming plural words to singular (or vice versa, depending on the form of the intended label - e.g. “sock” to “socks”). Some responses were a little more ambiguous, and necessitated comparison to the photographs they were in response to for additional clarity (e.g. a photograph depicting a plug that would fit into North American electrical sockets was labelled as “usplug” - given the nature of our UK-based sample, it’s likely the subject was responding: “U.S. (i.e. United States) plug”).

There were instances where subjects provided a sensible and correctly spelled English word, but that were clearly typos when examined against the photograph they were in response to (e.g. “dock” for a photograph depicting a duck, “frock” for a frog, and “beer” for a “bear”, etc). The most ambiguous spelling error to correct was “bittle”, which was provided by more than one

participant and to more than one item; separate inspections of the photographs participants were responding to made this easy to correct though, with one participant clearly meaning to respond “bottle”, whilst the other meant to respond “beetle”. Though participants were instructed to only give a single label for each item, some multiple word responses were found (without spaces) during the spell checking process. On such occasions, a judgement was made regarding whether multiple words were retained, or whether the response could be shortened into a single word. A general rule was applied whereby if the other words provided additional information, they were retained (e.g. “maledeer” - presumably “male deer” - was kept as a two-word answer). Multiple word responses were generally shortened into a single word when the intended label for the item was clearly present, and no information was lost in the process (e.g. “haircomb” was shortened to the intended answer “comb”). It is noted that there was some inherent subjectivity in this process, though as such items were not common among straightforward responses, their overall effects are estimated to be negligible.

Finally, there were some responses that were changed to “no” as they were clearly intended to signify that the responder did not know the name of the item shown; the experiment instructed participants to type “no” in these instances, though the labels “none” and “idk” (common abbreviation for “I don’t know) were provided instead. There was also a single response that was manually changed to “no”, as the provided label was a single letter and thus entirely unclear what the intended answer should be (see Appendix B for full list of manipulations to naming responses). This process yielded data that could be used to determine which photograph variation best matched the intended concepts (e.g. 100% of participants labelled the object “bottle”, indicating a perfect match), and which did not (e.g. only 50% of participants labelled the item “bottle”, whilst the other 50% gave the label “wine”, indicating a poor match). Photographs showing poor agreement across participant-generated labels, or those where the majority of labels differed from the intended concept, could be replaced with the variation demonstrating the most accurate depiction.

Analysis preparation A number of variables were calculated prior to analysis. For familiarity, visual complexity, colour diagnosticity, and mental imagery agreement, mean ratings were calculated for each (see Appendix C). Mean reaction times (RTs) were also calculated for each photograph / response variable, including naming responses. For naming responses, accuracy was defined as the proportion of subjects reporting the correct/intended label for any given item (e.g. 80% of subjects correctly labelled a photograph of the moon as “moon”). Percentage agreement was also calculated (i.e. the proportion of subjects providing the most frequent name, regardless of whether it matched the correct/intended label) in order to compute H values for each item. The H statistic also reflects naming agreement, but it takes into account the total number of unique labels given for an item. This is especially useful for comparing similar items, as it captures information not provided by simple agreement proportions. For instance, if the first variation of the photo moon ('moon-1') demonstrated 90% naming agreement among subjects, and the second variation ('moon-2') also demonstrated 90% naming agreement, it would appear as if both versions offer the same level of agreement among participants. However, 'moon-1' may have received a total of 2 unique names (e.g. moon, planet), while 'moon-2' received a total of 4 unique names (e.g. moon, planet, earth, comet). H values utilise this useful information to determine which item shows the best naming agreement (in other words, the item with the least number of unique names). The original formula by @snodgrass1980 was used to calculate H values:

$$H = \sum_{i=1}^k p_i \log_2 \frac{1}{p_i},$$

A H value of 0 indicates perfect naming agreement (all subjects responded with the same label for that item). Items showing a H value of 1 signify two unique names were provided, with identical proportions (e.g. 10 subjects responded “moon” and 10 subjects responded “planet”). As the H value increases, overall naming agreement decreases.

Results

Summary statistics (mean and *SD*) for each of the measured variables are shown in Table 4. Data for the grey and colour photographs are presented alongside previously obtained normative values for a number of other stimuli formats (all obtained from @rossion2004, who published revised norms for @snodgrass1980's (S&V) original line drawings, as well as their own re-drawn versions that contained shading and texture detail). The data from previous studies were not used in any statistical analyses. To examine whether the grey and colour photographs from the current study demonstrated any differences, a series of independent samples t-tests were run on each variable, as well as their corresponding reaction times (excluding scores of colour diagnosticity, which were obtained only in response to the colour items and thus cannot be compared). Mean (and *SD*) values for all x816 unique photograph items are presented in Appendix C.

Naming Naming accuracy was very high for all photographs ($M = 0.95$), indicating that overall, the selected items closely depicted the intended concepts. Compared with the other stimuli formats, there appears to be a steady increase in accuracy as items become more distinctive (see Table 4). Accuracy rates did not differ between the grey ($M = 0.94$) and colour ($M = 0.95$) versions of the photographs [$t(745.64) = -0.56, p = .576$].

H values were also low across all items ($M = 0.23$), showing that subjects generally agreed on how the items should be named. Similar to naming accuracy, naming agreement also appears to steadily increase as items become more distinctive (as indicated by decreasing *H* values - see Table 4. While @rossion2004 observed significantly better naming agreement for their colour - rather than greyscale - items, this pattern did not reach significance with the current set of photographs; *H* values did not differ between the grey ($M = 0.24$) and colour ($M = 0.22$) photographs [$t(743.66) = 0.62, p = .537$].

A mean reaction time (RT) of (3.9s) was observed for naming responses. While this was of little interest on its own, and could not be compared to those obtained in response to the other stimuli formats as our methodology was slightly different (RTs were only recorded when subjects had

typed their response *and* clicked the mouse to signify they had finished), they were useful for marking comparisons between the grey and colour items (though no difference was observed [M grey = 4s, M colour = 3.8s, $t(651.86) = 1.57$, $p = .117$]). Overall, these analyses suggest that the current photographs closely resemble the drawings they were designed to match, with high levels of naming accuracy and agreement among subjects. The absence of any colour differences indicates there were no naming advantages when photographs were made even more distinctive through the addition of colour.

Table 4: Summary statistics for each of the measured variables. Mean values are presented in bold (SDs are shown in parentheses).

	Rossion & Pourtois (2004)			Current study	
	S&V lines	Grey shaded	Colour shaded	Grey photos	Colour photos
Naming accuracy	88.2 (17.1)	89.2 (17.2)	90.3 (16.9)	0.94 (0.08)	0.95 (0.08)
Naming agreement (H)	0.44 (0.56)	0.38 (0.52)	0.32 (0.46)	0.24 (0.33)	0.22 (0.31)
Mental imagery agreement	3.73 (0.48)	3.76 (0.55)	3.74 (0.63)	3.46 (0.56)	3.74 (0.65)
Familiarity	3.59 (0.94)	3.52 (1.01)	3.44 (1.01)	4.13 (0.56)	4.19 (0.54)
Visual complexity	2.76 (1.03)	2.88 (1.03)	2.7 (0.94)	2.87 (0.62)	3.16 (0.63)
Colour diagnosticity	-	-	-	-	3.22 (0.84)

Mental imagery agreement Scores of mental imagery agreement were moderate across all items ($M = 3.6$). While no colour differences were previously observed between stimuli formats, the grey ($M = 3.46$) photographs in the current study showed significantly lower mental imagery agreement scores than the colour ($M = 3.74$) items [$t(800.06) = -6.54$, $p < .001$].

Comparisons with previous normative data also highlight how the grey photographs exhibited uniquely poorer mental imagery agreement scores than any of the other stimuli formats (see Table 4). RTs between the grey ($M = 3.04$) and colour ($M = 2.81$) items did not significantly differ [$t(571.37) = 2.14, p = .033$].

Familiarity Familiarity scores were high overall ($M = 4.16$), and like previous findings, there was no difference between the grey ($M = 4.13$) and colour ($M = 4.19$) items [$t(813.19) = -1.63, p = .103$]. However, familiarity scores for the current set of photographs were higher than those obtained for any of the other stimuli formats, and while there previously appeared to be a decline in familiarity as stimuli become more distinctive (from line drawings, to grey shaded, to colour shaded), such a pattern was not evident with the current photographs (see Table 4). RTs between the grey ($M = 0.97$) and colour ($M = 0.98$) items did not significantly differ [$t(783.66) = -0.30, p = .762$].

Visual complexity Visual complexity ratings were moderate across all of the items ($M = 3.3$). Colour ($M = 3.16$) photographs showed significantly higher scores of visual complexity than grey ($M = 2.87$) photographs [$t(813.51) = -6.65, p < .001$]. This finding is further demonstrated when compared to the scores from the other stimuli formats (see Table 4); where grey photographs show comparable levels of visual complexity, the colour photographs show higher scores than all of the other formats. There was no significant difference between the RTs of grey ($M = 3.26$) and colour ($M = 3.35$) items [$t(754.08) = -1.21, p = .228$].

Selection of final items For each concept represented in the photographs, one variation (e.g. shoe-1, shoe-2, or shoe-3) was selected for inclusion in a final list of stimuli that would be taken forward into subsequent recognition experiments. The normative naming data was assessed to establish which version best matched the existing line-drawn depictions of the concepts [@rossion2004]. Naming was favoured over all of the other variables as, if an item was found to primarily convey a different concept than was intended during the naming task (e.g. if a

photograph of the fruit ‘orange’ was labelled ‘grapefruit’ by the majority of subjects), then it could not be sufficiently compared to its line-drawn (and written-word) counterpart during recognition studies.

At least 20 unique naming responses were collected for each of the 816 photographs (408 grey items and 408 colour items). The proportion of ‘correct’ responses (i.e. names that were congruent with the intended concept) and the proportion of ‘don’t know’ responses were calculated for each item. Photographs were excluded if they:

1. received a high proportion of “don’t know” responses (20%; all of the photographs depicted common, everyday objects, and so if a number of subjects were unable to name the item, that particular photograph was considered to be a poor representation of the item);
2. were incorrectly named by the majority of subjects (i.e. if the proportion of correct responses equalled $\leq 50\%$, since it was essential for the photographs to depict the same concepts as those found in the line drawings and word stimuli);
3. had particularly poor naming agreement ($\leq 20\%$ subjects named the object similarly). Items may not have been flagged by the second criteria (e.g. if it received 4 different names, each with a 25% ratio), but could still be considered poor representations of the intended concepts.

54 photographs were found to meet at least one of the above criteria, and therefore excluded. Regardless of whether these items were grey or colour, it was also necessary to remove its grey or colour partner (since both versions were needed to make comparisons across recognition experiments). Thus, a total of 64 items (32 grey / 32 colour) were excluded at this stage (many items already had both grey and colour versions flagged by the original criteria).

Next, the proportion of correct responses were compared between grey and colour photographs in order to identify items showing the lowest difference. In order to manipulate colour in later recognition experiments, it was important to select items where naming was congruent across colour/grey items; in other words, it would be difficult to attribute particular recognition response patterns to the addition of colour (if a difference were found) when the grey version could not

be identified (or encoded) similarly. Variations exhibiting the least difference between colour and grey items (for the proportion of correct responses) were taken forward, while the rest were excluded. In a number of instances, multiple variations for the same object had the same ‘difference’ score. For example, all three variations of the item “balloon” exhibited perfect naming agreement, irrespective of whether they were presented in colour or grey (and thus “balloon1”, “balloon2”, and “balloon3” had a difference score of 0). For items where more than 1 variation remained, manual rankings were obtained from two of the researchers to determine which variation best depicted the intended concept. For each item, the researchers independently studied the remaining variations and provided a rank of which they thought was best (1) to worst (2 or 3, depending on the number of variations that remained). The ratings from both researchers were collated; items where there was agreement as to which variation best depicted the intended concept were selected for inclusion in the final stimuli list. For all the items where there was disagreement between the researchers rankings, one of the variations was simply selected at random.

Discussion

The role of colour For naming responses (accuracy, agreement [H], and RTs), no differences were observed between the grey and colour photographs. Such a result was expected for accuracy and agreement scores; the addition/absence of colour should not alter how participants identify (and thus label) items, except in rare instances whereby a lack of colour may lead to the misidentification of an object (e.g. incorrectly labelling a greyscale photograph of an orange as ‘grapefruit’). The data indicates, however, that this was not common, with the grey set of photographs exhibiting equally high levels of naming accuracy as the colour photographs. The absence of RT differences between the colour and greyscale sets was not expected for naming responses. It is reasonable to assume that colour photographs - with an additional layer of contextual information compared to grey items - would be identified (and therefore named) quicker than grey photographs (e.g. a colour photograph of an orange should avoid the potential ambiguity that might accompany a greyscale depiction, which could initially be confused for an-

other type of fruit). Indeed, @rossion2004 demonstrated RTs consistent with this hypotheses, with colour drawings showing significantly quicker RTs than grey items. The lack of difference in the current data could be attributable to ceiling effects, whereby all photographs were sufficiently unambiguous, and were quickly identified irrespective of whether they were presented in greyscale or colour. Examination of the other naming data, showing similarly high levels of accuracy and agreement across grey and colour, supports this notion.

Scores of mental imagery agreement produced particularly interesting results between the grey and colour items. Grey photographs exhibited a significantly poorer match with subjects imagined presentation of the objects than the colour items. Colour differences were not observed previously between drawings [@rossion2004], and comparing the current data with that obtained in other studies (see Table 4) demonstrates how the greyscale photographs show uniquely lower mental imagery agreement scores compared with any of the other stimuli formats. To imagine the objects, it seems likely that subjects would conjure an image of how they naturally see the item in their everyday lives - which for the majority of subjects, would presumably be a colour representation. Therefore, when presented with greyscale depictions, subjects may have been more inclined to report that that item did not align quite as well as those presented in colour. However, it is unclear why a similar pattern is not also evident when comparing grey and colour drawings [@rossion2004]. It may be that photographs promote stricter internal criteria when subjects must decide whether an item is a good match to their mental image. With line-drawn / illustrated items, subjects may simply accept that the items are baseline depictions, and that they will only able to match their real-world mental images to a certain degree - thus leading to a generally more liberal response bias throughout. The addition of colour may therefore do very little to further reconcile the match between the drawing and real-world mental representation. When subjects are responding only to photographs, the ecological nature of the items may facilitate deeper critical evaluation of whether they offer a good match to mental images, and thus promote a more conservative response bias. Colour may therefore be a far more important factor in photographs than it is in line drawings for allowing participants to decide whether an item matches well with their mental image.

There were no colour differences in familiarity scores. This result was expected - participants were asked to rate the degree to which they came in contact with, or think about, the concept itself rather than the particular depiction shown, and there is no apparent reason why colour should influence such ratings. Visual complexity, on the other hand, where participants were required to directly rate the amount of detail in the picture, did show an expected difference. Colour photographs were rated as significantly more visually complex than grey items, presumably due to their additional layer of contextual information. When compared to the previous data obtained for drawings, the greyscale photographs showed comparable levels of visual complexity, while the colour photographs showed higher levels than any of the other formats. It is unclear why the photographs of the current study showed colour differences, when grey and colour drawings did not differ, though it may tie in with the hypotheses proposed to explain the mental imagery agreement data. Subjects may apply stricter internal criteria when rating stimuli that are perceived as being closer to how they would be experienced in real life - when viewing a colour photograph of a rabbit, it is difficult to see how we could make the item any more visually complex than it already is (at least in a 2D medium). It's probable that subjects notice the absence of colour when viewing the greyscale items, since they depict the items in a way that they are not usually seen, and thus determine that these items could be made more complex if they were shown in colour (and so give lower visual complexity ratings as a result).

Establishing a new set of stimuli The objective of the current study was to establish a new set of ecological photograph stimuli to be taken forward into subsequent recognition memory experiments. Matching items with previously established drawings (and words) would allow for the effects of stimuli-format on recognition response patterns to be directly examined. A range of normative data was collected for 816 unique photograph items. These items may prove useful for a range of cognitive researchers that wish to utilise a set of high quality and realistic object stimuli, especially given the flexibility of items that can be filtered based on colour, naming agreement, familiarity, etc. For the needs of the current body of research, the naming data was used to determine which photographs best matched the intended concepts among a number of

possible variations. This allowed for the systematic comparison of recognition memory performance toward three distinct stimuli formats (words, drawings, and photographs) in the following study, in an effort to establish how stimuli of varying perceptual distinctiveness may affect recognition response patterns. Such comparisons might help to reconcile the inconsistencies present across recognition memory research, such as those attempted to determine whether familiarity processes are preserved in those with amnestic Mild Cognitive Impairment (aMCI).

Experiment: Effect of stimuli format (greyscale) and response option on recognition memory judgements.

For the recognition memory experiment, everyday objects were presented in three stimuli formats: i) words (written in simple, black ink); ii) drawings (shaded line-drawn illustrations); and iii) photographs (detail rich exemplars of the real world object). @rossion2004 demonstrated that naming agreement could be improved by adding surface texture and shading to the original @snodgrass1980 items; however, it is unclear how manipulations to distinctiveness actually impact performance in recognition memory paradigms. As well as general inconsistencies regarding the type of stimuli used in recognition memory experiments, there is also much variability in the response options available to participants when reporting their recognition, for example: Remember/Know (@lombardi2016), Recollection/Familiarity (???), or Low/Med/High confidence (????). In the current experiment, the availability of different response options when reporting recognition will also be examined by randomly assigning participants into a paradigm with three response options (Recollection / Familiarity / Guessing) or four response options (RFG + Both).

Based on the results of Experiment 1, which compared recognition to for words and drawings only, a number of hypotheses are proposed as to the potential effects of adding a third stimuli format (highly distinctive photograph stimuli). As stimuli become increasingly distinctive (from words, to drawings, to photographs), it seems likely that the number of hits (correctly recognised items) will increase, and the number of false alarms (FAs) will decrease. RFG responses are expected to show a similar pattern, with the most detailed stimuli showing the highest number of hits assigned “Recollection”, while the less detailed formats show increasing levels of “Famil-

arity” and “Guessing” hits. Whilst we expect the overall number of FAs to increase as stimuli become less distinctive (i.e. words will show the highest rate of FAs), there is no reason to believe that these FAs will be biased toward any particular RFG judgement across formats. It is also hypothesised that the rates of reported Recollection and Familiarity will differ across response option conditions (RFG / RFBG), though the direction of this difference is currently unclear.

Method

Participants A total of 158 subjects completed the online experiment (see Table 5 for a breakdown of the gender and age of the sample). To meet our YA requirements, all participants were required to be between 18-59 years of age (actual range: 18-58). As our experiment involved English word stimuli, we also asked subjects whether English was their first language; the vast majority (95.57%) reported that English was indeed their first language. Subjects were recruited from voluntary participation websites such as Prolific Academic¹⁰ (72.78%), where payment at the rate of £5/hr was given, and via the in-school research participation system¹¹ (15.19%), where they received course participation credits. A small number of participants were also recruited from Psychological Research on the Net¹² (12.03%). In order to detect a medium effect size of Cohen’s $f = 0.25$ with 80% power ($\alpha = .05$, two-tailed), GPower indicated that we would need 79 participants per group ($N^* = 158$) in a 3x2 mixed ANOVA.

Table 5: Gender and age (SD) of the current sample.

Gender	N	Age	
Female	96	29.53	(10.18)
Male	58	31.36	(11.19)
Questioning	1	21.00	(0)
Unspecified	3	50.33	(4.93)
Total	158	30.54	(10.84)

¹⁰<https://www.prolific.co/>

¹¹<https://keelepsychology.sona-systems.com/>

¹²<https://psych.hanover.edu/research/exponnet.html>

Materials A total of 126 innocuous, everyday objects (e.g. clock, rabbit, shoe) were presented across three individual stimuli formats: written words, line drawings, and photographs. The line drawings were obtained from @rossion2004, and consisted of greyscale shaded illustrations that contained some surface details. The word stimuli were simply the written word names of the line-drawn objects, presented in a clear Sans-serif typeface. The photograph stimuli were curated in the previous study; high quality photographs were sourced to similarly depict the same everyday objects as the line drawings. All objects in the photographs were isolated from their original background, converted to greyscale, and rotated to match the orientations shown in the line-drawn items.

Design The current study utilised a mixed design, with a 3-level within-subjects factor of stimuli format (words, drawings, photographs), and a 2-level between-subjects factor of response option (RFG, RFBG). Subjects passed through 2 levels of blocked randomization (equally sized, predetermined blocks); first, participants were randomly assigned one of six possible study lists (of equal length, and containing an even number of word, drawing, and photograph items) for counterbalancing purposes. Subjects were then either assigned into a recognition test with three possible response options (RFG: “Recollection”, “Familiarity”, “Guessing”), or four possible response options (RFBG: “Recollection”, “Familiarity”, “Guessing”, “Both”). These randomisation processes were completed automatically by the experiment software using balanced methods.

Words:	Drawings:	Photographs:
guitar		
mouse		
pumpkin		

Figure 10: Example stimuli from the three formats.

Procedure Data was collected online using Gorilla¹³ - a platform for the building and hosting of online experiments. The experiment consisted of three self-paced phases: i) study phase, ii) distractor task, and iii) recognition test. In the study phase, an even mix of word, drawing, and photograph stimuli were presented one-at-a-time on the computer screen. Subjects were instructed to learn the items in preparation for a later memory test. To ensure attention was directed to the presented stimuli, subjects were required to report whether each item was shown as a word, drawing, or photograph using the computer mouse. Following the study phase, participants completed some simple multiple choice mathematical questions (e.g. $6 \times 4 = ?$) as a distractor. Finally, participants' memory of the previously studied items was tested in the recognition task. An even mix of word, drawing, and photograph stimuli were again presented

¹³<https://gorilla.sc/>

one-at-a-time on the screen; half of the test items had been shown previously in the study phase, while the other half were new (and were not on the study list). For each item, subjects were instructed to press “Old” if they believed it was an item they had studied earlier, and “New” if they had not. “Old” responses led to a follow-up judgement, where participants reported whether they had experienced recognition through “Recollection”, “Familiarity”, or were simply taking an uninformed “Guess”. Participants that had been randomised into the RFBG test condition had a fourth option here, whereby they could report that they had experienced Recollection and Familiarity simultaneously (“Both”). Stimuli format was congruent across the study and test blocks (e.g. items presented as photos at study were also presented as photos at test). For each concept depicted across the three stimuli formats, subjects were only presented with one variation (in other words, if a subject saw a photograph for the item “shoe”, they did not see the word or line-drawn version of “shoe”).

Data processing Measured variables included the total number of hits and FAs, and the total number of hits and FAs assigned to each of the available response options (R/F/G and R/F/B/G). In order to create a common dependant variable, proportions were calculated from these variables slightly differently depending on the response option group. In the RFG-judgement group, simple proportions were created from the total number of R responses and the total number of F responses. In the RFBG condition, however, the proportion of Both responses was separately added to R proportions and F proportions. Additional DVs included: i) d' (d-prime, a signal detection measure of sensitivity); ii) c-value (a measure of response bias); iii) overall accuracy (hits / (hits + FAs)); iv) reaction times for all responses.

Participants were excluded from analysis if they showed poor performance during the encoding task; the relative ease of reporting whether each item was shown as a word, drawing, or photograph prompted a performance cut off of 90% accuracy. This allowed for some accidental clicks / incorrect responses toward potentially ambiguous items, though subjects scoring less than 90% were excluded on the assumption they did not dedicate their full attention to the task. Subjects with extreme z-scores were also excluded from analysis; those presenting z-scores of

+/- 3 (for total hits, total FAs, or overall recognition [hits minus FAs]) were considered outliers. These criteria resulted in the exclusion of 11 datasets.

Results

A series of 3x2 repeated measures ANOVAs were conducted on each of the DVs using a within-subjects factor of stimuli format (photographs / drawings / words) and a between-subjects factor of response option (RFG / RFBG). Significant main effects and interaction effects were followed-up with Bonferroni-adjusted pairwise comparisons.

Overall hits, false alarms (FAs), and recognition Separate 3 (stimuli format: words, drawings, photographs) x 2 (response option condition: RFG-judgements, RFBG-judgements) mixed ANOVAs were conducted on the mean proportion of hits and false alarms (FAs; see Table 6).

Table 6: Mean proportion of hits and FAs by stimuli format and response option condition.

	Hits	FAs
Stimuli format		
Words	0.55	0.21
Drawings	0.76	0.09
Photographs	0.86	0.05
Response option		
RFG	0.78	0.13
RFBG	0.74	0.11

For the proportion of total hits, there was a significant main effect of stimuli format [$F(1.75, 273.58) = 229.89, MSE = 0.02, p < .001, \hat{\eta}_G^2 = .346$]. The interaction effect was not significant [$F(1.75, 273.58) = 0.74, MSE = 0.02, p = .461, \hat{\eta}_G^2 = .002$]. Post-hoc comparisons for the main effect of stimuli format showed that photographs ($M = 0.86$) produced a significantly higher proportion of hits than both words ($M = 0.55$) [] and drawings ($M = 0.76$) []. Drawings ($M = 0.76$) also produced a significantly higher proportion of hits compared to words ($M = 0.55$) [].

There was also a significant main effect of stimuli format for the proportion of total FAs [$F(1.46, 227.29) = 106.64, MSE = 0.01, p < .001, \hat{\eta}_G^2 = .243$]; photographs ($M = 0.05$) produced a significantly lower proportion of FAs than both words (rexp3_apa_grey_anova_PROP.FA_words.mean) [] and drawings ($M = 0.09$) []; drawings ($M = 0.09$) produced a significantly lower proportion of FAs compared to words ($M = 0.21$). The interaction effect was not significant [$F(1.46, 227.29) = 1.22, MSE = 0.01, p = .287, \hat{\eta}_G^2 = .004$].

For overall performance accuracy, there was a significant main effect of stimuli format [$F(1.93, 300.97) = 586.13, MSE = 0.02, p < .001, \hat{\eta}_G^2 = .549$]. The interaction effect was not significant [$F(1.93, 300.97) = 2.02, MSE = 0.02, p = .136, \hat{\eta}_G^2 = .004$]. Post-hoc comparisons for the main effect of stimuli format showed that photographs ($M = 0.81$) produced significantly better performance accuracy than both words ($M = 0.34$) [] and draw-

ings ($M = 0.67$). Drawings ($M = 0.67$) also produced significantly better performance accuracy compared to words ($M = 0.34$).

Discrimination (d') and response bias (c) To assess the roles of discrimination and response bias, separate 3 (stimuli format: words, drawings, photographs) x 2 (response option condition: RFG-judgements, RFBG-judgements) mixed ANOVAs were conducted on the values of d' (d-prime; measure of sensitivity) and c (decision criterion; see Table 7).

Table 7: d' and c values by stimuli format and response option condition.

	d'	c
Stimuli format		
Words	1.06	0.39
Drawings	2.18	0.32
Photographs	2.78	0.22
Response option		
RFG	0.10	0.01
RFBG	0.10	0.01

For d' scores, there was a significant interaction between stimuli format and response option [$F(2.00, 311.73) = 3.60$, $MSE = 0.26$, $p = .029$, $\hat{\eta}_G^2 = .008$] (see Figure 11). Photographs ($M = 2.86$) facilitated better discrimination between hits / FAs than both words ($M = 1.13$) [$t(312.00) = -21.66$, $p < .001$] and drawings ($M = 2.13$) [$t(312) = -9.13$, $p < .001$] in the RFG group. Drawings ($M = 2.13$) also showed significantly higher d' scores compared to words ($M = 1.13$) [$t(312) = -12.53$, $p < .001$]. In the RFBG group, the same pattern was evident; photographs ($M = 2.69$) facilitated better discrimination than both words ($M = 0.98$) [$t(312) = -20.79$, $p < .001$] and drawings ($M = 2.24$) [$t(312) = -5.53$, $p < .001$]. Again, drawings ($M = 2.24$) also showed significantly higher d' scores than words ($M = 0.98$) [$t(319.09) = -10.22$, $p < .001$].

Comparisons of stimuli format across the response option groups showed that d' scores for photographs did not significantly differ between the RFG ($M = 2.86$) and RFBG ($M = 2.69$)

conditions [$t(319.09) = 1.43, p > .999$], nor for drawings (RFG: $M = 2.13$, RFBG: $M = 2.24$, [$t(319.09) = -0.96, p > .999$]), nor for words (RFG: $M = 1.13$, RFBG: $M = 0.98$, [$t(319.09) = 1.39, p > .999$]).

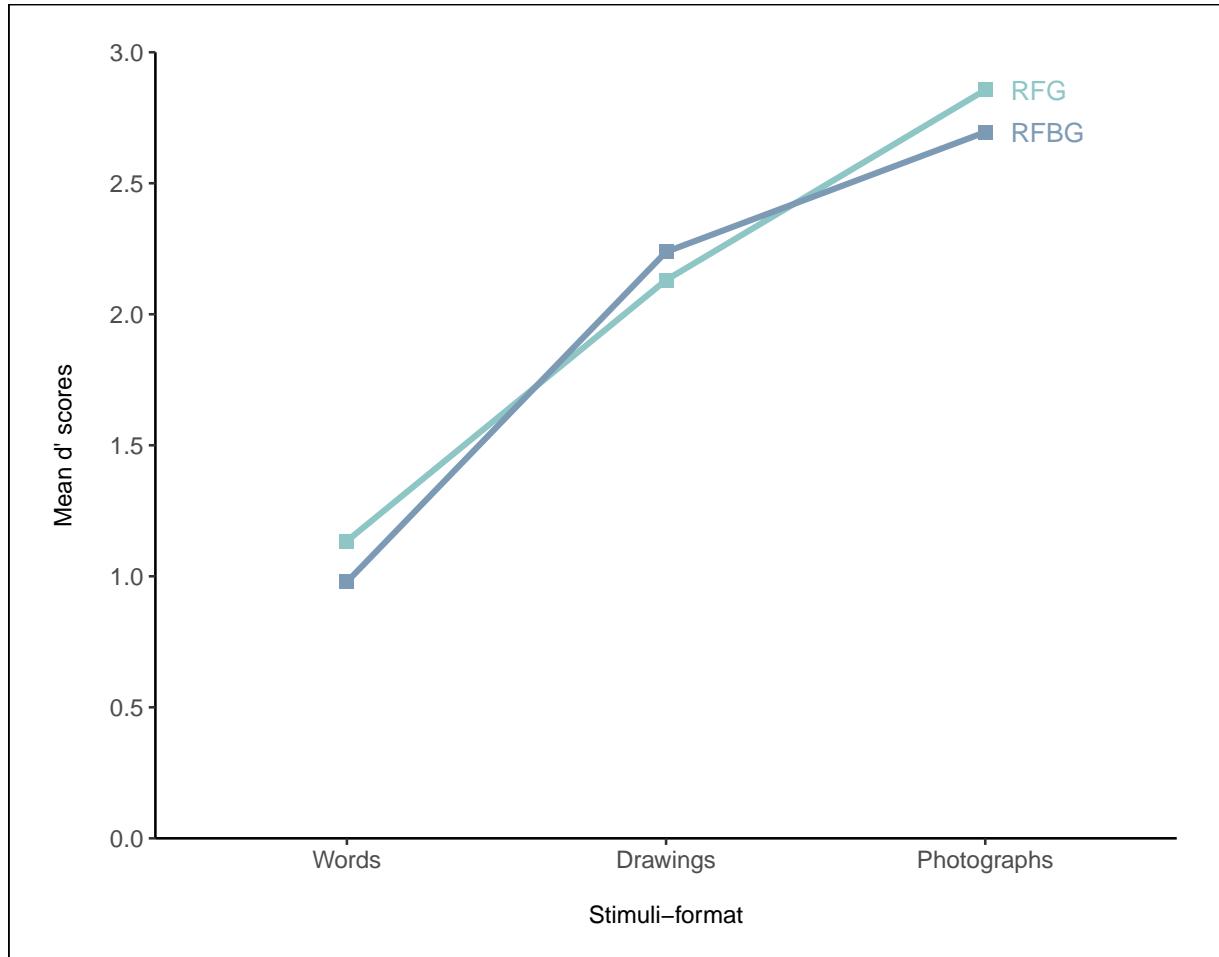


Figure 12: Interaction plot between stimuli format and response option for d' scores.

c-scores showed a significant main effect of stimuli format [$F(1.74, 272.11) = 10.25, MSE = 0.13, p < .001, \hat{\eta}_G^2 = .025$], with photographs ($M = 0.22$) demonstrating significantly lower c-scores (and thus a less conservative response bias) than words ($M = 0.39$) [] and drawings ($M = 0.32$). Photographs may have uniquely provided subjects with additional confidence that their old/new response was correct in comparison to the other stimuli formats; there was no difference in c-scores between drawings ($M = 0.32$) and words ($M = 0.39$). There

were no significant interaction effects [$F(1.74, 272.11) = 0.62, MSE = 0.13, p = .518, \hat{\eta}_G^2 = .002$].

Hits assigned Recollection, Familiarity, and Guessing To determine the effects of stimuli format and response option on accurate recognition memory judgements, separate 3 (stimuli format: words, drawings, photographs) \times 2 (response option condition: RFG-judgements, RFBG-judgements) mixed ANOVAs were conducted on the mean proportion of hits assigned Recollection, Familiarity, and Guessing (see Figure 13).

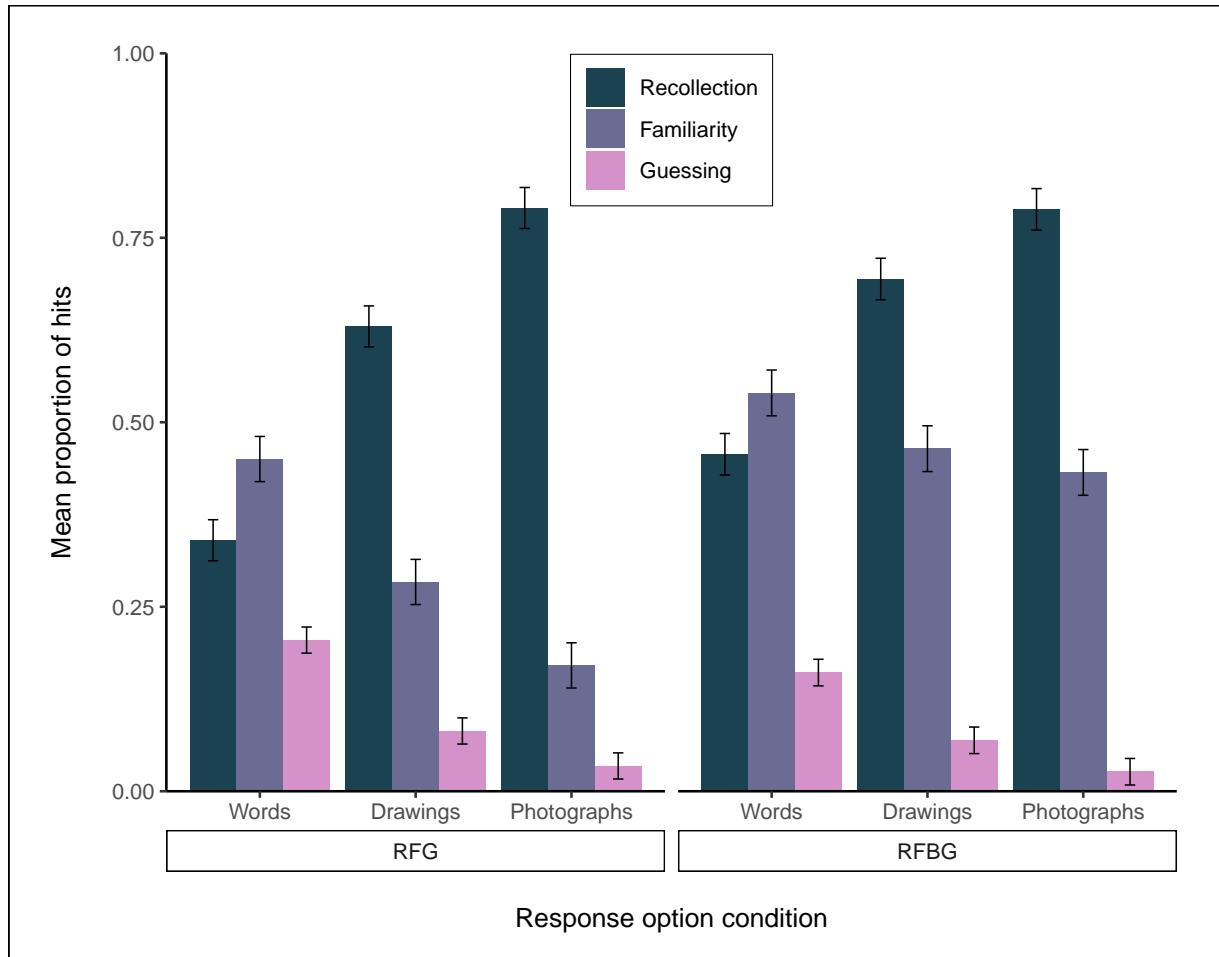


Figure 13: Proportion of hits assigned Recollection, Familiarity, and Guessing, by stimuli format and response option condition.

Recollection (hits) For hits assigned Recollection, there was a significant interaction effect [$F(1.77, 273.98) = 5.32$, $MSE = 0.03$, $p = .007$, $\hat{\eta}_G^2 = .010$] (see Figure 14). Photographs ($M = 0.79$) resulted in more Recollection hits than both words ($M = 0.34$) [$t(310) = -17.90$, $p < .001$] and drawings ($M = 0.63$) [$t(310) = -6.38$, $p < .001$] in the RFG group. Drawings ($M = 0.63$) also resulted in more Recollection hits than words ($M = 0.34$) [$t(310) = -11.52$, $p < .001$] in the RFG group. In the RFBG group, there was an identical pattern; photographs ($M = 0.79$) resulted in more Recollection hits than both words ($M = 0.46$) [$t(310) = -12.62$, $p < .001$] and drawings ($M = 0.69$) [$t(310) = -3.59$, $p = .006$]. Drawings ($M = 0.69$) also resulted in more Recollection hits than words ($M = 0.46$) [$t(310) = -9.03$, $p < .001$] in the RFBG group. Comparisons of stimuli format across the response option groups showed no difference in the number of Recollection hits for photograph stimuli between the RFG ($M = 0.79$) and RFBG ($M = 0.79$) conditions [$t(278.97) = 0.05$, $p > .999$]. The same pattern was evident for drawings (RFG: $M = 0.63$, RFBG: $M = 0.69$ [$t(278.97) = -1.62$, $p > .999$]). Word stimuli, however, produced significantly more Recollection hits in the RFBG group ($M = 0.46$) compared to the RFG ($M = 0.34$) group [$t(278.97) = -2.94$, $p = .053$].

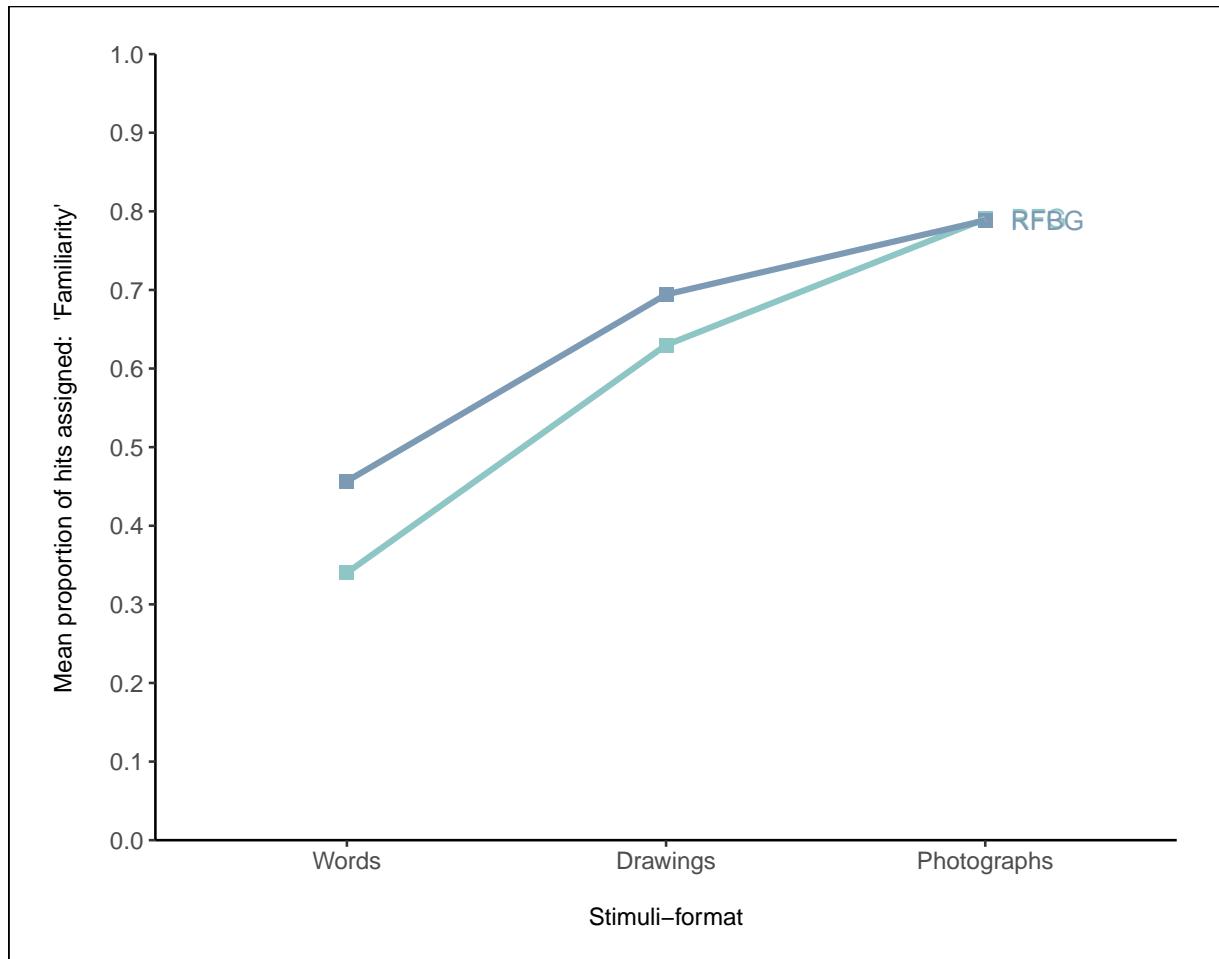


Figure 14: Interaction plot between stimuli format and response option for the mean proportion of hits assigned Recollection.

Familiarity (hits) For hits assigned Familiarity, there was a significant interaction between stimuli format and response option [$F(1.52, 236.21) = 8.68, MSE = 0.04, p = .001, \hat{\eta}_G^2 = .016$] (see Figure 15). Within the RFG condition, words ($M = 0.45$) resulted in more Familiarity hits than both drawings ($M = 0.28$) [$t(310) = 5.84, p < .001$] and photographs ($M = 0.17$) [$t(310) = 9.80, p < .001$]. Drawings ($M = 0.28$) also produced more Familiarity hits compared to photographs ($M = 0.17$) [$t(310) = 3.96, p = .001$].

Within the RFBG condition, words ($M = 0.54$) still produced more Familiarity hits than photographs ($M = 0.43$) [$t(310) = 3.61, p = .005$]. However, there was no difference in

the number of Familiarity hits when comparing words ($M = 0.54$) to drawings ($M = 0.46$) [$t(310) = 2.53, p = .178$]. Another difference from the within-RFG findings is the number of Familiarity hits for drawings ($M = 0.46$) did not differ from photographs ($M = 0.43$) in the RFBG condition [$t(310) = 1.08, p > .999$].

Comparisons across response option conditions showed that drawings produced significantly more Familiarity hits in the RFBG ($M = 0.46$) condition compared to RFG ($M = 0.28$) [$t(289.15) = -4.14, p = .001$]. A similar pattern was also evident for photographs (RFG: $M = 0.17$, RFBG: $M = 0.43$ [$t(289.15) = -6.00, p < .001$]). Words, however, showed no difference in the number of Familiarity hits between the RFG ($M = 0.45$) and RFBG ($M = 0.54$) conditions [$t(289.15) = -2.06, p = .611$].

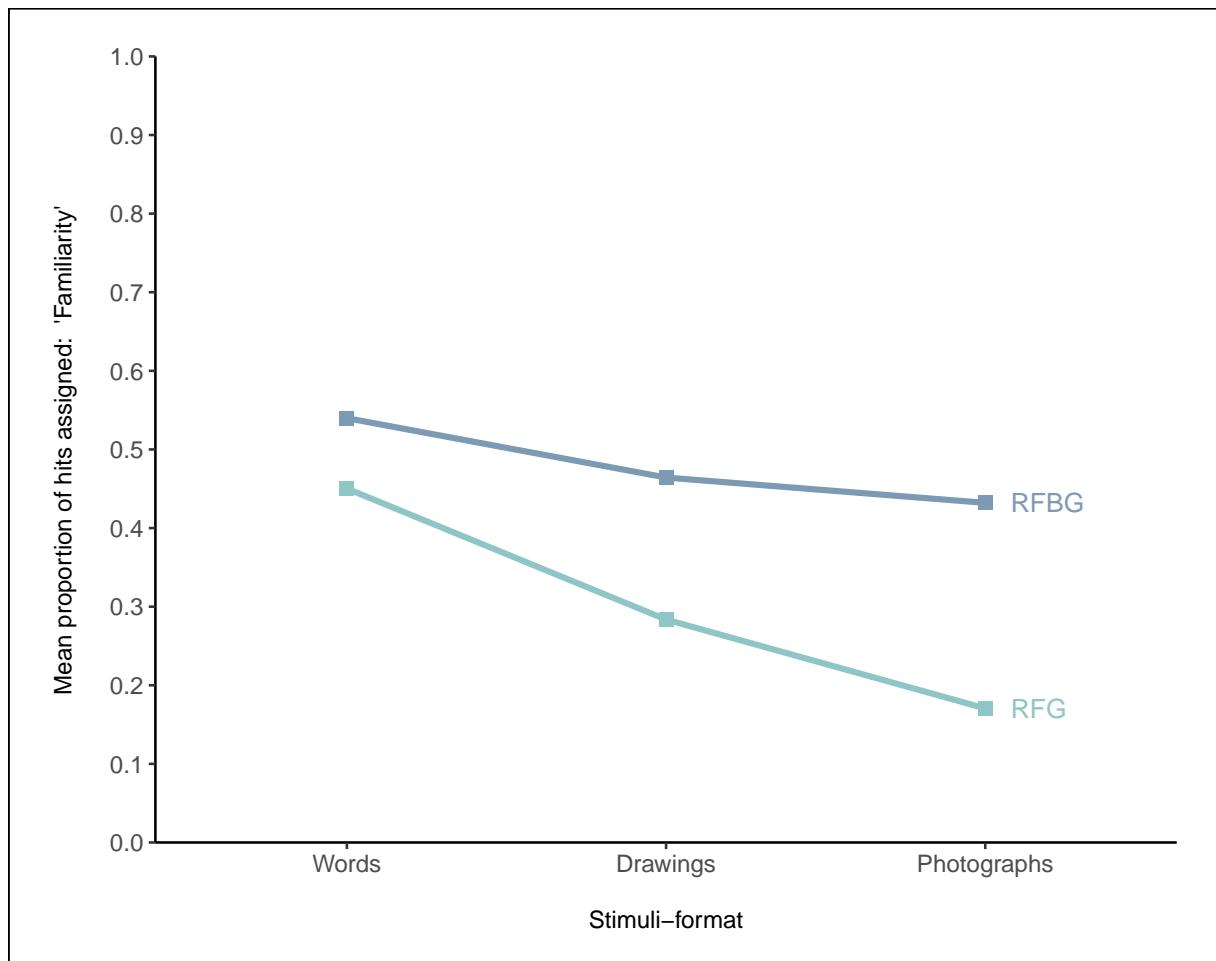


Figure 15: Interaction plot between stimuli format and response option for the mean proportion of hits assigned ‘Familiarity’.

Guessing (hits) For hits assigned Guessing, there was a significant main effect of stimuli format [$F(1.32, 204.20) = 71.22, MSE = 0.02, p < .001, \hat{\eta}_G^2 = .142$]. The interaction effect was not significant [$F(1.32, 204.20) = 1.12, MSE = 0.02, p = .308, \hat{\eta}_G^2 = .003$]. Post-hoc comparisons for the main effect of stimuli format showed that words ($M = 0.18$) produced significantly more Guessing hits than both drawings ($M = 0.08$) [] and photographs ($M = 0.03$) []. Drawings ($M = 0.08$) also produced significantly more Guessing hits compared to photographs ($M = 0.03$) [].

FAs assigned Recollection, Familiarity, and Guessing To determine the effects of stimuli format and response option on false alarms, separate 3 (stimuli format: words, drawings, photographs) x 2 (response option condition: RFG-judgements, RFBG-judgements) mixed ANOVAs were conducted on the mean proportion of FAs assigned Recollection, Familiarity, and Guessing (see Figure 16).

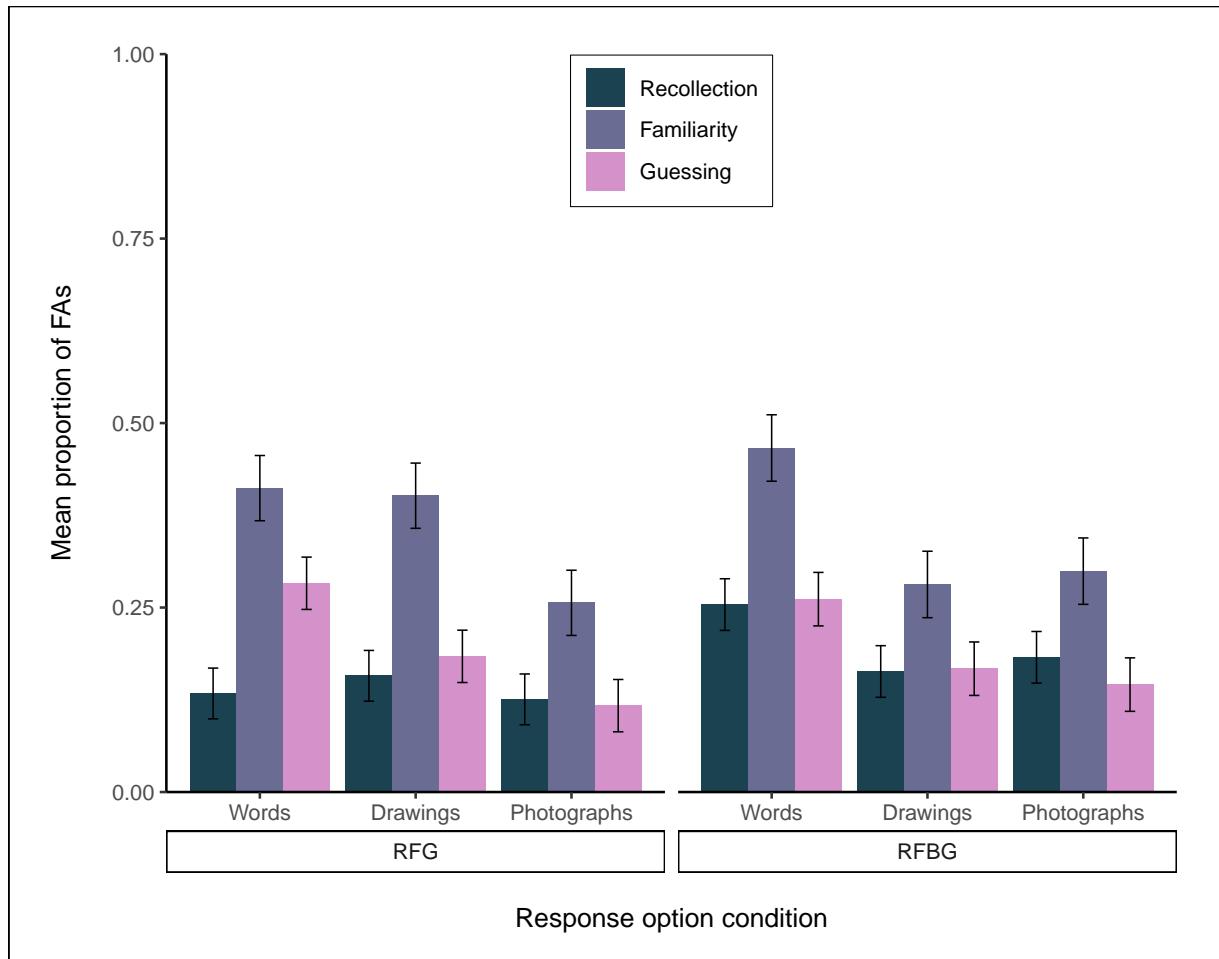


Figure 16: Proportion of FAs assigned Recollection, Familiarity, and Guessing, by stimuli format and response option condition.

Recollection (FAs) For FAs assigned Recollection, there was no significant main effect of stimuli format [$F(1.94, 301.96) = 1.11, MSE = 0.07, p = .328, \hat{\eta}_G^2 = .003$] or interaction [$F(1.94, 301.96) = 2.02, MSE = 0.07, p = .136, \hat{\eta}_G^2 = .006$].

Familiarity (FAs) For FAs assigned Familiarity, there was a significant interaction between stimuli format and response option [$F(1.98, 309.33) = 3.33, MSE = 0.11, p = .038, \hat{\eta}_G^2 = .010$] (see Figure 17). Within the RFG condition, words ($M = 0.41$) resulted in more Familiarity FAs than photographs ($M = 0.26$) [$t(312) = 2.96, p = .050$], but not drawings

($M = 0.40$) [$t(312) = 0.20, p > .999$]. The number of Familiarity FAs did not differ between drawings ($M = 0.40$) and photographs ($M = 0.26$) [$t(312) = 2.76, p = .091$] in the RFG condition.

Within the RFBG condition, words ($M = 0.47$) again produced more Familiarity FAs than photographs ($M = 0.30$) [$t(312) = 3.06, p = .036$]. However, words ($M = 0.47$) also produced more Familiarity FAs than drawings ($M = 0.28$) [$t(312) = 3.39, p = .012$]. As before, the number of Familiarity FAs did not differ between drawings ($M = 0.28$) and photographs ($M = 0.30$) in the RFBG condition [$t(312) = -0.33, p > .999$].

Comparisons across response option conditions showed no differences in the number of Familiarity FAs, for any stimuli format: words (RFG: $M = 0.41$, RFBG: $M = 0.47$ [$t(404.98) = -0.86, p > .999$]); drawings (RFG: $M = 0.40$, RFBG: $M = 0.28$ [$t(404.98) = 1.91, p = .859$]"; photographs (RFG: $M = 0.26$, RFBG: $M = 0.30$ [$t(404.98) = -0.68, p > .999$]).

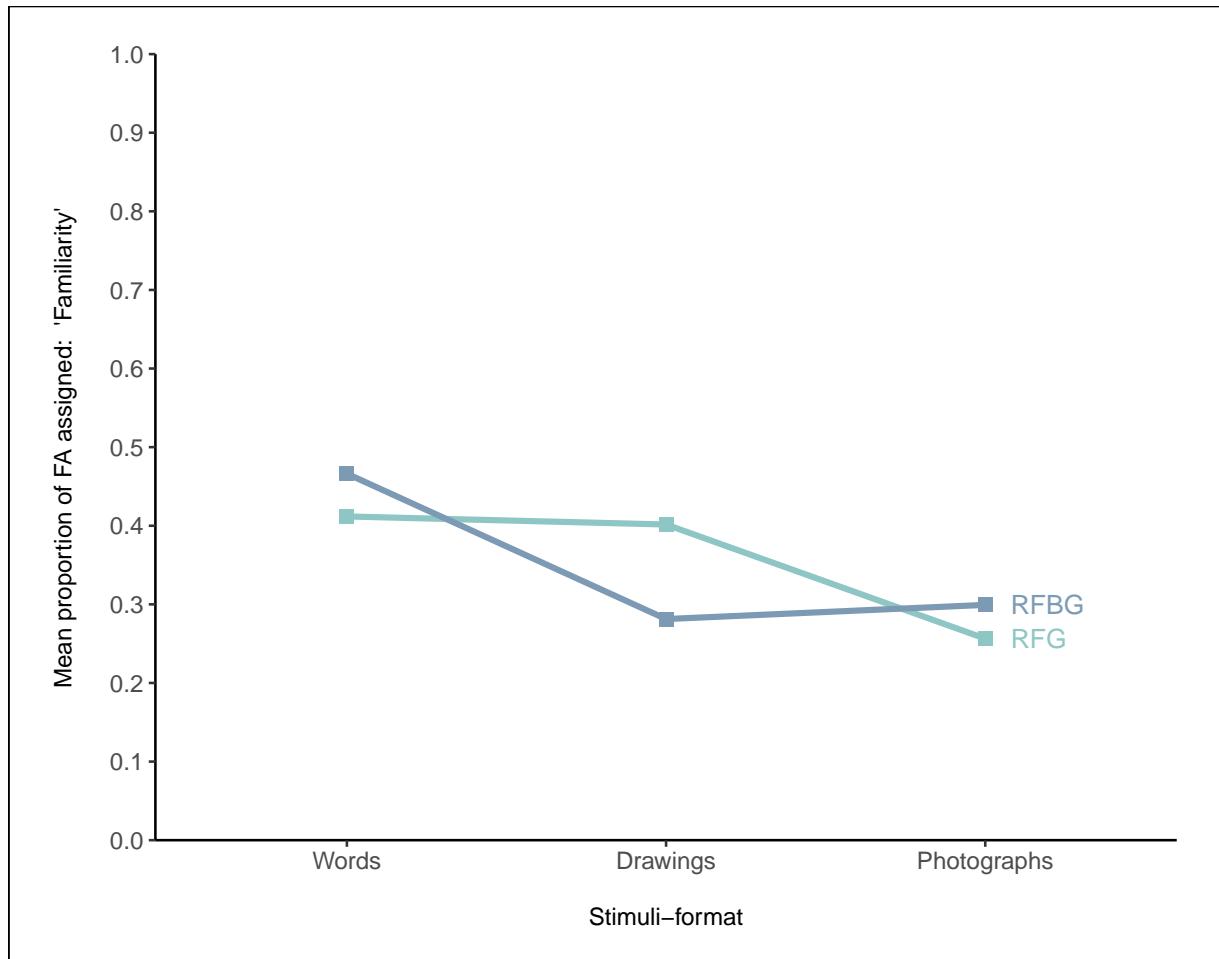


Figure 17: Interaction plot between stimuli format and response option for the mean proportion of FAs assigned ‘Familiarity’.

Guessing (FAs) For FAs assigned Guessing, there was a significant main effect of stimuli format [$F(1.93, 300.68) = 9.44, MSE = 0.09, p < .001, \hat{\eta}_G^2 = .033$]. The interaction effect was not significant [$F(1.93, 300.68) = 0.35, MSE = 0.09, p = .699, \hat{\eta}_G^2 = .001$]. Post-hoc comparisons for the main effect of stimuli format showed that words ($M = 0.27$) produced significantly more Guessing FAs than both drawings ($M = 0.18$) [] and photographs ($M = 0.13$) []. There was no difference in the proportion of FAs assigned Guessing between drawings ($M = 0.18$) and photographs ($M = 0.13$) [].

Discussion

Across a range of performance variables, the results show a clear effect of stimuli distinctiveness. As distinctiveness increased (from words, to drawings, to photographs), this produced more hits, less FAs, better overall recognition, and better discrimination between hits / FAs. The absence of any interaction effects across these variables demonstrates that the availability of different response options (i.e. the addition of a Both option) had little impact on overall performance. RF(B)G responses for accurate recognition displayed a similar pattern; as distinctiveness increased, the number of Recollected hits also increased, while the number of Familiarity and Guessing hits decreased. The rate of both Familiarity FAs and Guessing FAs was also highest for the least distinctive stimuli (words).

#####-----

Chapter 4 (The role of colour)

Experiment: Effect of stimuli format (colour) and response option on recognition memory judgements.

Method

Participants 161 participants completed the experiment online (see Table 8 for a breakdown of the age/gender of the current sample). All participants were required to be between the age of 18-59 years in order to meet our YA criteria (actual range: 18-57). As our experiment involved written words as to-be-remembered stimuli, we also asked that subjects first language be English; the vast majority (96.89%) reported that English was indeed their first language. Subjects were recruited from the voluntary participation website *Prolific Academic¹⁴ (86.34%), where payment at the rate of £5/hr was given, and via the in-school research participation system¹⁵ (13.66%), where they received course participation credits. *G*Power* software was used to calculate an appropriate sample size; to detect a medium effect size of Cohen's $f = 0.25$ with 80% power ($\alpha = .05$, two-tailed), 79 subjects were necessary per group ($N = 158$) in a 3x2 mixed ANOVA.

Table 8: Gender and age (SD) of the current sample.

Gender	N	Age	
Female	97	31.78	(11.18)
Male	60	31.77	(10.3)
Transgender	1	32.00	(0)
Non-binary	1	19.00	(0)
Unspecified	2	38.50	(3.54)
Total	161	31.78	(10.76)

Materials Stimuli were the same as those utilised in *Experiment 3*, only with the greyscale drawings and photographs substituted for their colour versions. Items consisted of 126 innocu-

¹⁴<https://www.prolific.co/>

¹⁵<https://keelepsychology.sona-systems.com/>

ous, everyday objects (e.g. clock, rabbit, shoe), presented across three individual stimuli formats: written words, line drawings, and photographs. Words and line drawings were sourced from @rossion2004; the drawings consisted of shaded, colour illustrations, and the words were simply the written names of the depicted objects (these were presented in a clear Sans-serif typeface in the current experiment). A matching set of photograph stimuli were curated in *Experiment 2*; high quality photographs were sourced to similarly depict the same everyday objects as those found in the @rossion2004 line drawings. In each photograph, the object of interest was isolated from its' original background and rotated to match the orientations shown in the line-drawn items. See Figure 18 for examples of each stimuli format.

Design A mixed 3x2 design was utilised, consisting of a within-subjects factor of stimuli format (words, drawings, photographs) and a between-subjects factor of response option (RFG, RFBG). Blocked randomisation presented participants with: 1) one of six possible study lists (equal length, with the same number of words, drawings, and photographs); 2) one of two possible recognition tests (either RFG: “Recollection”, “Familiarity”, “Guessing”), or RFBG: “Recollection”, “Familiarity”, “Guessing”, “Both”). All routes were of equal length, and subjects were randomly assigned into blocks according to balanced methods.

Words:	Drawings:	Photographs:
cloud		
lock		
penguin		

Figure 18: Example word, drawing, and photograph stimuli.

Procedure The procedure was identical to that of *Experiment 3*; data collection was conducted online using the experiment platform Gorilla¹⁶. All subjects completed three self-paced phases: i) study phase, ii) distractor task, and iii) recognition test. At study, subjects were instructed to learn each of the word, drawing, and photograph items (shown at random, one-at-a-time) in preparation for a later memory test. For each item, participants were required to report whether the current format was a word, drawing, or photograph - an encoding judgement that ensured attention was directed to the to-be-remembered stimuli. Next, subjects completed some simple multiple choice mathematical questions (e.g. $6 \times 4 = ?$) as a distractor task. Finally, participants were presented with the recognition test; word, drawing, and photograph items were once again shown one-at-a-time at random. Half of the test items had been shown previously in the study

¹⁶<https://gorilla.sc/>

phase, while the other half were new (not shown at study). Subjects were first required to make an “Old” / “New” judgement, based on whether they believed they had studied the item earlier or not. While “New” judgements simply led to the next item, “Old” judgements led to a follow-up screen where participants were asked whether they had recognised the item via “Recollection”, “Familiarity”, or were simply taking a “Guess” that it was old. Those in the RFBG response option condition had an additional “Both” option at this stage, where they could report that they had experienced Recollection and Familiarity simultaneously. Stimuli format stayed the same across study and test (e.g. if the item “penguin” was shown as a word at study, it was also shown as a word at test), and same concepts were not repeated across the other formats within-subjects (e.g. if the item “penguin” was shown as a word, that subject would not view the drawing or photo version).

Data processing The primary DVs consisted of the proportion of total hits, false alarms (FAs), and the total number of hits and FAs assigned to each of the available response options (R/F/G and R/F/B/G). Proportions of Recollection and Familiarity were calculated slightly differently depending on the response option condition; in the RFG-judgement group, simple proportions were created from the total number of R responses and the total number of F responses. In the RFBG condition, Both responses were separately added to R proportions and F proportions. Additional DVs included: i) d' (d-prime, a signal detection measure of sensitivity); ii) c-value (a measure of response bias); iii) overall accuracy ($\text{hits} / (\text{hits} + \text{FAs})$); iv) reaction times for all responses.

Subjects were excluded according to two key criteria; 1) less than 90% accuracy during the encoding task (“Is this a word, drawing, or photograph?”); 2) extreme z-scores (those presenting z-scores of $+/- 3$ for total hits, total FAs, or overall recognition [hits minus FAs]). A total of 3 participants were found to meet (at least) one of such criteria, and were thus considered outliers and excluded from analysis.

Results

Distinctiveness To determine whether colour photographs were more distinctive than colour drawings, and examine any potential interactions with response-option condition, a series of 2 (stimuli format: words, pictures) x 3 (response option condition: RFG-judgements, RFBG-judgements, RF-ratings) mixed ANOVAs were conducted on the proportion of overall hits, false alarms (FAs), and overall recognition (hits - FAs) [see Table 9]. To further examine response patterns between the stimuli formats, 2x3 mixed ANOVAs were also run on the signal detection measures of d' (sensitivity) and c (decision criterion). Significant main effects were followed up with single pairwise comparisons (drawings vs. photographs).

Table 9: Mean proportion of hits, FAs, and overall recognition, by stimuli format and response option condition.

	Hits	FAs	Overall recognition
Stimuli format			
Words	0.56	0.23	0.33
Drawings	0.73	0.08	0.65
Photographs	0.87	0.04	0.83
Response option			
RFG	0.74	0.13	0.60
RFBG	0.70	0.10	0.60

Visual inspection of the data (Table 9) shows some expected patterns with regard to stimuli distinctiveness. As the intended distinctiveness increases across the formats (from words, to drawings, to photographs), the mean proportion of hits increases, the mean proportion of FAs decreases, and overall scores of recognition (hits - FAs) also increases. These patterns support the notion that distinctiveness plays a key role in the memorability of stimuli, and also that the current stimuli indeed appear to capture different 'levels' of stimuli distinctiveness. The ANOVA on the proportion of hits demonstrated a significant main effect of stimuli-format, $F(1.70, 271.03) = 187.25$, $MSE = 0.02$, $p < .001$, $\hat{\eta}_G^2 = .311$; a distinctiveness effect was apparent, with

colour photographs ($M = 0.87$) showing a higher number of overall hits compared to colour drawings ($M = 0.73$), $t(318) = -8.53, p < .001$. Similarly, the ANOVA on the proportion of FAs also supported a distinctiveness effect; a significant main effect of stimuli-format $F(1.26, 200.79) = 123.14, MSE = 0.02, p < .001, \hat{\eta}_G^2 = .279$ showed that drawings ($M = 0.08$) produced more FAs than photographs ($M = 0.04$), $t(318) = 3.07, p = .002$. Overall recognition performance (taking into account hits and FAs) offered further support for a distinctiveness effect; a significant main effect of stimuli-format $F(1.79, 284.84) = 527.35, MSE = 0.02, p < .001, \hat{\eta}_G^2 = .550$ showed photographs ($M = 0.83$) produced better overall task performance compared to drawings ($M = 0.65$), $t(318) = -11.38, p < .001$. Taken together, the findings support the distinctiveness hypothesis, in that colour photographs were recognised better than colour drawings.

The ANOVA on d' -scores produced findings consistent with those found for hits, FAs, and overall recognition; there was a significant main effect of stimuli-format, $F(1.92, 305.18) = 497.69, MSE = 0.30, p < .001, \hat{\eta}_G^2 = .543$, whereby photographs ($M = 2.87$) facilitated better discrimination between hits and FAs than drawings ($M = 2.13$), $t(318) = -12.41, p < .001$. c -scores also showed a significant main effect of stimuli-format, $F(1.61, 256.18) = 7.21, MSE = 0.17, p = .002, \hat{\eta}_G^2 = .018$, with drawings ($M = 0.37$) showing higher c -scores (and thus a more conservative response bias) than photographs ($M = 0.22$), $t(318) = 3.57, p < .001$.

Response option In each of the previously discussed ANOVAs, the role of response option was also assessed. However, there were no significant interaction effects between stimuli format and response option for any of the examined DVs: i) proportion of hits; ii) proportion of FAs; iii) overall recognition; iv) d' -scores; v) c -scores. There was only a significant main effect of response option in the ANOVA on the mean proportion of FAs, $F(1, 159) = 5.25, MSE = 0.02, p = .023, \hat{\eta}_G^2 = .016$ (with the RFG group, $M = 0.13$, showing more FAs than the RFBG group, $M = 0.10$), $t(159) = 2.29, p = .023$, and the ANOVA on c -scores, $F(1, 159) = 4.85, MSE = 0.39, p = .029, \hat{\eta}_G^2 = .018$ (with the RFBG group, $M = 0.38$,

showing a more conservative response bias than the RFG group, $M = 0.25$, $t(159) = -2.20$, $p = .029$

Role of colour Visual inspection of the data from *Experiment 3*, which used greyscale drawings and photographs, alongside that obtained in the current study (Figure 19) suggests the effects of colour on distinctiveness are negligible. Performance, in terms of the proportion of hits, FAs, and overall recognition accuracy, appears the same irrespective of whether greyscale or colour pictures were used.

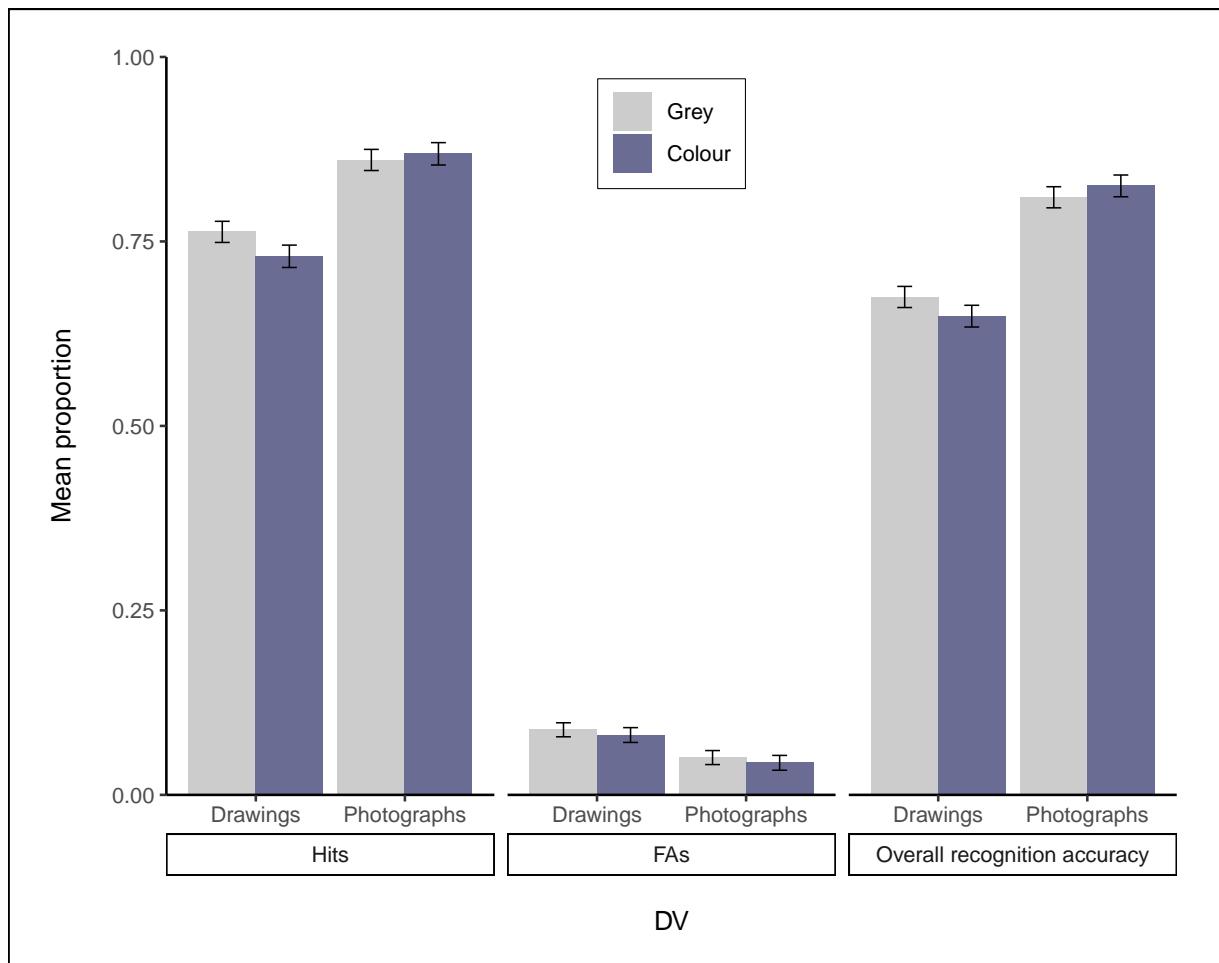
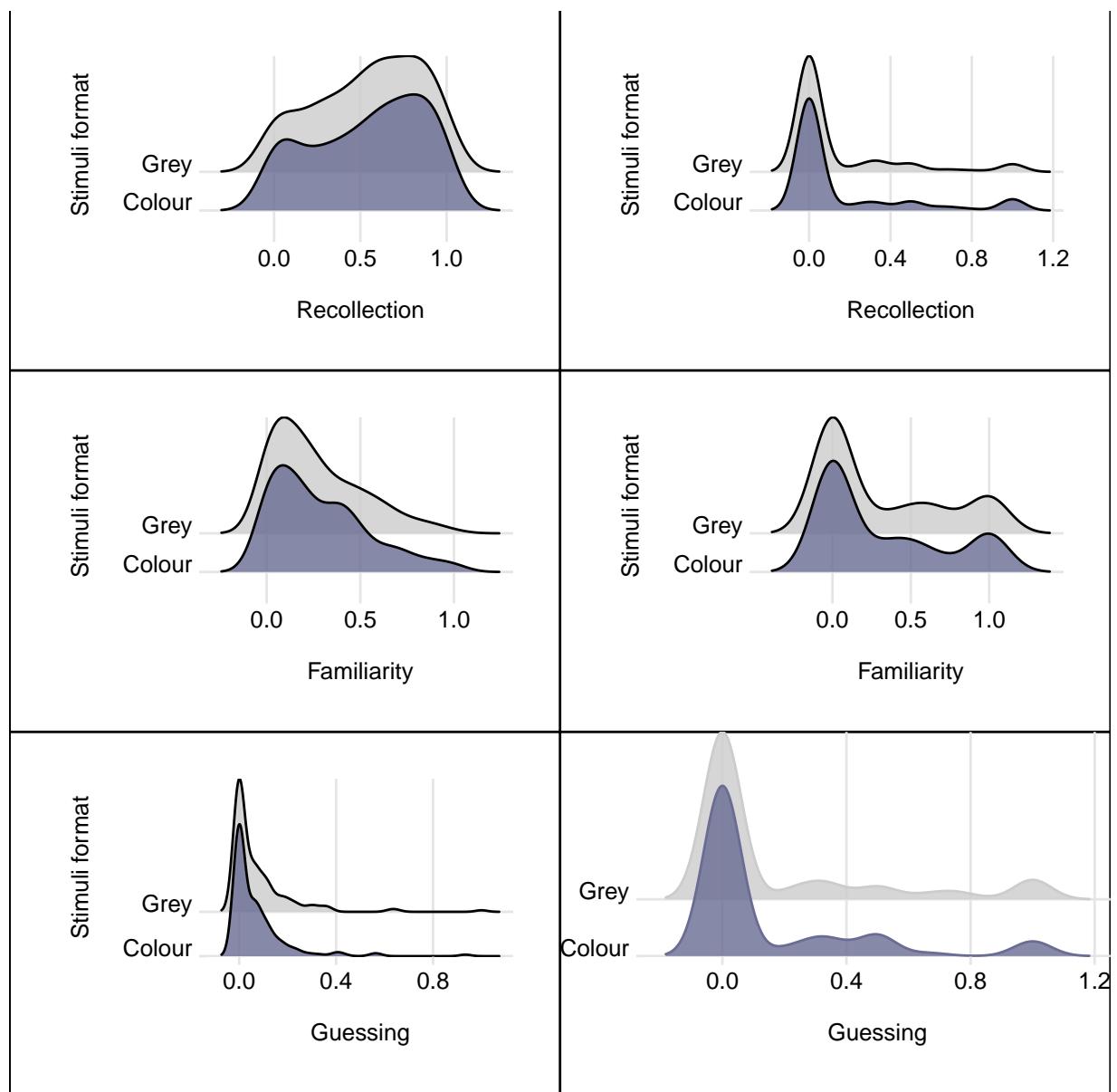


Figure 19: Grey vs. colour

In turn, this further implicates the stimuli format as the



Discussion

#####-----

Chapter 5 (Distinctiveness of word stimuli)

Experiment: Manipulating word distinctiveness

Method

All stimuli were presented as 500x500px image files at their actual size (i.e. without scaling) to ensure consistency across participants. Adobe Photoshop 2021 (22.0.0 Release) was used to create the image files, all on a plain white canvas. The Rossion and Pourtois (2004) line-drawings were sourced directly from the authors, and were trimmed of any preexisting background before being resized to fit the 500x500px canvas (for some items this was determined by a maximum width of 500px, whilst for others it was determined by a maximum height of 500px).

word stimuli were created in Roboto (light, 54pt)

Planned comparisons:

Distinctive words vs. grey words (are they actually more distinctive?)

vs. grey drawings	(are they as distinctive as grey drawings?)
vs. grey photos	(are they as distinctive as grey photos?)
vs. colour drawings	(are they as distinctive as colour drawings?)
vs. colour photos	(are they as distinctive as colour photos?)

Results

Proportion of total hits	Effect	df	MSE	F	ges	p.value
	stim_format	1.84, 289.13	0.03	282.85 ***	.323	<.001
	distinctiveness	1, 157	0.02	3.46 +	.002	.065
	stim_format:distinctiveness	1.71, 268.93	0.02	0.46	<.001	.604

Significant main effect of stimuli format: $F(1.84, 289.13) = 282.85$, $MSE = 0.03$, $p < .001$,

$$\hat{\eta}_G^2 = .323.$$

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	hit	158	0.579	0.243
Words	High	hit	158	0.606	0.233
Drawings	Low	hit	158	0.754	0.191
Drawings	High	hit	158	0.775	0.189
Photos	Low	hit	158	0.908	0.124
Photos	High	hit	158	0.915	0.113

No main effect of distinctiveness: $F(1, 157) = 3.46$, $MSE = 0.02$, $p = .065$, $\hat{\eta}_G^2 = .002$.

No interaction effects: $F(1.71, 268.93) = 0.46$, $MSE = 0.02$, $p = .604$, $\hat{\eta}_G^2 = .001$.

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	FA	158	0.175	0.187
Words	High	FA	158	0.051	0.079
Drawings	Low	FA	158	0.075	0.114
Drawings	High	FA	158	0.053	0.092
Photos	Low	FA	158	0.039	0.067
Photos	High	FA	158	0.020	0.052

Proportion of total FA	Effect	df	MSE	F	ges	p.value
	stim_format	1, 74, 273.37	0.01	53.57 ***	.093	<.001
distinctiveness		1, 157	0.01	86.43 ***	.062	<.001
stim_format:distinctiveness		1.63, 255.48	0.01	34.16 ***	.048	<.001

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	recog	158	0.404	0.245
Words	High	recog	158	0.555	0.230
Drawings	Low	recog	158	0.679	0.208
Drawings	High	recog	158	0.723	0.208
Photos	Low	recog	158	0.868	0.142
Photos	High	recog	158	0.895	0.125

	Effect	df	MSE	F	ges	p.value
	stim_format	1, 97, 310.07	0.03	425.75 ***	.410	<.001
Overall recognition:	distinctiveness	1, 157	0.03	47.83 ***	.034	<.001
	stim_format:distinctiveness	1.68, 263.84	0.03	16.23 ***	.019	<.001

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	d.prime	158	1.183	0.742
Words	High	d.prime	158	1.714	0.688
Drawings	Low	d.prime	158	2.059	0.698
Drawings	High	d.prime	158	2.223	0.729
Photos	Low	d.prime	158	2.746	0.584
Photos	High	d.prime	158	2.871	0.541

d' scores	Effect	df	MSE	F	ges	p.value
	stim_format	2.00, 313.36	0.35	424.00 ***	.410	<.001
	distinctiveness	1, 157	0.33	54.02 ***	.041	<.001
	stim_format:distinctiveness	1.81, 284.95	0.28	15.85 ***	.018	<.001

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	c	158	0.386	0.553
Words	High	c	158	0.566	0.412
Drawings	Low	c	158	0.302	0.409
Drawings	High	c	158	0.318	0.361
Photos	Low	c	158	0.105	0.304
Photos	High	c	158	0.151	0.263

C scores:

Effect	df	MSE	F	ges	p.value
stim_format	1.79, 280.46	0.14	74.32 ***	.115	<.001
distinctiveness	1, 157	0.10	14.90 ***	.010	<.001
stim_format:distinctiveness	1.84, 288.51	0.10	6.70 **	.008	.002

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	rec_hit	158	0.357	0.286
Words	High	rec_hit	158	0.480	0.304
Drawings	Low	rec_hit	158	0.563	0.306
Drawings	High	rec_hit	158	0.590	0.295
Photos	Low	rec_hit	158	0.713	0.322
Photos	High	rec_hit	158	0.733	0.325

Prop Rec hit:	Effect	df	MSE	F	ges	p.value
stim_format	1.42, 223.31	0.09	118.98 ***	.142	<.001	
distinctiveness	1, 157	0.03	24.81 ***	.008	<.001	
stim_format:distinctiveness	1.70, 266.57	0.04	8.59 ***	.006	<.001	

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	fam_hit	158	0.480	0.289
Words	High	fam_hit	158	0.407	0.288
Drawings	Low	fam_hit	158	0.385	0.285
Drawings	High	fam_hit	158	0.358	0.270
Photos	Low	fam_hit	158	0.270	0.311
Photos	High	fam_hit	158	0.245	0.315

Prop FAM hit:	Effect	df	MSE	F	ges	p.value
	stim_format	1.40, 220.37	0.10	40.99 ***	.064	<.001
	distinctiveness	1, 157	0.03	12.00 ***	.005	<.001
	stim_format:distinctiveness	1.62, 254.40	0.04	1.69	.001	.191

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	guess_hit	158	0.131	0.185
Words	High	guess_hit	158	0.100	0.192
Drawings	Low	guess_hit	158	0.052	0.102
Drawings	High	guess_hit	158	0.052	0.105
Photos	Low	guess_hit	158	0.017	0.053
Photos	High	guess_hit	158	0.023	0.058

Prop GUESS hit:	Effect	df	MSE	F	ges	p.value
	stim_format	1, 47, 230.35	0.02	46.07 ***	.089	<.001
	distinctiveness	1, 157	0.01	1.52	.001	.219
	stim_format:distinctiveness	1.31, 205.90	0.02	2.97 +	.004	.075

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	rec_FA	158	0.101	0.252
Words	High	rec_FA	158	0.041	0.178
Drawings	Low	rec_FA	158	0.109	0.300
Drawings	High	rec_FA	158	0.057	0.217
Photos	Low	rec_FA	158	0.066	0.247
Photos	High	rec_FA	158	0.027	0.149

Prop REC FA:	Effect	df	MSE	F	ges	p.value
	stim_format	1.96, 308.15	0.04	2.45 +	.004	.089
	distinctiveness	1, 157	0.04	13.91 ***	.012	<.001
	stim_format:distinctiveness	1.89, 297.40	0.04	0.21	<.001	.800

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	fam_FA	158	0.378	0.428
Words	High	fam_FA	158	0.185	0.376
Drawings	Low	fam_FA	158	0.234	0.396
Drawings	High	fam_FA	158	0.210	0.397
Photos	Low	fam_FA	158	0.139	0.333
Photos	High	fam_FA	158	0.061	0.231

Prop FAM FA:	Effect	df	MSE	F	ges	p.value
	stim_format	1.97, 309.05	0.12	22.29 ***	.041	<.001
	distinctiveness	1, 157	0.12	19.14 ***	.018	<.001
	stim_format:distinctiveness	1.81, 283.95	0.11	5.73 **	.009	.005

stim_format	distinctiveness	variable	n	mean	sd
Words	Low	guess_FA	158	0.185	0.323
Words	High	guess_FA	158	0.147	0.342
Drawings	Low	guess_FA	158	0.101	0.271
Drawings	High	guess_FA	158	0.074	0.242
Photos	Low	guess_FA	158	0.111	0.302
Photos	High	guess_FA	158	0.057	0.226

Prop GUESS FA:	Effect	df	MSE	F	ges	p.value
	stim_format	1.91, 300.51	0.08	9.45 ***	.017	<.001
	distinctiveness	1, 157	0.08	4.83 *	.005	.029
	stim_format:distinctiveness	2.00, 313.38	0.08	0.19	<.001	.826

Results

Manipulation check

To examine whether the distinctiveness manipulation for word stimuli was effective, a series of two-way repeated measures ANOVAs (Greenhouse-Geisser corrected) were run on the mean proportion of hits, mean proportion of false alarms (FAs), and overall recognition scores (hits - FAs).

While there was no significant main effect of distinctiveness in the ANOVA on the mean proportion of hits [$F(1, 157) = 3.46, MSE = 0.02, p = .065, \hat{\eta}_G^2 = .002$], significant effects were demonstrated in the ANOVA on the mean proportion of FAs [$F(1, 157) = 86.43, MSE = 0.01, p < .001, \hat{\eta}_G^2 = .062$] and overall recognition scores [$F(1, 157) = 47.83, MSE = 0.03, p < .001, \hat{\eta}_G^2 = .034$].

#####-----

References

Appendices

Appendix A: d' calculations

Appendix B: Spelling corrections / manipulations to naming responses.

Response	Correction	Response	Correction	Response	Correction
;ashtray	ashtray	draw	drawers	lobster	lobster
a	no	drawer	drawers	onions	onion
acoop	scoop	draws	drawers	onions	onion
ancher	anchor	drums	drum	osterich	ostrich
anchore	anchor	eagal	eagle	ostrage	ostrich
ancor	anchor	eclipes	eclipse	ostridge	ostrich
anker	anchor	eclipsse	eclipse	ostrigie	ostrich
aparagus	asparagus	eeagle	eagle	ostrisge	ostrich
aprincote	apricot	eele	seal	ostritch	ostrich
ashtry	ashtray	eyeglass	eyeglasses	pair	pear
ballon	balloon	falg	flag	paper	pepper
ballone	balloon	feet	foot	peacck	peacock
balloone	balloon	fencing	fence	pecock	peacock
ballun	balloon	footstall	footstool	peguin	penguin
baloon	balloon	fott	foot	peneut	peanut
bamnna	banana	frock	frog	pengiuin	penguin
bananaa	banana	frog/	frog	pengiuin	penguin
bananna	banana	geese	goose	penguine	penguin
bannan	banana	giaffee	giraffe	penquin	penguin
bannana	banana	giaraffe	giraffe	peper	pepper
barel	barrel	giraffe	giraffe	pestleandmorter	pestleandmortar
barrell	barrel	girafe	giraffe	piccalo	piccolo
barlle	barrel	giraff	giraffe	pilers	pliers
barrow	barrel	giraffee	giraffe	piars	pliers
beatle	beetle	girafffe	giraffe	plier	pliers
beer	bear	girrafe	giraffe	plyers	pliers

(continued)

Response	Correction	Response	Correction	Response	Correction
bellpepp34	bellpepper	girraffe	giraffe	plugin	plug
bettle	beetle	gitaur	guitar	plyers	pliers
bicucle	bicycle	gitter	guitar	potatoe	potato
bicyle	bicycle	glases	glasses	pottato	potato
bittle	bottle	glass	glasses	pumkin	pumpkin
bittle	beetle	glassesbottle	bottle	pumpkim	pumpkin
blueberrys	blueberries	gloves	glove	pumpkin	pumpkin
bolw	bowl	grape	grapes	rabit	rabbit
bootle	bottle	gutair	guitar	racoon	raccoon
broon	broom	haircomb	comb	rubarb	rhubarb
broon	broom	hamp	harp	rule	ruler
brum	broom	hand5	hand	seel	seal
busket	basket	harper	hamper	showel	shovel
bycycle	bicycle	hemmar	hammer	snakw	snake
camal	camel	hose	house	soak	socks
canddle	candle	idk	no	specs	spectacles
canle	candle	kacket	jacket	spon	spoon
canon	cannon	kangroo	kangaroo	ssnowman	snowman
carrott	carrot	ladders	ladder	steplader	stepladder
carrots	carrot	lader	ladder	sterss	step
carrott	carrot	latter	ladder	sweetcurn	sweetcorn
celary	celery	leafe	leaf	tabaccopipe	tobaccopipe
celeary	celery	leamon	lemon	teakettle	kettle
cellary	celery	leema	lemur	thinbell	thimble
cerlery	celery	lettace	lettuce	thmble	thimble
chain2	chain	lip	lips	thunb	thumb
chestofdrawerserss	chest of drawers	longdress	dress	timbil	thimble

(continued)

Response	Correction	Response	Correction	Response	Correction
chestofdraws	chest of drawers	maledear	maledeer	timble	thimble
chisle	chisel	meercat	meerkat	toitouse	tortoise
chissel	chisel	mercat	meerkat	tomatoe	tomato
chizel	chisel	mice	mouse	tomoato	tomato
claranet	clarinet	mit	mitten	tortise	tortoise
clouds	cloud	mittens	mitten	tortiste	tortoise
cochroach	cockroach	monkeybut	monkeynut	tortus	tortoise
cock	cockeral	mortle	mortar	usplug	plug
cockaroach	cockroach	mousse	moose	vulture	vulture
cockrel	cockerel	muscat	muskrat	vicescripts	vicegrips
combe	comb	nectarin	nectarine	violin	violin
cycle	bicycle	nectarinee	nectarine	violen	violin
dear	deer	nectrine	nectarine	volion	violin
deere	deer	neddle	needle	waistcoast	waistcoat
dock	duck	noise	nose	wale	well
dolly	doll	none	no	wasteccoat	waistcoat
doormouse	dormouse	oencil	pencil	whisell	whistle
				whistel	whistle
				white	whistle
				whsitle	whistle
				windown	window

Appendix C: Normative data for all photograph items.

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
anchor				
anchor1-photo-colour	3.55 (1.32)	3 (1.11)	2.77 (1.19)	4.05 (1.05)
anchor1-photo-grey	3.03 (1.47)	2.81 (0.93)		3.95 (0.89)
anchor2-photo-colour	3.38 (1.56)	3.45 (1.23)	3.5 (1.19)	3.75 (1.25)
anchor2-photo-grey	2.82 (1.53)	3 (0.97)		3.52 (1.33)
anchor3-photo-colour	3.5 (1.3)	2.48 (0.93)	2.76 (1.3)	3.86 (1.01)
anchor3-photo-grey	3.04 (1.43)	2.46 (1.22)		4.27 (0.77)
apple				
apple1-photo-colour	4.8 (0.48)	3.1 (1.14)	3.43 (1.08)	4.15 (0.88)
apple1-photo-grey	4.85 (0.49)	2.32 (1.21)		3.05 (1.23)
apple2-photo-colour	4.91 (0.29)	3 (0.97)	3.5 (1.32)	3.76 (1.04)
apple2-photo-grey	4.24 (1.14)	2.4 (0.99)		2.75 (1.45)
apple3-photo-colour	4.77 (0.53)	3.16 (1.37)	3.68 (1.25)	4.27 (0.98)
apple3-photo-grey	4.73 (0.63)	2.38 (1.07)		3.38 (1.28)
ashtray				
ashtray1-photo-colour	3.77 (1.43)	3.05 (1.16)	2.71 (1.1)	3.9 (1.07)
ashtray1-photo-grey	4.15 (1.14)	2.95 (1.13)		3.5 (1.28)
ashtray2-photo-colour	3.86 (1.46)	3 (1.08)	2.25 (1.29)	4.1 (1.3)
ashtray2-photo-grey	3.52 (1.47)	3 (1.12)		3.5 (1.28)
ashtray3-photo-colour	3.36 (1.47)	3.75 (1.15)	3.21 (1.32)	3.5 (1.19)
ashtray3-photo-grey	4.14 (0.99)	3.09 (1.11)		3.76 (1)
balloon				
balloon1-photo-colour	4.4 (1.1)	1.63 (1)	2.35 (1.69)	4.62 (0.92)
balloon1-photo-grey	4.15 (1.18)	1.8 (1.11)		3.41 (1.1)
balloon2-photo-colour	4.45 (1)	2.14 (1.08)	1.82 (1.3)	4.5 (1)
balloon2-photo-grey	4.35 (0.81)	1.9 (1)		4.05 (0.94)
balloon3-photo-colour	4.09 (1.02)	1.86 (1.04)	1.86 (1.49)	4.46 (0.88)
balloon3-photo-grey	4.24 (0.94)	1.68 (0.99)		3.57 (1.4)
banana				
banana1-photo-colour	4.65 (0.99)	2.55 (1.5)	4.55 (0.76)	4.45 (0.86)
banana1-photo-grey	4.8 (0.7)	2.23 (1.1)	2.36 (1.5)	3.76 (0.94)
banana2-photo-colour	4.8 (0.41)	2.05 (1.24)	4 (1.3)	4.85 (0.37)
banana2-photo-grey	4.9 (0.45)	2.36 (1.36)		3.3 (1.42)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
banana3-photo-colour	4.33 (1.02)	2.05 (1.09)	4.59 (0.96)	4.67 (0.66)
banana3-photo-grey	4.86 (0.47)	2 (0.87)		3.58 (1.1)
barrel				
barrel1-photo-colour	3.53 (1.25)	3.57 (0.93)	3.95 (0.97)	4.9 (0.45)
barrel1-photo-grey	3.9 (1.29)	3 (1.02)		4.05 (1.05)
barrel2-photo-colour	4 (1.02)	3 (1.08)	3.5 (1.15)	4.43 (0.81)
barrel2-photo-grey	3.81 (1.54)	2.95 (1.1)		4.15 (0.93)
barrel3-photo-colour	3.5 (1.26)	3.38 (1.17)	2.96 (1.3)	4.22 (0.95)
barrel3-photo-grey	3.45 (1.5)	2.68 (1.13)		4 (0.95)
basket				
basket1-photo-colour	4.13 (1.04)	3.86 (0.96)	3.62 (1.24)	4.3 (1.13)
basket1-photo-grey	4.5 (0.83)	3.36 (1.09)		3.85 (1.14)
basket2-photo-colour	4.14 (1.08)	3.05 (1)	2.8 (1.36)	3.95 (0.97)
basket2-photo-grey	4.48 (0.93)	2.9 (1.45)		3.55 (1.23)
basket3-photo-colour	4.27 (0.83)	3.8 (1.08)	3.48 (1.53)	3.64 (1.14)
basket3-photo-grey	4.45 (0.86)	3.18 (1.1)		3.95 (0.92)
bear				
bear1-photo-colour	3.36 (1.53)	3.9 (0.79)	3.9 (0.72)	4.15 (1.09)
bear1-photo-grey	3.9 (1.34)	3.5 (1.15)		3.93 (1.08)
bear2-photo-colour	3.85 (1.53)	3.25 (0.91)	4.3 (0.73)	4.24 (1.22)
bear2-photo-grey	3.25 (1.45)	3.25 (1.16)		4.27 (1.12)
bear3-photo-colour	3.9 (1.37)	3.81 (1.25)	3.71 (1.1)	4.09 (0.87)
bear3-photo-grey	3.75 (1.48)	3 (1.02)		4 (1.23)
beetle				
beetle1-photo-colour	3.1 (1.59)	4 (0.79)	3.15 (1.5)	2.95 (1.07)
beetle1-photo-grey	3.1 (1.45)	3.1 (1.21)		2.95 (1.21)
beetle2-photo-colour	3.2 (1.4)	3.73 (1.16)	3 (1.38)	2.95 (1.19)
beetle2-photo-grey	3 (1.34)	2.95 (1.24)		2.7 (1.13)
beetle3-photo-colour	2.68 (1.25)	3.41 (0.96)	3.05 (1.13)	3.17 (1.05)
beetle3-photo-grey	3 (1.26)	3.18 (1.05)		3.24 (1.04)
bell				
bell1-photo-colour	4 (1.21)	2.47 (1.41)	3.55 (1.39)	4 (1.05)
bell1-photo-grey	3.8 (1.28)	2.45 (1.05)		3.64 (0.95)
bell2-photo-colour	3.9 (1.14)	3.32 (0.95)	3.41 (1.3)	3.95 (1.1)

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
bell2-photo-grey	3.5 (1.15)	2.24 (0.62)		3.3 (0.86)
bell3-photo-colour	3.43 (1.41)	3.09 (1.02)	2.91 (1.34)	3.5 (1.06)
bell3-photo-grey	3.62 (1.47)	2.82 (0.96)		3.05 (1.36)
belt				
belt1-photo-colour	4.8 (0.52)	2.86 (1.21)	2.68 (1.46)	3.95 (0.83)
belt1-photo-grey	4.6 (0.89)	3.05 (0.97)		4 (0.92)
belt2-photo-colour	4.71 (0.72)	3 (1.38)	2.85 (1.46)	4.05 (1.05)
belt2-photo-grey	4.64 (0.79)	2.4 (1.1)		4.48 (0.87)
belt3-photo-colour	4.86 (0.64)	2.43 (1.29)	3.05 (1.43)	4.24 (0.77)
belt3-photo-grey	4.35 (1.03)	2.21 (1.28)		3.95 (1.09)
bicycle				
bicycle1-photo-colour	4.3 (1.13)	3.55 (1.1)	2.45 (1.47)	3.14 (1.17)
bicycle1-photo-grey	4.5 (1)	3.6 (1.04)	1.73 (1.19)	3.33 (1.2)
bicycle2-photo-colour	4.65 (0.67)	2.95 (1.12)	1.57 (0.98)	4.05 (0.94)
bicycle2-photo-grey	4.76 (0.62)	3.41 (1.05)		3.65 (0.93)
bicycle3-photo-colour	3.86 (1.15)	3.27 (0.98)	1.91 (1.27)	3.38 (1.16)
bicycle3-photo-grey	4.05 (1.05)	3.48 (0.99)		3.33 (1.46)
book				
book1-photo-colour	4.85 (0.37)	3.15 (1.23)	2.4 (1.27)	3.59 (1.14)
book1-photo-grey	4.9 (0.45)	2.7 (0.99)	1.45 (0.93)	3.62 (0.8)
book2-photo-colour	4.75 (0.72)	3.1 (1.09)	1.71 (1.06)	3.65 (1.35)
book2-photo-grey	4.75 (0.55)	2.86 (1.08)		3.2 (1.15)
book3-photo-colour	4.33 (0.91)	3.05 (1.05)	2.27 (1.24)	3 (1.14)
book3-photo-grey	4.04 (1.36)	2.45 (0.51)		3 (1.18)
boot				
boot1-photo-colour	4.15 (1.18)	2.95 (1.05)	2.5 (1.47)	3.27 (1.2)
boot1-photo-grey	4.7 (0.47)	2.93 (1.36)	2.82 (1.6)	2.76 (0.94)
boot2-photo-colour	4.6 (0.6)	2.95 (1.2)	2.1 (0.94)	4.25 (0.85)
boot2-photo-grey	4.7 (0.66)	3.41 (1.01)		3.75 (1.02)
boot3-photo-colour	4.29 (0.96)	2.95 (1.09)	2.45 (1.44)	4.1 (1)
boot3-photo-grey	4.52 (0.67)	2.95 (1)		3.79 (1.22)
bottle				
bottle1-photo-colour	4.57 (0.73)	3.33 (1.35)	3.14 (1.2)	3.7 (0.98)
bottle1-photo-grey	4.85 (0.37)	2.45 (1.1)		3.15 (0.99)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
bottle2-photo-colour	4.82 (0.5)	2.25 (1.07)	2.25 (1.21)	2.81 (1.5)
bottle2-photo-grey	4.48 (1.12)	1.85 (0.93)		3.2 (1.11)
bottle3-photo-colour	4.45 (0.91)	1.92 (1.28)	2.04 (1.23)	2.55 (1.18)
bottle3-photo-grey	4.5 (0.96)	1.5 (1.01)		3.33 (1.43)
bowl				
bowl1-photo-colour	4.68 (0.78)	2.25 (1.16)	1.65 (0.99)	3.7 (1.13)
bowl1-photo-grey	4.81 (0.51)	2 (0.86)		3.03 (1)
bowl2-photo-colour	4.8 (0.41)	2.5 (1.1)	1.9 (1.55)	3.43 (1.16)
bowl2-photo-grey	4.75 (0.44)	1.9 (1.45)		4.05 (1.09)
bowl3-photo-colour	4.62 (0.67)	2.19 (1.25)	1.9 (1.45)	3.27 (1.28)
bowl3-photo-grey	4.52 (0.92)	1.73 (0.98)		3.09 (1.06)
bread				
bread1-photo-colour	4.82 (0.66)	3.45 (0.94)	3.85 (1.04)	3.8 (1.15)
bread1-photo-grey	4.52 (0.87)	2.8 (1.06)		3.07 (1.01)
bread2-photo-colour	5 (0)	3.2 (1.24)	3.75 (1.02)	3.81 (1.21)
bread2-photo-grey	4.9 (0.31)	2.35 (1.31)		3.36 (1.18)
bread3-photo-colour	4.71 (0.64)	3.48 (1.36)	3.29 (1.27)	3.41 (1.37)
bread3-photo-grey	4.67 (0.64)	2.59 (0.96)		3.26 (1.18)
broom				
broom1-photo-colour	4.09 (1.19)	2.3 (1.08)	2.55 (1.19)	3.35 (1.35)
broom1-photo-grey	4.29 (1.01)	2.55 (1.05)		3.27 (1.14)
broom2-photo-colour	4.2 (1.15)	2.85 (1.09)	3.3 (1.3)	3.14 (1.46)
broom2-photo-grey	4.05 (0.76)	2.1 (1.02)		3.41 (1.37)
broom3-photo-colour	4 (1)	2.38 (1.02)	2.67 (1.28)	3.73 (1.45)
broom3-photo-grey	4.12 (1.12)	2.27 (1.24)		3.32 (1.36)
brush				
brush1-photo-colour	4.33 (0.91)	3 (1.17)	2.15 (1.42)	2.27 (1.36)
brush1-photo-grey	4.41 (0.8)	2.8 (0.83)		2.85 (1.39)
brush2-photo-colour	4.1 (0.97)	3.5 (1.24)	3.65 (1.23)	2.18 (1.22)
brush2-photo-grey	4 (0.97)	3.2 (1.11)		2.24 (1.3)
brush3-photo-colour	4.38 (0.97)	3.13 (1.18)	2.87 (1.42)	2.64 (1.29)
brush3-photo-grey	4 (0.95)	3.14 (1.06)		2.55 (1.22)
button				
button1-photo-colour	4.45 (1)	2.77 (1.27)	1.91 (1.23)	3.05 (1.54)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
button1-photo-grey	4.57 (0.77)	2.33 (1.02)		4.05 (1.23)
button2-photo-colour	4.71 (0.78)	1.55 (0.89)	1.45 (0.83)	4.45 (1.1)
button2-photo-grey	4.82 (0.5)	1.5 (0.83)		4.24 (1.34)
button3-photo-colour	4.68 (0.57)	2.14 (1.46)	1.77 (1.15)	3.67 (1.28)
button3-photo-grey	4.55 (0.74)	2 (1.5)		3.55 (1.14)
cake				
cake1-photo-colour	4.62 (0.59)	3.35 (1.04)	2.5 (1.57)	3.27 (1.01)
cake1-photo-grey	4.73 (0.63)	3.9 (0.91)		3.35 (1.27)
cake2-photo-colour	4.5 (0.69)	4.4 (0.99)	2.6 (1.93)	3.68 (1.21)
cake2-photo-grey	4.7 (0.57)	4 (0.92)		2.62 (1.2)
cake3-photo-colour	4.79 (0.41)	3.7 (0.97)	2.91 (1.2)	3.14 (1.32)
cake3-photo-grey	4.19 (1.03)	2.9 (1)		2.55 (1.1)
camel				
camel1-photo-colour	3.4 (1.57)	3.6 (1.19)	3.75 (1.21)	4.18 (1.14)
camel1-photo-grey	3.65 (1.6)	3.8 (1.03)	2.36 (1.5)	3.95 (1.07)
camel2-photo-colour	3.9 (1.17)	3.52 (1.44)	3.95 (1.16)	3.85 (0.93)
camel2-photo-grey	3.3 (1.72)	3.82 (1.01)		3.35 (0.99)
camel3-photo-colour	3.29 (1.42)	3.5 (1.01)	4.41 (0.96)	4.29 (0.9)
camel3-photo-grey	2.87 (1.55)	3.18 (1.01)		3.92 (1.14)
candle				
candle1-photo-colour	4.6 (0.68)	2.33 (0.92)	3 (1.21)	3.67 (1.06)
candle1-photo-grey	4.45 (0.89)	1.8 (0.89)		2.95 (1.09)
candle2-photo-colour	4.71 (0.64)	2.59 (1.01)	2.45 (1.5)	3.65 (1.14)
candle2-photo-grey	4.35 (0.93)	1.76 (1)		3.8 (1.15)
candle3-photo-colour	3.5 (1.06)	2.95 (1)	2.55 (1.14)	3 (0.98)
candle3-photo-grey	4.14 (0.79)	2.41 (0.85)		2.62 (1.28)
cannon				
cannon1-photo-colour	3.2 (1.51)	3.32 (1.13)	3.32 (1.21)	3.7 (0.92)
cannon1-photo-grey	3.1 (1.49)	3.43 (0.87)		3.8 (1.36)
cannon2-photo-colour	3.9 (1.34)	3.3 (1.03)	3.75 (1.02)	3.95 (0.69)
cannon2-photo-grey	3.45 (1.44)	2.8 (0.95)		3.71 (1.19)
cannon3-photo-colour	3.32 (1.59)	3.55 (1.1)	2.91 (1.34)	3.1 (1.26)
cannon3-photo-grey	3 (1.38)	3.24 (1.13)		2.59 (1.5)
carrot				

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
carrot1-photo-colour	4.91 (0.29)	3 (1.12)	4.15 (1.27)	4.25 (0.85)
carrot1-photo-grey	4.57 (0.6)	2.25 (0.97)		2.87 (1.04)
carrot2-photo-colour	4.9 (0.31)	3.1 (1.17)	4.6 (0.68)	4.1 (1.26)
carrot2-photo-grey	4.75 (0.55)	2.35 (0.99)		3.55 (1.1)
carrot3-photo-colour	4.71 (0.56)	3.52 (1.21)	4.19 (0.98)	3.91 (1.11)
carrot3-photo-grey	4.6 (0.58)	3.23 (1.07)		3.64 (1.09)
celery				
celery1-photo-colour	4.4 (0.99)	2.33 (1.35)	4.5 (0.89)	4.1 (1)
celery1-photo-grey	4 (1.38)	1.95 (1.23)		2.91 (1.23)
celery2-photo-colour	4.15 (1.27)	3.32 (0.99)	4.55 (0.91)	3.9 (1.12)
celery2-photo-grey	3.9 (1.12)	3 (1.26)		2.75 (1.21)
celery3-photo-colour	3.22 (1.54)	3 (1.02)	4.27 (1.28)	4.04 (1.16)
celery3-photo-grey	3.71 (1.38)	3.14 (1.08)		2.86 (1.35)
chain				
chain1-photo-colour	4.1 (1.25)	2.3 (1.12)	3.1 (1.37)	3.67 (1.15)
chain1-photo-grey	3.55 (1.39)	2.1 (1.17)		3.27 (1.55)
chain2-photo-colour	4.2 (0.89)	2.55 (1.1)	2.82 (1.26)	3.6 (1.35)
chain2-photo-grey	3.75 (1.16)	1.9 (0.89)		4.55 (0.69)
chain3-photo-colour	3.77 (1.41)	2.7 (0.88)	3.39 (1.27)	4.08 (1.21)
chain3-photo-grey	3.67 (1.39)	2.32 (1.09)		3.86 (1.11)
chair				
chair1-photo-colour	4.76 (0.77)	3.15 (1.14)	2.9 (1.45)	3.53 (1.28)
chair1-photo-grey	4.64 (0.66)	3.1 (0.72)		3.4 (1.31)
chair2-photo-colour	4.6 (0.75)	3.1 (1.25)	3 (1.56)	3.77 (1.23)
chair2-photo-grey	4.95 (0.22)	2.7 (1.08)		3.38 (1.36)
chair3-photo-colour	5 (0)	2.13 (1.18)	3.04 (1.33)	4.32 (0.89)
chair3-photo-grey	4.76 (0.54)	2.57 (0.98)		3.41 (1.22)
cherry				
cherry1-photo-colour	4.4 (0.94)	2.55 (1.32)	4.3 (0.8)	4.18 (1.18)
cherry1-photo-grey	4.2 (1.11)	2.27 (1.11)	1.73 (1.1)	3.48 (1.03)
cherry2-photo-colour	4.3 (0.8)	2.19 (1.36)	4.05 (1.2)	4.7 (0.57)
cherry2-photo-grey	4.45 (0.83)	2.36 (1)		3.25 (1.25)
cherry3-photo-colour	4.14 (1.06)	2.59 (1.18)	4.41 (0.96)	4.48 (0.75)
cherry3-photo-grey	3.87 (1.18)	2.23 (0.97)		3.62 (1.06)

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
chicken				
chicken1-photo-colour	4.33 (0.96)	4.14 (1.06)	3.81 (1.12)	3.85 (0.93)
chicken1-photo-grey	4.25 (1.12)	3.59 (1.14)		2.9 (1.17)
chicken2-photo-colour	4.36 (1.05)	4 (0.92)	3.35 (1.04)	4.43 (0.98)
chicken2-photo-grey	4.38 (0.86)	3.4 (1.47)		3.7 (1.13)
chicken3-photo-colour	4.14 (0.94)	4.29 (0.86)	3.79 (1.35)	3.59 (1.1)
chicken3-photo-grey	4 (1.11)	3.59 (1.26)		3.29 (1.15)
chisel				
chisel1-photo-colour	3.91 (1.27)	2.8 (1.15)	2.4 (1.31)	3.7 (1.17)
chisel1-photo-grey	3.86 (1.2)	2.45 (1.15)		3.33 (1.35)
chisel2-photo-colour	3.45 (1.36)	3.6 (1.1)	2.7 (1.45)	3.62 (1.4)
chisel2-photo-grey	2.6 (1.5)	2.5 (1.4)		3.91 (0.75)
chisel3-photo-colour	3.1 (1.58)	3.14 (1.11)	2.62 (1.2)	3.73 (1.16)
chisel3-photo-grey	2.62 (1.31)	2.27 (1.12)		3.59 (1.33)
clock				
clock1-photo-colour	5 (0)	3.18 (0.96)	2.82 (1.22)	3.55 (1.19)
clock1-photo-grey	4.73 (0.74)	3.24 (1)		4.2 (0.83)
clock2-photo-colour	4.71 (0.64)	3.15 (1.14)	3.2 (1.51)	2.9 (1.12)
clock2-photo-grey	4.77 (0.43)	2.6 (0.88)		2.9 (1.3)
clock3-photo-colour	4.82 (0.39)	3.24 (1.04)	2.38 (0.97)	3.24 (1.34)
clock3-photo-grey	4.36 (0.9)	3.42 (1.38)		2.59 (1.18)
cloud				
cloud1-photo-colour	4.2 (1.15)	2.9 (1.25)	3.15 (1.04)	4.05 (0.9)
cloud1-photo-grey	4.75 (0.91)	2.7 (1.26)	3.64 (1.29)	4 (1.14)
cloud2-photo-colour	4.55 (0.94)	2.86 (1.2)	2.9 (1.34)	3.15 (1.35)
cloud2-photo-grey	4.2 (1.2)	2.23 (1.15)		2.55 (1.36)
cloud3-photo-colour	4.48 (0.81)	1.77 (0.92)	4.32 (0.95)	3.67 (1.32)
cloud3-photo-grey	4.14 (1.17)	2.18 (1.1)		3.71 (1.04)
comb				
comb1-photo-colour	4.55 (0.74)	2.45 (1.23)	2.25 (1.48)	3.45 (1.23)
comb1-photo-grey	4.67 (0.73)	2.05 (1)		3.6 (1.25)
comb2-photo-colour	4.5 (0.89)	2.85 (1.27)	1.8 (1.32)	4 (0.89)
comb2-photo-grey	4.6 (0.75)	1.85 (1.04)		3.91 (0.87)
comb3-photo-colour	4.57 (0.6)	2.33 (1.06)	1.81 (1.33)	3.82 (1.22)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
comb3-photo-grey	4.54 (0.78)	1.87 (0.97)		4.23 (0.87)
corn				
corn1-photo-colour	4.27 (0.88)	3.8 (0.95)	4.6 (0.75)	4.25 (1.07)
corn1-photo-grey	4.71 (0.56)	2.9 (1.29)		3.4 (1.13)
corn2-photo-colour	4.55 (0.69)	3.5 (1.19)	4.55 (0.89)	4.29 (1.27)
corn2-photo-grey	4.45 (0.69)	3 (1.34)		4.18 (0.73)
corn3-photo-colour	4.43 (0.81)	3.57 (1.16)	4.62 (0.59)	4 (1.2)
corn3-photo-grey	4.42 (0.83)	3.05 (1.09)		4.09 (0.92)
crown				
crown1-photo-colour	3.77 (1.41)	4.57 (0.6)	4.38 (0.67)	4.2 (0.95)
crown1-photo-grey	4.2 (1.2)	4.27 (1.08)		3.4 (1.19)
crown2-photo-colour	3.91 (1.48)	4 (1.08)	3.1 (1.41)	4.19 (1.03)
crown2-photo-grey	3.95 (1.47)	3.5 (0.95)		3.55 (1.23)
crown3-photo-colour	3.27 (1.45)	4.56 (0.71)	4 (1.32)	3.14 (1.21)
crown3-photo-grey	3.45 (1.57)	3.68 (1.21)		3 (1.05)
deer				
deer1-photo-colour	3.55 (1.54)	3.55 (1.06)	3.64 (1.18)	3.15 (1.31)
deer1-photo-grey	3.47 (1.5)	3.29 (0.9)		3.65 (1.18)
deer2-photo-colour	3.43 (1.4)	3.8 (1.15)	4.4 (0.94)	3.5 (1)
deer2-photo-grey	3.36 (1.26)	3.3 (1.22)		3.1 (1.51)
deer3-photo-colour	3.45 (1.34)	3.64 (1.05)	3.64 (1.26)	3.86 (0.91)
deer3-photo-grey	3.45 (1.3)	3.42 (1.1)		3.09 (1.16)
doll				
doll1-photo-colour	4.4 (1.23)	2.91 (1.06)	2.73 (1.35)	2.95 (1)
doll1-photo-grey	3.8 (1.37)	3.05 (0.74)		3.2 (1.11)
doll2-photo-colour	4.29 (1.06)	4.05 (1.05)	2.8 (1.28)	3.35 (1.14)
doll2-photo-grey	4.09 (1.19)	3.4 (1.14)		3.43 (1.33)
doll3-photo-colour	4 (1.2)	3.57 (1.25)	3.05 (1.12)	2.57 (1.16)
doll3-photo-grey	3.55 (1.3)	3.85 (1.35)		2.68 (1.21)
donkey				
donkey1-photo-colour	3.73 (1.2)	3.55 (1.1)	4.05 (0.94)	4.45 (0.76)
donkey1-photo-grey	4.14 (1.01)	3.55 (1.19)		3.93 (0.94)
donkey2-photo-colour	4.15 (1.09)	3.7 (1.13)	3.6 (1.1)	4.43 (0.98)
donkey2-photo-grey	3.65 (1.27)	3.05 (1.23)		4.23 (0.97)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
donkey3-photo-colour	3.95 (1.12)	3.62 (1.16)	3.19 (1.21)	4.32 (0.72)
donkey3-photo-grey	3.46 (1.53)	3.14 (1.17)		4.09 (0.97)
door				
door1-photo-colour	4.95 (0.22)	2.33 (1.12)	2.8 (1.51)	3.43 (1.12)
door1-photo-grey	4.8 (0.52)	2.2 (1.15)		3.27 (1.16)
door2-photo-colour	4.8 (0.52)	3.68 (1.09)	1.91 (1.44)	2.9 (1.17)
door2-photo-grey	4.5 (0.89)	2.71 (1.06)		2.9 (0.97)
door3-photo-colour	4.82 (0.39)	2.32 (0.95)	1.82 (1.05)	3.96 (1.08)
door3-photo-grey	4.81 (0.51)	1.82 (0.8)		3.57 (0.98)
dress				
dress1-photo-colour	4.09 (1.11)	3.45 (0.83)	1.6 (1.1)	2.8 (1.44)
dress1-photo-grey	4.38 (0.74)	2.7 (0.98)		2.37 (1.03)
dress2-photo-colour	4.7 (0.8)	3.9 (0.97)	2.2 (1.64)	2.19 (1.21)
dress2-photo-grey	4.15 (1.27)	3.4 (1.31)		2.82 (1.3)
dress3-photo-colour	4 (1.05)	2.71 (1.01)	1.62 (0.97)	2.77 (0.97)
dress3-photo-grey	4.42 (0.97)	2.05 (0.95)		3.27 (0.88)
dresser				
dresser1-photo-colour	4.25 (1.25)	2.4 (1.27)	2.15 (1.14)	3.14 (1.32)
dresser1-photo-grey	4.8 (0.52)	2.27 (1.08)	2.27 (1.42)	3.1 (1.34)
dresser2-photo-colour	4.5 (0.69)	2.76 (0.89)	1.62 (1.12)	2.65 (1.31)
dresser2-photo-grey	4.6 (0.6)	2.77 (0.92)		2.9 (1.45)
dresser3-photo-colour	4.33 (0.8)	3 (1.02)	2.68 (1.17)	3 (1.26)
dresser3-photo-grey	4.55 (0.6)	2.95 (1.05)		3.52 (1.29)
drum				
drum1-photo-colour	3.9 (1.41)	3.59 (1.05)	2.91 (1.31)	3.6 (1.19)
drum1-photo-grey	4 (1.02)	3.48 (0.87)		3.95 (0.89)
drum2-photo-colour	3.81 (1.36)	3.5 (1.1)	2.45 (1.23)	3.55 (1.15)
drum2-photo-grey	4.05 (1.25)	3.45 (1.05)		4.1 (1.18)
drum3-photo-colour	3.86 (1.46)	3.24 (1.18)	2.76 (1.26)	3.9 (0.89)
drum3-photo-grey	3.96 (1.22)	3.25 (1.03)		3.23 (1.27)
duck				
duck1-photo-colour	4.71 (0.64)	3.75 (1.25)	3.85 (1.18)	4.03 (0.96)
duck1-photo-grey	4.09 (1.23)	3.9 (1.07)		3.55 (0.94)
duck2-photo-colour	4.35 (0.99)	4.65 (0.75)	3.75 (1.41)	4.23 (0.97)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
duck2-photo-grey	4.4 (0.82)	4.1 (0.91)		3.67 (1.32)
duck3-photo-colour	4.04 (1.12)	3.23 (1.23)	3.59 (1.01)	3.59 (1.18)
duck3-photo-grey	4.62 (0.59)	3.38 (1.07)		3.18 (1.14)
eagle				
eagle1-photo-colour	3.65 (1.5)	3.87 (1.14)	4.3 (1.17)	4 (1.05)
eagle1-photo-grey	3.15 (1.63)	2.95 (1.19)		3.23 (1.19)
eagle2-photo-colour	3.4 (1.5)	4.05 (1.33)	4.23 (1.07)	3.95 (1.15)
eagle2-photo-grey	3.4 (1.19)	3.71 (1.01)		3.75 (1.07)
eagle3-photo-colour	3.13 (1.46)	3.5 (1.01)	4.36 (0.95)	4.46 (0.88)
eagle3-photo-grey	3.14 (1.35)	3.73 (1.03)		3.48 (1.03)
fence				
fence1-photo-colour	4.5 (0.74)	2.25 (1.07)	2.6 (1.23)	3.5 (1.05)
fence1-photo-grey	4.57 (0.68)	1.95 (1)		3.23 (0.94)
fence2-photo-colour	4.75 (0.55)	3.1 (1.02)	2.75 (1.48)	3.33 (1.2)
fence2-photo-grey	4.5 (0.61)	2.9 (1.41)		2.86 (1.28)
fence3-photo-colour	4.57 (0.87)	3.05 (1.43)	2.52 (1.36)	3.86 (1.21)
fence3-photo-grey	4.46 (0.83)	2.41 (1.26)		3.36 (1.29)
fish				
fish1-photo-colour	4.62 (0.59)	3.6 (1.14)	3 (1.41)	3.53 (1.17)
fish1-photo-grey	4 (1.11)	3.05 (0.94)		3.3 (1.49)
fish2-photo-colour	4.2 (0.95)	4.05 (0.94)	2.9 (1.37)	3.23 (1.41)
fish2-photo-grey	4.4 (0.75)	4 (0.92)		3.05 (1.32)
fish3-photo-colour	4.38 (1.1)	3.32 (1.25)	2.91 (1.38)	3.5 (1.34)
fish3-photo-grey	4.1 (0.89)	3.29 (1.15)		3.95 (0.95)
flag				
flag1-photo-colour	4.62 (0.8)	2.4 (1.1)	2.1 (1.33)	2.9 (1.42)
flag1-photo-grey	3.73 (1.39)	2.15 (0.88)		3.3 (1.38)
flag2-photo-colour	4.05 (1.1)	2.85 (1.35)	4.25 (1.25)	2.95 (1.46)
flag2-photo-grey	4.6 (0.68)	2.25 (0.91)		2.48 (1.33)
flag3-photo-colour	4.28 (1.02)	2.5 (1.1)	2.09 (1.31)	3.09 (1.19)
flag3-photo-grey	3.81 (1.17)	2.29 (0.96)		2.73 (1.52)
flower				
flower1-photo-colour	4.45 (1)	3.8 (1.36)	3.1 (1.41)	3.5 (1.34)
flower1-photo-grey	4.65 (0.67)	3.37 (1.03)	1.82 (1.08)	3.14 (1.31)

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
flower2-photo-colour	4.25 (1.07)	3.62 (1.16)	1.81 (0.98)	3.2 (1.2)
flower2-photo-grey	3.9 (1.21)	3.95 (1)		2.4 (0.99)
flower3-photo-colour	4.14 (0.85)	3.82 (0.85)	3.05 (1.33)	3.14 (1.42)
flower3-photo-grey	4.22 (0.8)	3.59 (1.01)		2.92 (1.1)
flute				
flute1-photo-colour	3.35 (1.23)	3.8 (1.03)	4.1 (1.07)	3.95 (1.24)
flute1-photo-grey	3.3 (1.45)	3.25 (1.16)		3.77 (1.11)
flute2-photo-colour	2.55 (1.39)	2.86 (1.08)	2.55 (1.26)	2.8 (1.36)
flute2-photo-grey	2.95 (1.5)	2.48 (1.17)		3.05 (1.43)
flute3-photo-colour	3.04 (1.72)	3.32 (0.95)	3.59 (1.14)	3.92 (1.35)
flute3-photo-grey	3.05 (1.75)	2.95 (1.05)		4.05 (1.07)
foot				
foot1-photo-colour	4.95 (0.22)	2.8 (1.24)	3.3 (1.42)	4.27 (0.98)
foot1-photo-grey	5 (0)	2.97 (0.93)	1.82 (0.98)	4.05 (0.8)
foot2-photo-colour	4.85 (0.49)	2.62 (1.36)	2.29 (1.27)	3.55 (1.32)
foot2-photo-grey	4.65 (0.81)	3.36 (1.14)		3.2 (1.2)
foot3-photo-colour	4.81 (0.51)	2.18 (1.01)	2.95 (1.36)	3.9 (1.37)
foot3-photo-grey	4.91 (0.43)	2.59 (1.14)		3.96 (1.12)
frog				
frog1-photo-colour	4.2 (1.32)	3.86 (1.28)	3.77 (1.07)	3.95 (1.05)
frog1-photo-grey	4.27 (0.87)	3.9 (0.94)		3.75 (0.97)
frog2-photo-colour	4.05 (1.16)	3.85 (1.09)	3.95 (1.1)	4 (1.03)
frog2-photo-grey	4 (1.31)	3.55 (1.19)		3.81 (1.25)
frog3-photo-colour	4 (1.07)	3.86 (1.15)	3.71 (1.38)	4.14 (0.96)
frog3-photo-grey	4.05 (1.13)	3.88 (1.2)		3.32 (0.99)
giraffe				
giraffe1-photo-colour	4.43 (1.03)	4.1 (1.07)	4.85 (0.37)	4.5 (0.97)
giraffe1-photo-grey	3.41 (1.59)	3.65 (1.09)		4.15 (1.18)
giraffe2-photo-colour	3.5 (1.36)	4.25 (0.91)	4.75 (0.55)	4.59 (0.96)
giraffe2-photo-grey	4 (1.41)	3.8 (0.77)		4.24 (1.09)
giraffe3-photo-colour	3.83 (1.43)	3.78 (1.09)	4.39 (0.94)	4.86 (0.47)
giraffe3-photo-grey	4.14 (1.2)	3.71 (1.19)		4.18 (1.05)
glasses				
glasses1-photo-colour	4.57 (0.6)	2.6 (1.19)	1.95 (1.32)	3.87 (1.04)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
glasses1-photo-grey	4.64 (0.85)	2.1 (0.79)		4 (0.97)
glasses2-photo-colour	4.8 (0.41)	2.35 (1.23)	2.35 (1.63)	4.09 (0.92)
glasses2-photo-grey	4.65 (0.81)	2.15 (1.04)		3.71 (1.45)
glasses3-photo-colour	4.88 (0.45)	2.68 (0.84)	2.18 (1.14)	3.05 (1.09)
glasses3-photo-grey	4.33 (0.91)	2.43 (1.03)		3.32 (1.21)
goat				
goat1-photo-colour	4.19 (1.03)	3.9 (1.07)	3.8 (0.95)	4 (1.05)
goat1-photo-grey	3.23 (1.31)	3.75 (0.55)		3.7 (1.08)
goat2-photo-colour	3.6 (1.1)	4 (0.97)	3.25 (1.45)	4.09 (0.97)
goat2-photo-grey	3.95 (1)	3.7 (1.03)		3.9 (1.14)
goat3-photo-colour	3.38 (1.53)	3.91 (1.04)	3.83 (1.15)	3.32 (1.13)
goat3-photo-grey	3.43 (1.21)	3.81 (1.25)		2.86 (1.21)
grapes				
grapes1-photo-colour	4.65 (0.75)	3.53 (1.14)	4.2 (0.89)	3.76 (1)
grapes1-photo-grey	4.7 (0.57)	2.85 (1.31)		3 (1.31)
grapes2-photo-colour	4.45 (0.76)	3.82 (1.18)	3.64 (1.14)	3.25 (1.16)
grapes2-photo-grey	4.2 (1.01)	3.48 (0.98)		3 (0.97)
grapes3-photo-colour	4.61 (0.66)	3.32 (0.99)	3.68 (0.95)	3.96 (1.04)
grapes3-photo-grey	4.43 (0.68)	3.45 (0.91)		3.14 (1.01)
guitar				
guitar1-photo-colour	4.15 (1.27)	3.35 (1.09)	3.05 (1.23)	4.45 (0.86)
guitar1-photo-grey	4.35 (0.99)	3.3 (1.02)	2 (1.1)	3.9 (0.7)
guitar2-photo-colour	4.45 (0.89)	2.86 (1.2)	2.1 (1)	4.3 (1.03)
guitar2-photo-grey	4.25 (1.12)	3.5 (1.1)		3.85 (1.09)
guitar3-photo-colour	3.67 (1.46)	2.86 (1.04)	3.32 (1.25)	3.95 (1.16)
guitar3-photo-grey	4.3 (0.93)	3.05 (0.72)		4.17 (1.05)
hammer				
hammer1-photo-colour	4.5 (0.89)	2.4 (1.22)	3.7 (1.22)	4.62 (0.5)
hammer1-photo-grey	4.5 (0.76)	2.1 (1.02)		3.5 (1.19)
hammer2-photo-colour	4.75 (0.55)	2.95 (1.05)	3.23 (1.19)	4.4 (0.88)
hammer2-photo-grey	4.55 (0.83)	2.43 (1.12)		4.55 (1)
hammer3-photo-colour	4.32 (0.95)	2.77 (0.97)	2.68 (1.52)	4.25 (1.03)
hammer3-photo-grey	4.29 (1.01)	2.36 (0.85)		3.71 (1.19)
hand				

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
hand1-photo-colour	4.9 (0.4)	3.33 (0.91)	3.14 (1.56)	4.6 (0.75)
hand1-photo-grey	5 (0)	3.14 (1.28)		4.2 (0.89)
hand2-photo-colour	4.73 (0.88)	3.4 (1.23)	3.15 (1.39)	4.62 (0.67)
hand2-photo-grey	5 (0)	3.05 (1.32)		3.9 (0.91)
hand3-photo-colour	4.86 (0.64)	3.62 (1.44)	2.92 (1.28)	2.73 (1.24)
hand3-photo-grey	4.86 (0.47)	3.24 (1.04)		3.43 (1.25)
harp				
harp1-photo-colour	3.25 (1.77)	3.7 (1.02)	3.55 (1.32)	3.81 (0.75)
harp1-photo-grey	2.75 (1.48)	3.15 (1.31)		3.68 (1.09)
harp2-photo-colour	3.15 (1.57)	4.09 (1.06)	2.91 (1.27)	4.15 (0.88)
harp2-photo-grey	3.4 (1.14)	3.14 (1.06)		4.25 (0.85)
harp3-photo-colour	2.64 (1.47)	3.41 (0.96)	2.91 (1.31)	4.46 (0.78)
harp3-photo-grey	3 (1.58)	3.09 (0.92)		3.76 (1.22)
horse				
horse1-photo-colour	4.27 (0.94)	3.45 (1.23)	2.45 (1.15)	4.3 (0.86)
horse1-photo-grey	4.48 (0.93)	3.35 (0.99)		3.73 (0.94)
horse2-photo-colour	4.4 (0.88)	3.75 (0.97)	3.55 (1)	3.86 (1.06)
horse2-photo-grey	3.9 (1.17)	3.55 (1.32)		3.68 (1.17)
horse3-photo-colour	4.14 (1.06)	3.62 (1.2)	3.19 (1.33)	4.14 (1.13)
horse3-photo-grey	3.96 (1.37)	3.23 (1.19)		3.73 (0.98)
house				
house1-photo-colour	4.57 (0.98)	2.95 (1.05)	2.15 (1.23)	2.57 (1.17)
house1-photo-grey	4.5 (1.06)	2.45 (0.89)		2.8 (1.4)
house2-photo-colour	4.75 (0.44)	3.55 (1.19)	2.6 (1.6)	2.86 (0.83)
house2-photo-grey	4.75 (0.55)	3.35 (0.93)		2.33 (1.11)
house3-photo-colour	4.67 (0.87)	4.09 (1.2)	3.04 (1.52)	2.64 (1.26)
house3-photo-grey	4.29 (1.01)	3.95 (1.32)		2 (0.93)
iron				
iron1-photo-colour	4.37 (1.07)	3.71 (1.06)	2.9 (1.45)	3.95 (1.39)
iron1-photo-grey	4.9 (0.45)	2.95 (1.25)		4.3 (0.98)
iron2-photo-colour	4.59 (0.73)	3.45 (1.28)	2.65 (1.57)	3.9 (1.45)
iron2-photo-grey	4.33 (1.32)	3.75 (1.37)		4.2 (0.83)
iron3-photo-colour	4.14 (0.99)	2.72 (1.17)	2.48 (1.42)	2.73 (1.39)
iron3-photo-grey	4.27 (1.12)	1.95 (1.05)		3.9 (0.83)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
jacket				
jacket1-photo-colour	4.5 (0.89)	3.15 (1.14)	2.95 (1.36)	2.41 (1.14)
jacket1-photo-grey	4.8 (0.41)	3.33 (1.06)	2.27 (1.56)	2.52 (1.12)
jacket2-photo-colour	4.55 (0.51)	2.76 (0.89)	2.71 (1.27)	3 (1.21)
jacket2-photo-grey	4.7 (0.47)	3.14 (0.94)		3.1 (1.17)
jacket3-photo-colour	4.43 (0.75)	2.55 (0.91)	1.73 (0.94)	2.71 (1.19)
jacket3-photo-grey	4.32 (0.84)	2.91 (0.9)		3.29 (1.16)
kettle				
kettle1-photo-colour	4.37 (1.22)	3.05 (0.8)	2.81 (1.54)	3.3 (1.03)
kettle1-photo-grey	4.75 (0.64)	2.68 (1.25)		2.7 (1.17)
kettle2-photo-colour	4.5 (0.8)	2.85 (1.23)	2.6 (1.05)	3.1 (1.41)
kettle2-photo-grey	4.29 (1.06)	2.75 (1.29)		2.6 (1.23)
kettle3-photo-colour	4.09 (1.19)	2.64 (1.11)	2.2 (1.08)	2.14 (1.13)
kettle3-photo-grey	4.5 (0.67)	2.32 (1.04)		2.05 (0.97)
kite				
kite1-photo-colour	4.25 (1.16)	3.1 (1.06)	2.45 (1.67)	4.05 (0.74)
kite1-photo-grey	4.2 (1.06)	2.4 (1.23)		3.32 (1.13)
kite2-photo-colour	4.2 (1.11)	3.05 (1.17)	2 (1.38)	3.85 (0.99)
kite2-photo-grey	3.6 (1.05)	2.33 (1.06)		3.45 (1.15)
kite3-photo-colour	3.23 (1.45)	3.48 (0.73)	2.09 (1.5)	3.08 (1.35)
kite3-photo-grey	3.29 (1.35)	3.14 (0.77)		2.14 (1.01)
knife				
knife1-photo-colour	4.57 (0.86)	2.67 (0.91)	3.43 (0.93)	3.5 (1.15)
knife1-photo-grey	4.95 (0.22)	2.32 (1.09)		3.6 (0.99)
knife2-photo-colour	4.91 (0.29)	2.6 (0.82)	3.15 (1.18)	3.67 (1.2)
knife2-photo-grey	4.67 (0.91)	2.2 (0.77)		3.7 (1.3)
knife3-photo-colour	4.23 (1.23)	3.21 (1.02)	3.08 (1.35)	2.14 (0.99)
knife3-photo-grey	4.27 (0.94)	2.81 (1.03)		2.62 (0.92)
ladder				
ladder1-photo-colour	4.75 (0.55)	2 (1.07)	2.95 (1.4)	3.1 (1.37)
ladder1-photo-grey	4.3 (1.06)	1.62 (0.67)		3.95 (1.05)
ladder2-photo-colour	4.52 (0.98)	1.95 (1.28)	1.8 (1.01)	4.1 (0.97)
ladder2-photo-grey	4.5 (0.74)	1.6 (0.94)		4.24 (1.09)
ladder3-photo-colour	4.5 (0.96)	2.41 (1.22)	2.91 (1.41)	3.19 (1.29)

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
ladder3-photo-grey	4.27 (0.98)	1.71 (0.95)		3.5 (1.06)
lamp				
lamp1-photo-colour	4.5 (0.86)	2.76 (0.77)	2.9 (1.14)	3.55 (1.28)
lamp1-photo-grey	4.85 (0.49)	2.41 (0.85)		3.2 (1.2)
lamp2-photo-colour	4.64 (0.9)	3.35 (0.67)	2.55 (1.28)	4.05 (1.02)
lamp2-photo-grey	4.71 (0.56)	2.95 (1)		3.6 (1.05)
lamp3-photo-colour	4.45 (1.06)	3.92 (1.1)	2.62 (1.24)	2.91 (1.15)
lamp3-photo-grey	4.77 (0.53)	3.32 (0.95)		2.81 (1.03)
leaf				
leaf1-photo-colour	4.8 (0.52)	3.55 (1.1)	3.14 (1.28)	3.35 (1.35)
leaf1-photo-grey	4.63 (0.89)	3.43 (1.21)		3.4 (1.19)
leaf2-photo-colour	4.76 (0.54)	2.85 (1.31)	3.05 (1.15)	4.1 (1.07)
leaf2-photo-grey	4.86 (0.64)	2.9 (1.02)		3.48 (1.33)
leaf3-photo-colour	4.68 (0.57)	3.1 (1.26)	2.9 (0.94)	3.38 (1.32)
leaf3-photo-grey	4.55 (0.96)	2.62 (1.24)		2.77 (1.27)
lemon				
lemon1-photo-colour	4.76 (0.7)	3.05 (1.32)	4.5 (0.76)	3.8 (1.19)
lemon1-photo-grey	4.27 (0.98)	2.45 (0.94)		3.35 (1.35)
lemon2-photo-colour	4.65 (0.59)	3.05 (1.43)	4.7 (0.57)	4.91 (0.29)
lemon2-photo-grey	4.5 (0.69)	2.95 (1.23)		3.76 (1.14)
lemon3-photo-colour	4.75 (0.61)	2.26 (1.48)	4.26 (1.18)	5 (0)
lemon3-photo-grey	3.81 (1.17)	2.62 (1.02)		3.45 (1.34)
lion				
lion1-photo-colour	4.1 (1.21)	3.64 (1.05)	4.32 (0.99)	4.7 (0.57)
lion1-photo-grey	4.2 (1)	3.33 (1.24)		4.35 (0.81)
lion2-photo-colour	4.1 (1.37)	3.75 (0.91)	4.65 (0.67)	4.65 (0.93)
lion2-photo-grey	4.05 (1.21)	3.4 (1.1)		4.29 (0.78)
lion3-photo-colour	3.95 (1.29)	3.95 (1.09)	4.32 (1.09)	4.67 (0.48)
lion3-photo-grey	3.41 (1.37)	3.75 (1.39)		3.41 (1.18)
lips				
lips1-photo-colour	4.8 (0.62)	3.23 (1.07)	3.09 (0.97)	3.8 (0.95)
lips1-photo-grey	4.87 (0.51)	3.29 (1.01)		3.8 (0.95)
lips2-photo-colour	4.86 (0.65)	2.15 (0.81)	3.25 (1.29)	4.55 (0.6)
lips2-photo-grey	4.82 (0.85)	2.15 (0.99)		4.14 (1.28)

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
lips3-photo-colour	4.82 (0.66)	2.41 (1.05)	3.27 (1.49)	4.05 (0.92)
lips3-photo-grey	4.77 (0.53)	1.96 (1.21)		3.32 (1.25)
lobster				
lobster1-photo-colour	3.45 (1.5)	4.17 (1.15)	4.15 (0.99)	4.43 (0.68)
lobster1-photo-grey	3.5 (1.5)	3.65 (1.27)		3.36 (1.09)
lobster2-photo-colour	4.24 (1.14)	4.05 (1)	4.18 (1.05)	4.15 (0.81)
lobster2-photo-grey	3.75 (1.33)	3 (1.3)		3.35 (1.23)
lobster3-photo-colour	2.95 (1.53)	3.5 (0.91)	4.27 (0.83)	4.44 (0.71)
lobster3-photo-grey	2.95 (1.4)	3.18 (0.91)		3.29 (1.31)
lock				
lock1-photo-colour	4.3 (0.99)	3.81 (1.21)	4.05 (0.92)	4.15 (1.18)
lock1-photo-grey	4.5 (0.69)	3.32 (1.25)		3.45 (1.39)
lock2-photo-colour	4.27 (1.08)	3.4 (1.19)	2.7 (1.03)	3.48 (1.4)
lock2-photo-grey	4.29 (1.1)	2.6 (1.35)		4.1 (0.97)
lock3-photo-colour	4.04 (1.02)	3.54 (1.22)	3.75 (1.26)	3.82 (1.14)
lock3-photo-grey	4.36 (0.85)	2.86 (0.94)		3.57 (1.25)
mitten				
mitten1-photo-colour	4.05 (1.25)	3.95 (0.94)	1.85 (1.14)	3.35 (1.18)
mitten1-photo-grey	4.48 (0.68)	3.2 (1.11)		3.67 (1.06)
mitten2-photo-colour	4.1 (1.17)	3.35 (0.88)	2 (1.59)	3.38 (1.32)
mitten2-photo-grey	3.75 (1.21)	2.3 (1.26)		3.82 (1.05)
mitten3-photo-colour	4.14 (0.96)	2.43 (1.33)	1.86 (1.49)	3.68 (1.36)
mitten3-photo-grey	4.33 (0.87)	1.91 (1.06)		3.77 (0.97)
monkey				
monkey1-photo-colour	3.62 (1.24)	3.95 (1.32)	3.85 (1.14)	3.47 (1.2)
monkey1-photo-grey	3.09 (1.6)	3.9 (0.97)		3.1 (1.25)
monkey2-photo-colour	3.1 (1.25)	4.15 (1.14)	4.15 (1.14)	3.59 (1.01)
monkey2-photo-grey	3.65 (1.42)	3.85 (0.99)		3.1 (1.18)
monkey3-photo-colour	3.79 (1.41)	3.36 (1.22)	3.59 (1.01)	3.55 (0.96)
monkey3-photo-grey	3.62 (1.24)	3.29 (1.15)		3.23 (1.23)
moon				
moon1-photo-colour	4.6 (0.68)	2.6 (1)	2.8 (1.15)	2.71 (1.27)
moon1-photo-grey	4.05 (1.47)	2.25 (1.02)		2.95 (1.4)
moon2-photo-colour	4.1 (1.17)	2.45 (1.14)	3.36 (1.05)	2.55 (1.28)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
moon2-photo-grey	3.55 (1.39)	1.9 (0.94)		2.6 (1.27)
moon3-photo-colour	4.26 (1.01)	2.73 (1.28)	3.32 (0.99)	2.46 (1.25)
moon3-photo-grey	4.14 (1.11)	2.09 (0.97)		2.24 (1.37)
mouse				
mouse1-photo-colour	3.68 (1.25)	3.75 (0.91)	3.6 (1.19)	3.75 (1.29)
mouse1-photo-grey	4.43 (0.87)	3.35 (1.09)		3.5 (1.17)
mouse2-photo-colour	4.15 (0.99)	3.9 (0.91)	3.7 (1.03)	3.76 (1.45)
mouse2-photo-grey	3.8 (1.24)	3.1 (1.37)		3.95 (1.05)
mouse3-photo-colour	3.76 (1.04)	3.9 (1.26)	3.33 (1.24)	3.41 (1.18)
mouse3-photo-grey	3.71 (1.43)	3.13 (1.18)		3 (1.38)
nail				
nail1-photo-colour	4.17 (1.21)	2.14 (1.31)	3.57 (1.43)	3.7 (1.75)
nail1-photo-grey	4.5 (0.95)	2.18 (1.14)		3.4 (1.7)
nail2-photo-colour	4.32 (0.95)	2.05 (1.05)	2.7 (1.22)	3.33 (1.77)
nail2-photo-grey	4.67 (0.58)	1.5 (0.76)		3.85 (1.27)
nail3-photo-colour	4.18 (1.14)	2.12 (1.17)	3.88 (1.51)	3.77 (1.69)
nail3-photo-grey	4.41 (1.05)	1.77 (1.07)		4.14 (1.39)
needle				
needle1-photo-colour	3.87 (1.22)	2.48 (1.4)	3.48 (1.08)	4.6 (0.82)
needle1-photo-grey	4.45 (1.05)	2.27 (1.24)		4.4 (0.99)
needle2-photo-colour	4.09 (1.38)	1.9 (0.85)	2.4 (1.23)	4 (1.34)
needle2-photo-grey	4.38 (0.97)	1.7 (1.03)		3.95 (1.15)
needle3-photo-colour	3.64 (1.18)	1.92 (1.32)	3.76 (1.54)	4.05 (1.05)
needle3-photo-grey	3.82 (1.22)	1.95 (0.74)		3.81 (1.29)
nose				
nose1-photo-colour	5 (0)	2.83 (1.42)	3.15 (1.76)	4 (0.84)
nose1-photo-grey	5 (0)	2.55 (1.19)		3.82 (1.26)
nose2-photo-colour	4.95 (0.22)	3.23 (1.11)	2.68 (1.52)	3.85 (1.09)
nose2-photo-grey	4.7 (0.66)	2.38 (1.02)		3.6 (0.99)
nose3-photo-colour	4.86 (0.47)	3.17 (0.94)	3.13 (1.36)	4.21 (0.98)
nose3-photo-grey	4.76 (0.7)	2.55 (1.26)		3.29 (1.19)
onion				
onion1-photo-colour	4.9 (0.31)	2.83 (1.29)	3.9 (1.33)	4.67 (0.73)
onion1-photo-grey	4.8 (0.52)	2.55 (1.15)		3.36 (1.22)

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
onion2-photo-colour	4.8 (0.7)	3.45 (1.37)	3.82 (0.96)	4.25 (1.07)
onion2-photo-grey	4.5 (0.69)	2.95 (1.2)		3.65 (0.93)
onion3-photo-colour	4.48 (0.79)	2.59 (1.22)	3.86 (1.08)	4.75 (0.68)
onion3-photo-grey	4.38 (0.92)	2.23 (0.97)		3.24 (1.18)
orange				
orange1-photo-colour	4.68 (0.65)	2.3 (1.45)	4.7 (0.73)	4.75 (0.55)
orange1-photo-grey	4.48 (0.87)	2.35 (1.04)		2.83 (1.29)
orange2-photo-colour	4.9 (0.31)	3.25 (1.37)	4.85 (0.37)	4.81 (0.68)
orange2-photo-grey	4.25 (0.91)	1.9 (1.37)		3.68 (1.04)
orange3-photo-colour	4.86 (0.36)	3.57 (1.33)	4.67 (0.58)	4.45 (1.06)
orange3-photo-grey	4.17 (1.2)	2.26 (1.14)		3.23 (1.07)
ostrich				
ostrich1-photo-colour	3.47 (1.38)	3.81 (0.68)	4.19 (1.12)	4.7 (0.47)
ostrich1-photo-grey	3.4 (1.57)	3.18 (1.01)		4.3 (0.73)
ostrich2-photo-colour	3.27 (1.35)	3.85 (0.99)	3.55 (0.89)	4.52 (0.68)
ostrich2-photo-grey	3.24 (1.41)	3.5 (1.43)		3.9 (1.07)
ostrich3-photo-colour	3.18 (1.59)	3.8 (1)	4.12 (1.01)	3.77 (1.15)
ostrich3-photo-grey	2.95 (1.7)	3.36 (1.26)		3.43 (1.25)
peach				
peach1-photo-colour	4.45 (1)	3.14 (1.13)	3.91 (1.02)	3.95 (1.1)
peach1-photo-grey	4.23 (1.1)	2.95 (1.2)		2.95 (1.39)
peach2-photo-colour	4.19 (1.29)	3.2 (1.28)	4.2 (1.01)	4.45 (0.76)
peach2-photo-grey	4.18 (1.05)	2.2 (0.95)		2.9 (1.41)
peach3-photo-colour	4.14 (1.08)	2.29 (0.96)	3.95 (1.16)	4 (1.05)
peach3-photo-grey	3.68 (1.52)	1.83 (1.01)		1.73 (0.98)
peacock				
peacock1-photo-colour	3.87 (1.36)	4.43 (0.68)	4.76 (0.44)	4.4 (0.68)
peacock1-photo-grey	3.7 (1.56)	3.59 (1.3)		3.1 (1.12)
peacock2-photo-colour	3.73 (1.49)	4.55 (0.83)	4.2 (1.06)	4 (1.1)
peacock2-photo-grey	3.52 (1.47)	3.85 (1.31)		2.55 (1.05)
peacock3-photo-colour	3.45 (1.41)	4.44 (0.65)	4.56 (0.92)	4.05 (0.95)
peacock3-photo-grey	3.5 (1.47)	3.82 (1.33)		3.24 (1.14)
peanut				
peanut1-photo-colour	4.38 (0.86)	3.2 (1.2)	4.1 (1.02)	3.97 (1.38)

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
peanut1-photo-grey	3.91 (1.11)	2.75 (1.12)		3.35 (1.31)
peanut2-photo-colour	4 (0.97)	3.45 (1.15)	4.4 (0.75)	3.95 (1.46)
peanut2-photo-grey	4.2 (0.89)	3.45 (0.94)		3.86 (1.28)
peanut3-photo-colour	4.25 (1.03)	2.83 (1.19)	4.13 (1.1)	4.05 (1.17)
peanut3-photo-grey	4.05 (1.12)	3.05 (0.86)		4.14 (1.04)
pear				
pear1-photo-colour	4.57 (0.86)	3.14 (1.24)	4.48 (0.6)	5 (0)
pear1-photo-grey	4.8 (0.41)	2.59 (1.22)		3.55 (1.23)
pear2-photo-colour	4.68 (0.78)	2.8 (1.2)	3.6 (1.1)	4.33 (0.73)
pear2-photo-grey	4.33 (1.06)	2.45 (1.05)		3.55 (1.05)
pear3-photo-colour	4.32 (0.99)	3.08 (1.19)	4.12 (1.13)	4.09 (0.92)
pear3-photo-grey	4.41 (1.05)	2.32 (1.13)		3.14 (0.96)
pencil				
pencil1-photo-colour	4.95 (0.22)	2.25 (1.55)	2.15 (1.6)	3.77 (1.19)
pencil1-photo-grey	4.5 (0.91)	1.75 (1.02)		3.85 (1.04)
pencil2-photo-colour	4.8 (0.7)	2.8 (1.54)	2.3 (1.49)	3.86 (1.13)
pencil2-photo-grey	4.9 (0.31)	2.6 (1.23)		3.71 (1.01)
pencil3-photo-colour	4.83 (0.48)	1.86 (1.04)	2.41 (1.14)	3.77 (0.97)
pencil3-photo-grey	4.52 (0.75)	1.86 (0.96)		2.64 (1.22)
penguin				
penguin1-photo-colour	4 (1.45)	3.4 (1.33)	4.5 (0.95)	4.62 (0.59)
penguin1-photo-grey	3.5 (1.4)	2.65 (1.18)		3.91 (0.87)
penguin2-photo-colour	3.85 (1.57)	3.82 (1.26)	4.55 (0.6)	4.65 (0.59)
penguin2-photo-grey	4.05 (1.15)	2.9 (1.26)		4.5 (0.69)
penguin3-photo-colour	3.59 (1.53)	3.45 (0.91)	4.18 (0.91)	4.5 (0.78)
penguin3-photo-grey	3.52 (1.44)	2.77 (0.75)		3.71 (1.1)
pepper				
pepper1-photo-colour	4.73 (0.46)	2.2 (1.36)	2.95 (1.1)	3.1 (1.48)
pepper1-photo-grey	4.71 (0.64)	2.4 (1.1)		2.77 (1.3)
pepper2-photo-colour	4.65 (0.67)	3.1 (1.29)	3.5 (1.36)	2.52 (1.66)
pepper2-photo-grey	4.55 (0.6)	2.15 (1.39)		3.05 (1.33)
pepper3-photo-colour	4.38 (0.67)	3.33 (1.32)	3 (1.18)	2.82 (1.22)
pepper3-photo-grey	4.29 (1)	2.35 (0.78)		2.95 (1.46)
piano				

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
piano1-photo-colour	4.48 (0.75)	3.9 (1.25)	3.55 (1.19)	3.87 (1.11)
piano1-photo-grey	3.95 (1.13)	3.95 (1.15)		4.2 (0.95)
piano2-photo-colour	4.15 (0.81)	4.15 (1.09)	3.6 (1.27)	4.45 (0.74)
piano2-photo-grey	4.15 (1.23)	3.4 (0.94)		4.14 (1.15)
piano3-photo-colour	4.29 (1.04)	3.77 (0.97)	3.5 (1.1)	3.86 (0.89)
piano3-photo-grey	4.14 (1.01)	3.71 (1.15)		3.95 (0.9)
pipe				
pipe1-photo-colour	3.43 (1.25)	4.1 (1.33)	3.2 (1.11)	2.8 (1.71)
pipe1-photo-grey	2.82 (1.5)	4.35 (0.88)		2.8 (1.47)
pipe2-photo-colour	3 (1.38)	2.75 (1.62)	3.55 (1.23)	2.77 (1.6)
pipe2-photo-grey	3.65 (1.23)	2.6 (1.1)		2.86 (1.53)
pipe3-photo-colour	3 (1.53)	2.26 (1.01)	3.17 (1.3)	2.86 (1.49)
pipe3-photo-grey	3.29 (1.52)	2.38 (1.12)		2.68 (1.7)
pitcher				
pitcher1-photo-colour	4.25 (1.07)	3.3 (1.26)	3.45 (1.36)	3.18 (1.4)
pitcher1-photo-grey	4.5 (0.76)	2.5 (0.9)	2.27 (1.1)	2.9 (1.18)
pitcher2-photo-colour	4.45 (0.6)	2.81 (1.17)	1.57 (1.12)	2.3 (0.98)
pitcher2-photo-grey	3.85 (1.14)	2.64 (1)		3.1 (1.25)
pitcher3-photo-colour	3.67 (1.2)	2.18 (1.14)	1.55 (0.96)	2.29 (1.01)
pitcher3-photo-grey	3.39 (1.08)	2.09 (0.81)		2.67 (1.27)
pliers				
pliers1-photo-colour	4.09 (1.31)	3 (1.08)	2.4 (1.05)	3.9 (0.97)
pliers1-photo-grey	3.95 (1.16)	3.05 (1.05)		3.4 (1.16)
pliers2-photo-colour	4.25 (0.79)	2.75 (1.33)	2.9 (1.37)	4.19 (0.98)
pliers2-photo-grey	3.6 (1.31)	2.45 (1.23)		3.73 (1.2)
pliers3-photo-colour	4.05 (1.16)	2.67 (1.06)	2.33 (1.32)	3.82 (1.05)
pliers3-photo-grey	3.58 (1.18)	2.09 (0.97)		3.82 (1.1)
plug				
plug1-photo-colour	4.05 (1.36)	2.65 (1.14)	3 (1.17)	2.09 (1.27)
plug1-photo-grey	4.45 (0.83)	2.7 (1.15)	3.18 (1.4)	2.33 (1.39)
plug2-photo-colour	4.1 (1.29)	2.19 (0.93)	2.86 (1.31)	2 (1.49)
plug2-photo-grey	4.5 (1.15)	3.18 (1.1)		2.35 (1.5)
plug3-photo-colour	3.81 (1.4)	2.5 (0.86)	3.5 (1.19)	2.24 (1.3)
plug3-photo-grey	3 (1.45)	2.73 (1.12)		2.48 (1.42)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
potato				
potato1-photo-colour	4.67 (1.03)	2.71 (1.45)	4 (1.26)	5 (0)
potato1-photo-grey	4.85 (0.37)	2.45 (1.1)		4 (1.12)
potato2-photo-colour	4.82 (0.5)	2.75 (1.12)	3.7 (1.3)	4.52 (0.81)
potato2-photo-grey	4.38 (1.24)	2.45 (1.15)		3.5 (1.24)
potato3-photo-colour	5 (0)	2.6 (1.55)	4.08 (1.08)	4.55 (0.67)
potato3-photo-grey	4.82 (0.5)	2.48 (1.08)		3.76 (1)
pumpkin				
pumpkin1-photo-colour	4.23 (1.23)	3 (1.17)	4.2 (1.28)	4.6 (0.75)
pumpkin1-photo-grey	4.1 (1.04)	2.55 (1)		3.2 (1.27)
pumpkin2-photo-colour	4 (1.08)	2.95 (1.23)	4.45 (0.83)	4.14 (1.2)
pumpkin2-photo-grey	3.5 (1.19)	2.25 (1.02)		3.18 (1.05)
pumpkin3-photo-colour	4.05 (0.97)	3.24 (1.37)	4.52 (0.98)	4.55 (0.74)
pumpkin3-photo-grey	3.79 (1.25)	2.35 (1.15)		3.55 (1.06)
rabbit				
rabbit1-photo-colour	4.35 (1.18)	3.45 (1.37)	3.64 (0.95)	3.9 (0.85)
rabbit1-photo-grey	4.3 (0.99)	3.38 (0.97)		4.05 (1.05)
rabbit2-photo-colour	4.14 (1.28)	3.2 (1.06)	3.6 (1.1)	3.7 (1.22)
rabbit2-photo-grey	4 (1.23)	3.1 (1.17)		3.43 (1.12)
rabbit3-photo-colour	4.23 (0.97)	3.95 (1.2)	3.81 (1.25)	3.9 (1.09)
rabbit3-photo-grey	4.09 (1.11)	3.8 (1.04)		3.77 (0.81)
raccoon				
raccoon1-photo-colour	3.45 (1.73)	3.53 (1.14)	4.15 (0.88)	4.33 (0.86)
raccoon1-photo-grey	3.45 (1.36)	3.25 (1.33)		4 (1.02)
raccoon2-photo-colour	3.6 (1.47)	4.09 (1.06)	4.55 (0.67)	3.95 (1.19)
raccoon2-photo-grey	3.65 (1.27)	3.62 (1.4)		4.3 (0.73)
raccoon3-photo-colour	2.83 (1.56)	3.59 (1.01)	4.32 (1.04)	4.46 (0.72)
raccoon3-photo-grey	2.86 (1.35)	3.27 (1.12)		3.62 (1.12)
ring				
ring1-photo-colour	4.62 (0.67)	2.35 (0.75)	2.9 (1.17)	3.3 (1.29)
ring1-photo-grey	4.23 (1.07)	2.7 (0.98)		3.4 (1.39)
ring2-photo-colour	2.55 (1.61)	3.45 (1.19)	3.2 (1.36)	2.09 (0.87)
ring2-photo-grey	2.75 (1.71)	2.85 (1.04)		2 (1.18)
ring3-photo-colour	4.08 (1.15)	1.82 (0.85)	2.14 (1.13)	3.27 (1.24)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
ring3-photo-grey	4 (1.14)	1.9 (0.94)		2.59 (1.3)
ruler				
ruler1-photo-colour	4.76 (0.44)	2 (1.26)	2.15 (1.39)	3.33 (1.24)
ruler1-photo-grey	4.45 (1.06)	2.05 (1.05)		3.3 (1.17)
ruler2-photo-colour	4.4 (0.82)	3 (1.59)	3.2 (1.61)	3.86 (1.08)
ruler2-photo-grey	4.55 (0.69)	2.65 (1.04)		4.1 (0.89)
ruler3-photo-colour	4.4 (1.12)	2.55 (1.14)	2.45 (1.14)	3.27 (1.49)
ruler3-photo-grey	4.1 (1.04)	2.62 (1.02)		3.68 (0.99)
screw				
screw1-photo-colour	4.8 (0.62)	3.09 (1.15)	3.27 (1.39)	4.6 (0.6)
screw1-photo-grey	4.4 (1)	3.38 (1.16)		4.9 (0.31)
screw2-photo-colour	4.67 (0.66)	3.1 (1.41)	3.4 (1.43)	4.45 (0.69)
screw2-photo-grey	4.77 (0.53)	2.6 (1.05)		4.38 (0.92)
screw3-photo-colour	4.5 (0.74)	3.05 (1.13)	3.23 (1.45)	4.48 (0.68)
screw3-photo-grey	4.09 (1.35)	2.67 (1.34)		4.41 (0.8)
seal				
seal1-photo-colour	3.8 (1.44)	3.23 (1.23)	3.59 (1.33)	3.95 (1.1)
seal1-photo-grey	3.63 (1.43)	3.48 (1.08)		3.8 (1.15)
seal2-photo-colour	3.76 (1.41)	2.95 (1.15)	4.15 (1.04)	3.95 (1.15)
seal2-photo-grey	3.68 (1.36)	3.25 (1.07)		4.38 (0.97)
seal3-photo-colour	3.64 (1.47)	3.57 (1.16)	4.1 (1.3)	3.71 (1.01)
seal3-photo-grey	3.32 (1.43)	2.84 (1.07)		3.27 (1.16)
sheep				
sheep1-photo-colour	4.3 (1.03)	3.64 (1.29)	3.64 (1.14)	3.95 (1.15)
sheep1-photo-grey	4.1 (1.16)	3.43 (0.87)		4 (0.97)
sheep2-photo-colour	4 (1.22)	3.75 (1.25)	4.3 (0.66)	3.85 (1.09)
sheep2-photo-grey	4.18 (0.91)	3.6 (1.1)		4 (0.89)
sheep3-photo-colour	3.91 (1.02)	3.41 (1.26)	3.95 (1.29)	3.95 (0.97)
sheep3-photo-grey	3.95 (1.05)	3.32 (1.11)		3.68 (1.21)
shirt				
shirt1-photo-colour	4.6 (0.75)	3.35 (1.14)	2.25 (1.45)	3.36 (1.22)
shirt1-photo-grey	4.75 (0.64)	3.17 (0.87)	2.09 (1.22)	3.71 (1.01)
shirt2-photo-colour	4.45 (0.89)	2.52 (1.29)	2.33 (1.2)	3.1 (1.41)
shirt2-photo-grey	4.75 (0.55)	3.27 (1.12)		2.6 (1.39)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
shirt3-photo-colour	4.57 (0.51)	3.09 (1.23)	1.82 (1.18)	3.38 (1.12)
shirt3-photo-grey	4.59 (0.73)	3.18 (0.66)		3.76 (1.01)
shoe				
shoe1-photo-colour	4.8 (0.61)	3.19 (1.08)	2.67 (1.2)	4.05 (1.32)
shoe1-photo-grey	4.95 (0.22)	2.64 (0.9)		3.6 (1.14)
shoe2-photo-colour	4.77 (0.69)	3.55 (0.94)	2.95 (1.1)	3.52 (1.44)
shoe2-photo-grey	4.76 (0.89)	3.05 (1.36)		3.35 (1.35)
shoe3-photo-colour	4.64 (0.73)	3.67 (1.24)	2.79 (1.56)	3.41 (1.59)
shoe3-photo-grey	4.91 (0.29)	2.95 (1.21)		3.86 (0.91)
skirt				
skirt1-photo-colour	4.15 (1.09)	3.23 (1.3)	2.3 (1.63)	2.81 (1.08)
skirt1-photo-grey	3.95 (1.36)	3.05 (1.23)		2.77 (1.15)
skirt2-photo-colour	4.15 (1.18)	3.23 (1.19)	1.68 (1.39)	2.6 (1.27)
skirt2-photo-grey	4 (1.08)	3.05 (1.02)		2.4 (0.94)
skirt3-photo-colour	3.59 (1.26)	2.95 (0.79)	1.77 (1.02)	3.25 (1.19)
skirt3-photo-grey	4 (0.95)	2.73 (0.94)		2.67 (1.06)
skunk				
skunk1-photo-colour	3.19 (1.66)	3.65 (0.93)	4.4 (0.99)	4.17 (1.34)
skunk1-photo-grey	3.09 (1.66)	3.7 (1.08)		4.15 (0.88)
skunk2-photo-colour	2.85 (1.53)	3.65 (1.04)	4.35 (0.88)	3.91 (0.87)
skunk2-photo-grey	3 (1.45)	3.35 (1.14)		3.43 (1.25)
skunk2-photo-grey2	NaN NA	NaN NA		2 NA
skunk3-photo-colour	2.54 (1.35)	3.45 (1.1)	3.55 (1.44)	3.23 (1.15)
skunk3-photo-grey	2.76 (1.45)	3.27 (1.08)		2.82 (1.18)
snail				
snail1-photo-colour	4.14 (0.94)	4.1 (0.72)	3.25 (1.21)	4.15 (1.04)
snail1-photo-grey	4.57 (0.6)	3 (1.03)		3.63 (1.03)
snail2-photo-colour	4.55 (0.69)	3.95 (1.1)	4.25 (0.72)	4.24 (1)
snail2-photo-grey	4.2 (1.2)	3.5 (1.32)		3.95 (1.09)
snail3-photo-colour	4.14 (1.24)	4.05 (1.2)	2.95 (1.2)	3.5 (1.1)
snail3-photo-grey	4.29 (1.04)	3.7 (1.06)		4.05 (0.79)
snake				
snake1-photo-colour	4.1 (1.45)	3.05 (1.17)	2.82 (0.91)	3.65 (1.09)
snake1-photo-grey	4 (1.29)	3.1 (0.94)		3.55 (1)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
snake2-photo-colour	4.38 (1.02)	2.95 (1.1)	2.6 (1.23)	3.45 (0.94)
snake2-photo-grey	3.77 (1.48)	2.65 (0.99)		3.86 (1.11)
snake3-photo-colour	3.5 (1.41)	3.05 (0.97)	2.71 (1.15)	3.57 (1.03)
snake3-photo-grey	3.5 (1.41)	3.08 (1.18)		3.5 (0.96)
snowman				
snowman1-photo-colour	4.25 (1.07)	2.41 (1.18)	3.73 (1.52)	3.05 (1.19)
snowman1-photo-grey	4.17 (0.99)	2.33 (1.15)		3.6 (1.19)
snowman2-photo-colour	4.14 (1.28)	3.4 (1.14)	4.4 (1.05)	3.9 (0.91)
snowman2-photo-grey	4.05 (1.25)	2.6 (1.14)		4.1 (0.94)
snowman3-photo-colour	4.23 (1.02)	3.05 (1.12)	4.29 (1.23)	3.33 (0.97)
snowman3-photo-grey	3.74 (1.39)	2.88 (1.19)		2.68 (1.21)
socks				
socks1-photo-colour	4.9 (0.45)	2.27 (1.12)	2 (1.38)	2.9 (1.29)
socks1-photo-grey	4.8 (0.81)	1.81 (0.87)		4.05 (1.1)
socks2-photo-colour	4.81 (0.87)	2.05 (0.94)	1.65 (0.99)	3.55 (1.39)
socks2-photo-grey	4.91 (0.43)	2 (0.86)		3.29 (1.45)
socks3-photo-colour	5 (0)	2.67 (1.35)	1.9 (1.34)	3.14 (1.2)
socks3-photo-grey	4.73 (0.88)	2.25 (1.26)		3.04 (1.19)
spider				
spider1-photo-colour	4.37 (1.13)	4 (0.77)	3.38 (1.16)	3.3 (1.45)
spider1-photo-grey	4.35 (0.99)	3.68 (1.21)		3.15 (1.27)
spider2-photo-colour	4.59 (0.85)	4.25 (0.91)	3.2 (1.2)	3.33 (1.49)
spider2-photo-grey	4.48 (0.75)	3.95 (1.1)		2.85 (1.04)
spider3-photo-colour	4.05 (1.17)	4.54 (0.93)	3.5 (1.38)	2.27 (1.39)
spider3-photo-grey	4.18 (1.22)	4.05 (1.09)		3.14 (1.24)
spoon				
spoon1-photo-colour	4.95 (0.22)	2.65 (1.42)	3.85 (1.46)	4.37 (0.85)
spoon1-photo-grey	4.73 (0.88)	2.5 (1.24)		4.35 (0.88)
spoon2-photo-colour	4.6 (0.68)	2.7 (1.26)	2.95 (1.39)	3.82 (1.14)
spoon2-photo-grey	4.9 (0.31)	2.55 (1.1)		3.67 (1.28)
spoon3-photo-colour	4.92 (0.28)	2.14 (1.28)	3.64 (1.43)	4.3 (1.06)
spoon3-photo-grey	4.76 (0.54)	2.62 (1.24)		4.36 (1.09)
stool				
stool1-photo-colour	3.5 (1.3)	2.85 (1.04)	2.85 (1.18)	2.45 (1.36)

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
stool1-photo-grey	3.52 (1.25)	2.35 (0.88)		1.93 (1.05)
stool2-photo-colour	4.3 (0.8)	3.1 (1.02)	2.8 (1.4)	3.71 (1.19)
stool2-photo-grey	4.3 (0.86)	2.55 (1)		3.23 (1.15)
stool3-photo-colour	4.19 (0.93)	2.86 (1.15)	2.24 (1.26)	3.14 (1.21)
stool3-photo-grey	4.42 (0.65)	2.32 (0.78)		3.41 (1.22)
swan				
swan1-photo-colour	4.62 (0.59)	3.45 (1.28)	4.4 (0.99)	4.27 (1.01)
swan1-photo-grey	3.91 (1.11)	3.2 (1.15)		3.65 (1.35)
swan2-photo-colour	3.9 (1.17)	4.15 (0.88)	4.45 (1.1)	4.64 (0.58)
swan2-photo-grey	4.3 (1.03)	3.9 (0.85)		3.81 (1.17)
swan3-photo-colour	4 (1.35)	2.59 (1.22)	4 (1.2)	4.45 (0.96)
swan3-photo-grey	4.33 (0.8)	2.62 (1.16)		4.05 (0.84)
swing				
swing1-photo-colour	4.03 (1.38)	2.33 (1.06)	3.05 (1.28)	3.65 (0.93)
swing1-photo-grey	4.25 (1.12)	2.09 (0.87)		2.95 (1.28)
swing2-photo-colour	4.05 (1.25)	2.35 (0.88)	2.8 (1.28)	3.14 (1.35)
swing2-photo-grey	4.05 (1.2)	1.9 (0.64)		2.95 (1.39)
swing3-photo-colour	3.45 (1.41)	2.72 (1.14)	3.28 (1.34)	2.86 (1.28)
swing3-photo-grey	3.82 (1.37)	2.14 (0.83)		3.29 (0.96)
table				
table1-photo-colour	4.91 (0.43)	2.15 (1.27)	2.8 (1.36)	3.95 (1)
table1-photo-grey	4.9 (0.44)	1.9 (0.91)		3.47 (1.14)
table2-photo-colour	4.95 (0.22)	1.95 (1.15)	1.8 (1.32)	3.52 (1.36)
table2-photo-grey	4.8 (0.62)	1.5 (0.95)		3.32 (1.13)
table3-photo-colour	4.14 (1.24)	3.48 (1.08)	2.19 (1.21)	2.27 (0.98)
table3-photo-grey	4.29 (0.91)	3.27 (1.12)		2.23 (0.97)
thimble				
thimble1-photo-colour	3.4 (1.64)	2.87 (1.04)	3.4 (1.31)	3.95 (1.28)
thimble1-photo-grey	3.45 (1.64)	2.4 (1.1)		3.91 (1.38)
thimble2-photo-colour	3.25 (1.62)	3.86 (1.13)	2.64 (1.47)	3.7 (1.03)
thimble2-photo-grey	3.15 (1.14)	3.33 (1.46)		4.55 (0.89)
thimble3-photo-colour	2.73 (1.39)	3.09 (1.34)	3.41 (1.4)	4.08 (1.35)
thimble3-photo-grey	2.86 (1.56)	2.59 (1.3)		3.48 (1.6)
thumb				

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
thumb1-photo-colour	4.9 (0.31)	2.95 (1.28)	3.2 (1.44)	3.77 (1.34)
thumb1-photo-grey	5 (0)	2.63 (1.22)	1.82 (0.87)	3.52 (1.29)
thumb2-photo-colour	4.55 (0.69)	2.57 (1.12)	2.81 (1.54)	4.15 (0.99)
thumb2-photo-grey	4.67 (0.91)	3.05 (0.9)		3.3 (1.45)
thumb3-photo-colour	4.67 (0.73)	3.32 (1.09)	2.77 (1.34)	3.52 (1.44)
thumb3-photo-grey	4.43 (1.08)	3.14 (1.32)		4.04 (1.12)
tiger				
tiger1-photo-colour	3.45 (1.71)	4.1 (0.91)	4.45 (0.76)	4.55 (0.83)
tiger1-photo-grey	4.1 (1.18)	3.55 (1.23)		3.53 (1.07)
tiger2-photo-colour	4.15 (1.18)	4.05 (0.83)	4.65 (0.59)	4.57 (0.81)
tiger2-photo-grey	3.6 (1.35)	3.2 (1.28)		4.18 (0.73)
tiger3-photo-colour	4 (1.26)	4.05 (1.2)	4.14 (1.06)	4.59 (0.67)
tiger3-photo-grey	3.75 (1.57)	3.45 (1.22)		4.18 (0.85)
toaster				
toaster1-photo-colour	4.7 (0.66)	2.9 (1.02)	2.6 (1.35)	3.32 (1.13)
toaster1-photo-grey	4.9 (0.31)	2.87 (1.07)	2.55 (1.69)	3.48 (0.93)
toaster2-photo-colour	4.75 (0.55)	3.05 (1.28)	1.9 (1.04)	3.95 (1.23)
toaster2-photo-grey	4.4 (0.75)	3.68 (1.13)		4.1 (0.91)
toaster3-photo-colour	4.57 (0.6)	3.36 (1.05)	1.73 (1.2)	3.19 (1.33)
toaster3-photo-grey	4.41 (0.59)	3.5 (0.74)		3.12 (1.3)
tomato				
tomato1-photo-colour	4.85 (0.49)	2.5 (1.36)	4.35 (1.14)	4.67 (0.58)
tomato1-photo-grey	4.5 (0.95)	2.2 (1.15)		3 (1.27)
tomato2-photo-colour	4.76 (0.54)	3.45 (1.3)	4.18 (0.91)	4.3 (0.73)
tomato2-photo-grey	4.3 (1.08)	2.43 (0.93)		3.05 (1.32)
tomato3-photo-colour	4.35 (0.93)	2.36 (1.22)	4.41 (0.85)	4.92 (0.28)
tomato3-photo-grey	4.38 (0.97)	2.27 (1.16)		3.14 (1.39)
train				
train1-photo-colour	4.36 (0.9)	4.55 (0.6)	2.6 (1.35)	3.3 (1.69)
train1-photo-grey	4.05 (1.02)	4.1 (1.33)		2.57 (1.1)
train2-photo-colour	4.15 (0.93)	4.3 (0.47)	2.3 (1.53)	2.81 (1.25)
train2-photo-grey	3.85 (1.14)	3.5 (1.5)		3.05 (1.25)
train3-photo-colour	4.52 (0.68)	4 (1.26)	2.57 (1.36)	2.91 (0.87)
train3-photo-grey	4.42 (1.02)	3.95 (1)		3.18 (1.22)

(continued)

Photograph	Familiarity	Visual	Colour	Mental Imagery
		Complexity	Diagnosticity	
tree				
tree1-photo-colour	4.9 (0.3)	4.05 (1.28)	3.7 (1.03)	3.8 (1.24)
tree1-photo-grey	4.64 (0.95)	4.2 (0.83)		3.55 (1.23)
tree2-photo-colour	4.9 (0.31)	4.1 (1.07)	4.05 (1.19)	4.59 (0.59)
tree2-photo-grey	4.9 (0.31)	3.4 (1.1)		3.67 (1.46)
tree3-photo-colour	4.88 (0.34)	3.64 (1.18)	4.14 (0.94)	4.05 (1.21)
tree3-photo-grey	4.38 (1.07)	3.29 (1.27)		3.27 (1.12)
trumpet				
trumpet1-photo-colour	3.3 (1.59)	3.6 (1.14)	3.5 (0.95)	3.77 (1.11)
trumpet1-photo-grey	3.55 (1.54)	3.43 (1.01)	2.82 (0.98)	3.81 (1.17)
trumpet2-photo-colour	3.8 (1.2)	2.9 (1.3)	4.33 (0.97)	4.9 (0.31)
trumpet2-photo-grey	3.5 (1.54)	3.5 (1.01)		4.05 (1)
trumpet3-photo-colour	3.38 (1.53)	3.36 (0.85)	3.95 (1.25)	4.19 (0.81)
trumpet3-photo-grey	2.77 (1.31)	3.27 (0.83)		4.12 (1.12)
turtle				
turtle1-photo-colour	3.45 (1.43)	3.25 (1.12)	3.65 (1.14)	3.41 (1.22)
turtle1-photo-grey	3.65 (1.39)	3.43 (0.97)	2 (1.1)	3.24 (1.18)
turtle2-photo-colour	3.95 (1.19)	3.24 (1.41)	4.29 (1.01)	3.1 (1.37)
turtle2-photo-grey	3.35 (1.39)	3.68 (1.21)		2.55 (1.39)
turtle3-photo-colour	3.52 (1.44)	4.09 (0.92)	4.18 (1.1)	3.81 (1.54)
turtle3-photo-grey	3.36 (1.43)	3.35 (1.11)		3.5 (1.5)
vest				
vest1-photo-colour	4.05 (1.28)	1.95 (1.13)	1.86 (1.21)	1.85 (1.09)
vest1-photo-grey	3.73 (1.26)	2 (1.22)		2.7 (1.38)
vest2-photo-colour	4.05 (1.16)	2.05 (0.94)	2.05 (1.15)	2.2 (1.28)
vest2-photo-grey	4 (1.15)	2.1 (0.91)		2.76 (1.48)
vest3-photo-colour	3.68 (1.25)	3.05 (1.05)	2.18 (1.1)	2.33 (1.2)
vest3-photo-grey	3.23 (1.38)	2.8 (1.15)		1.95 (1.17)
violin				
violin1-photo-colour	4 (0.89)	3.7 (1.17)	4.2 (0.95)	4.47 (1.07)
violin1-photo-grey	3.36 (1.47)	3.1 (0.91)		4 (1.12)
violin2-photo-colour	3.6 (1.54)	3.4 (1.14)	4.05 (1.1)	4.5 (0.91)
violin2-photo-grey	3.7 (1.53)	3.25 (0.85)		4.19 (0.87)
violin3-photo-colour	3.08 (1.56)	3.41 (1.22)	3.77 (1.27)	4.05 (1)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
violin3-photo-grey	3.57 (1.21)	3.52 (1.33)		3.32 (1.17)
watch				
watch1-photo-colour	4.6 (0.5)	3.05 (1.05)	2.85 (1.5)	3.59 (1.18)
watch1-photo-grey	4.8 (0.7)	2.87 (0.97)	2.45 (1.63)	3.19 (1.08)
watch2-photo-colour	4.85 (0.37)	3.33 (0.97)	1.86 (1.2)	3.8 (1.06)
watch2-photo-grey	4.7 (0.66)	3.36 (1.14)		3.15 (1.31)
watch3-photo-colour	4.43 (0.81)	3.09 (1.15)	2.05 (1.17)	3.43 (1.12)
watch3-photo-grey	4.41 (0.85)	3.23 (1.02)		3.83 (1.2)
well				
well1-photo-colour	3.71 (1.1)	3.9 (1.25)	3.45 (1.36)	4.23 (1.14)
well1-photo-grey	2.95 (1.43)	3.65 (0.75)		4.1 (1.02)
well2-photo-colour	3.25 (1.52)	3.65 (1.23)	3 (1.69)	4 (1.2)
well2-photo-grey	3.95 (1.1)	3.45 (0.94)		4.05 (1.07)
well3-photo-colour	2.88 (1.42)	3.91 (0.75)	3 (1.07)	3.86 (1.08)
well3-photo-grey	3.14 (1.31)	3.67 (1.11)		3.59 (1.1)
whistle				
whistle1-photo-colour	4.05 (1.15)	2.75 (1.33)	2.9 (1.21)	4.27 (0.83)
whistle1-photo-grey	4.25 (1.21)	2.2 (1.13)	3 (1)	4.24 (0.94)
whistle2-photo-colour	4.3 (1.13)	1.95 (0.8)	2.9 (1.18)	4.35 (0.99)
whistle2-photo-grey	4 (1.08)	2.95 (1)		4.25 (1.02)
whistle3-photo-colour	3.43 (1.29)	2.18 (1.01)	3.27 (1.35)	4.29 (1.06)
whistle3-photo-grey	3.48 (1.34)	2.55 (0.96)		4.46 (1.14)
window				
window1-photo-colour	4.7 (0.57)	2.75 (1.25)	2.3 (1.22)	2.82 (1.26)
window1-photo-grey	4.9 (0.45)	2.87 (1.11)	1.91 (1.3)	2.81 (1.17)
window2-photo-colour	4.75 (0.79)	3.62 (0.97)	2 (1.14)	4 (1.12)
window2-photo-grey	4.55 (0.69)	3.45 (1.3)		3.4 (1.19)
window3-photo-colour	4.57 (0.6)	2.64 (0.9)	2.27 (1.2)	3.05 (1.32)
window3-photo-grey	4.82 (0.39)	2.64 (0.9)		3.92 (0.72)
zebra				
zebra1-photo-colour	3.41 (1.56)	3.8 (1.11)	4.75 (0.44)	4.6 (0.6)
zebra1-photo-grey	4.29 (1.06)	3.25 (1.12)		4.13 (1.17)
zebra2-photo-colour	3.9 (1.37)	3.8 (1.01)	4.55 (0.89)	4.52 (0.81)
zebra2-photo-grey	3.45 (1.5)	3.9 (1.33)		4.73 (0.55)

(continued)

Photograph	Familiarity	Visual Complexity	Colour	Mental Imagery
			Diagnosticity	
zebra3-photo-colour	4.1 (1.26)	3.62 (1.07)	4.76 (0.44)	4.23 (1.11)
zebra3-photo-grey	3.5 (1.59)	3.23 (1.23)		4.68 (0.65)