

# Investigating the role of working memory resources across aesthetic and non-aesthetic judgements



Ionela Bara<sup>1</sup> , Richard J Binney<sup>1</sup> and Richard Ramsey<sup>2</sup>

Quarterly Journal of Experimental Psychology  
1–19

© Experimental Psychology Society 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/17470218221101876

qjep.sagepub.com



## Abstract

Aesthetic judgements dominate much of daily life by guiding how we evaluate objects, people, and experiences in our environment. One key question that remains unanswered is the extent to which more specialised or largely general cognitive resources support aesthetic judgements. To investigate this question in the context of working memory, we examined the extent to which a working memory load produces similar or different response time interference on aesthetic compared with non-aesthetic judgements. Across three pre-registered experiments that used Bayesian multi-level modelling approaches ( $N > 100$  per experiment), we found clear evidence that a working memory load produces similar response time interference on aesthetic judgements relative to non-aesthetic (motion) judgements. We also showed that this similarity in processing across aesthetic versus non-aesthetic judgements holds across variations in the form of art (people vs. landscape; Experiments 1–3), medium type (artwork vs. photographs; Experiment 2), and load content (art images vs. letters; Experiments 1–3). These findings suggest that across a range of experimental contexts, as well as different processing streams in working memory (e.g., visual vs. verbal), aesthetic and motion judgements commonly rely on a domain-general cognitive system, rather than a system that is more specifically tied to aesthetic judgements. In doing so, these findings shine new light on the working memory resources that support aesthetic judgements, as well as on how domain-general cognitive systems operate more generally in cognition.

## Keywords

Aesthetic judgement; working memory; dual-task paradigm; working memory load

Received: 13 December 2021; revised: 31 March 2022; accepted: 3 May 2022

## General introduction

Fascination with art is a universal and timeless human phenomenon (Davies, 2012; Dissanayake, 1995; Dutton, 2009; Hoffmann et al., 2018; Lamarque, 1999; Marshack, 1996; Martindale et al., 2019; Morriss-Kay, 2010; White et al., 2012). From creating art to visiting galleries and attending live performances, people are frequently captivated by the aesthetic appeal of art. Likewise, interest in studying art across the scientific community has led to a programme of research that investigates aesthetic experiences from psychological and neuroscientific perspectives (Augustin et al., 2012; Berlyne, 1971; Cattaneo, 2019; Chatterjee, 2003; Fechner, 1876; Iigaya et al., 2020; Jacobsen, 2006; Kirsch et al., 2016; Nadal & Chatterjee, 2019; Nadal & Skov, 2015; Palmer et al., 2013; Pearce et al., 2016; Van de Cruys & Wagemans, 2011; Zeki, 1999). Yet understanding of the cognitive processes that support

aesthetic judgements remains in its infancy. Given the vital role of aesthetics in guiding how we appraise objects, people, and experiences in our environment, the current work investigates the type of cognitive processes that support aesthetic judgements.

<sup>1</sup>Wales Institute for Cognitive Neuroscience, School of Human and Behavioural Sciences, Bangor University, Bangor, UK

<sup>2</sup>School of Psychological Sciences, Macquarie University, Sydney, NSW, Australia

### Corresponding authors:

Ionela Bara, Wales Institute for Cognitive Neuroscience, School of Human and Behavioural Sciences, Bangor University, Bangor LL57 2AS, Gwynedd, UK.

Email: ionela.bara@bangor.ac.uk

Richard Ramsey, School of Psychological Sciences, Macquarie University, Sydney, NSW 2109, Australia.

Email: richard.ramsey@mq.edu.au

Despite some variability in terms of emphasis, most of the previous models have often characterised the cognitive processes that underpin aesthetic judgements using dual-processing frameworks, which distinguish automatic from more controlled processing stages (Chatterjee, 2003; Chatterjee & Vartanian, 2014; Graf & Landwehr, 2015; Leder et al., 2004; Leder & Nadal, 2014; Locher et al., 2007, 2010; Pearce et al., 2016; Pelowski & Akiba, 2011; Redies, 2015). For example, Leder and colleagues (2004) proposed that aesthetic judgements represent the end-product of a sequential cascade of five information processing stages that include automatic and controlled processes and span sensory-perceptual signals, cognitive mastering, and deliberate evaluation stages. Although it seems likely, or even necessary, that a form of executive control would be required during aesthetic judgements, the type and structure of such executive control remains largely unknown. Moreover, the extent to which aesthetic judgements rely upon general executive control mechanisms, which operate across many domains, or more specific mechanisms, which are partially tied to aesthetic contexts, remains unclear.

One way to probe the operation of executive functions is to use dual-task paradigms, whereby a demanding secondary task is performed alongside a main task of interest (Lavie et al., 2004; Satpute & Lieberman, 2006). For example, participants may hold in memory one letter (low load) or six letters (high load) while quickly and accurately performing the primary task. According to load theory (Lavie et al., 2005, 2010), executive functions, such as working memory, help to maintain response priorities throughout a task. Consequently, when working memory resources are loaded with a demanding secondary task, the control capacity that maintains task priority is reduced, leading to increased distractor interference that perturbs the main task response. In cases when higher load interferes with the main task, it has been suggested that mental operations required during the main task are relatively resource-intensive and reliant on controlled and effortful processes. In contrast, in cases when higher load does not interfere with the main task, it suggests that mental operations required during the main task are resource-light, relatively efficient, and less reliant on controlled or effortful processes. As such, dual-task paradigms are a useful way to characterise the type of working memory resources that are relied upon in a given context.

A small number of previous studies have investigated the type of cognitive systems that are involved in aesthetic judgements using dual-task paradigms. Briellmann and Pelli (2017) found that adding a secondary two-back task decreased aesthetic judgements for beautiful stimuli, but not for non-beautiful stimuli. Likewise, Che and colleagues (2021) showed that a demanding secondary task delayed judgements of beauty, but not liking. These findings suggest that, during aesthetic judgements, different

stimulus and task features place more demands on effortful operations of executive control. Conversely, Mullennix and colleagues (2013) showed that aesthetic ratings were not affected by a secondary load task. This latter finding suggests that, at least in some instances, aesthetic judgements remain unaffected by higher load and can be processed in a relatively automatic manner.

These prior studies of aesthetics using secondary tasks have all used art-based stimuli and aesthetically oriented tasks, such as judgements of beauty and liking. Such experimental designs are useful for probing information processing structures within aesthetic contexts. However, these designs are unable to address the extent to which common or distinct forms of executive control are deployed across aesthetic compared with non-aesthetic judgements. As such, new and unexplored questions remain concerning domain specificity in aesthetic judgements compared with non-aesthetic judgements, and contrasting theoretical possibilities exist. A domain-specific account would suggest that aesthetic judgements draw upon distinct sets of cognitive control processes (Goldman, 2001; Guyer, 2005). One prediction that follows from this account is that aesthetic judgements may rely on partially distinct executive resources compared with non-aesthetic judgements. In contrast, a domain-general account would suggest that the same set of executive resources will be deployed in a similar manner across aesthetic and non-aesthetic contexts. For example, a semantic cognition account of aesthetics predicts that similar cognitive and brain systems that are engaged in extracting meaning from the environment in general (i.e., non-aesthetic contexts), such as modality-specific conceptual representations and controlled executive processes, would be similarly involved in aesthetic judgements (Bara, Binney, et al., 2021).

Although “aesthetic” and “non-aesthetic” are familiar terms in neuroaesthetics research, establishing unambiguous boundaries between these terms is not straightforward. Avoiding the need for categorical divisions, feature mapping, or dimensional approaches, which have previously been used in social cognition and psychopathology (Brown & Barlow, 2009; Cross & Ramsey, 2021; Oosterhof & Todorov, 2008), could provide a fruitful alternative perspective in defining “aesthetic” and “non-aesthetic.” According to dimensional perspectives, different stimulus or task features could be more or less aesthetically oriented. For example, the assessment of visual clarity (Whittlesea et al., 1990), implied motion (Bara, Darda, et al., 2021), or symmetry (Jacobsen & Höfel, 2003; Jacobsen et al., 2006) could be regarded as less aesthetically oriented than assessing liking, preference, or beauty. Furthermore, this definition means that stimuli, tasks, and contexts that possess fewer aesthetic features are not necessarily devoid of any aesthetic features. Accordingly, the current work uses “aesthetic” and “non-aesthetic” in a relative sense, where the former has more aesthetic features than the latter.

In the current work, we contrast an aesthetic judgement with a particular type of non-aesthetic judgement, which involves assessing implied motion. We chose to focus on this distinction because different lines of prior research have suggested that aesthetic judgements rely on more elaborate cognitive processes than motion judgements. For example, aesthetic models characterise aesthetic judgements as requiring complex, continuous, and dynamic integration of perception, memory, attention, action, and affective resources (e.g., Chatterjee & Vartanian, 2014; Leder & Nadal, 2014; Pearce et al., 2016). In contrast, motion sensitivity has been associated with specialised and fast processing in select patches of the visual cortex (Beauchamp et al., 2002; Mather et al., 1992; Maunsell & Van Essen, 1983). Given these differences in cognitive processing between aesthetic and motion judgements, we thought it was reasonable to hypothesise that aesthetic judgements may rely on a distinctive type of working memory resource that motion judgements do not require to the same degree.

Therefore, the overarching aim of the current study is to investigate the extent to which domain-general or domain-specific working memory resources are deployed during aesthetic compared with non-aesthetic judgements. More specifically, the novel question we address here is the extent to which working memory load produces similar or different response time interference on aesthetic judgements compared with non-aesthetic judgements. By using a Bayesian analytical framework (rather than null-hypothesis significance testing), we can provide supporting evidence for the domain-general or domain-specific accounts that we have outlined. In other words, we can provide support for a similarity in interference, as well as a difference in interference, between aesthetic and non-aesthetic contexts. Across three pre-registered experiments, we test these hypotheses by varying the type of judgement, the type of stimuli, and the type of load content between aesthetic and non-aesthetic categories. By doing so, we are able to test the extent to which the pattern of results generalises across different stimulus features and task contexts.

## Experiment 1

### Introduction

In Experiment 1, we investigated to what extent high working memory load produces greater response time interference on aesthetic judgements relative to non-aesthetic judgements. To do so, we compared aesthetic with implied motion judgements towards the same art stimuli. Greater interference in aesthetic than non-aesthetic judgements would support the view that somewhat specialised working memory resources are deployed during aesthetic judgements. In contrast, equivalent interference between aesthetic and non-aesthetic judgements would support the








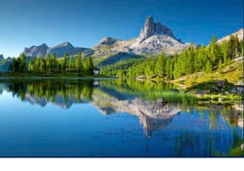
view that largely general working memory resources support aesthetic judgements.

### Method

**Pre-registration and Open Science statement.** Across all three experiments, the research questions, hypotheses, planned analyses, sample sizes, and exclusion criteria were pre-registered. For Experiment 1, the pre-registration can be found at <https://aspredicted.org/4td5q.pdf>. In addition, following open science initiatives (Munafò et al., 2017), all raw data, stimuli, and analysis code for each experiment are available online on the open science framework (<https://osf.io/9q5jx/>).

We note a few minor deviations from the pre-registered analysis. We pre-registered that prior to building regression models, we would remove trials from the data with response times less than 10ms, as they are likely to reflect a response error. Due to the type of modelling we performed, which involved shifted lognormal models, low response times can make model fitting and model comparison more difficult. As such, the reported models in all experiments have a 100ms cut-off point, rather than a 10ms one. We did run the models both ways and there were no meaningful differences between the models. In fact, there were only a few data points that were between 10 and 100ms. For example, in Experiment 1, there were only 16 data points in this range out of approximately 16,000 datapoints in total. However, given that the models were easier to work with when using a 100ms cut-off, we chose to use this throughout all of the the experiments. Furthermore, we pre-registered a shifted lognormal model which naturally requires a non-decision time parameter (ndt); however, in the full model, we additionally allowed ndt parameter to vary by participant. Another small deviation refers to the number of participants in Experiment 1. We pre-registered 50 usable data files, but in the context of the Covid pandemic and the easier access to online testing platforms, we decided to test 100 participants.

**Participants.** One hundred and two participants took part in this study for course credit (21 males,  $M_{age}=21.09$ ,  $SD_{age}=5.30$ , age range=18–44). All participants provided informed consent before completing the experiment. The experiment was granted ethical approval by the Research Ethics and Governance Committee of the School of Human and Behavioural Sciences at Bangor University. According to our pre-registered exclusion criteria, we excluded participants who had an average working memory accuracy that was less than 55% (four participants were removed). The reason for doing so is the difficulty in assessing whether a close to 50% memory accuracy is indicative of the task's difficulty or lack of engagement in completing the task. Trials with response times less than 100ms on the main judgement tasks were also removed. The final sample included 98 participants and a total of 15,154 trials.

	People		Landscape	
	Dynamic	Static	Dynamic	Static
<b>Art Stimuli</b> (Exp. 1-3)				
<b>Photographic Stimuli</b> (Exp. 2)				

**Figure 1.** A representation of the four different stimulus categories used in Experiments 1, 2, and 3: people dynamic ( $N=20$ ), people static ( $N=20$ ), landscape dynamic ( $N=20$ ), and landscape static ( $N=20$ ) across art images and non-art images.

### Stimuli, tasks, and procedure

**Art stimuli.** The art stimuli dataset consisted of 80 images of representational paintings depicting either human bodies (40 images) or landscapes (40 images). The stimuli were validated previously across a range of dimensions: familiarity, aesthetic appreciation, implied dynamism, and evocativeness (Bara, Darda, et al., 2021). The images were characterised by a realistic representational style in the 19th–20th century European and American pictorial tradition. All art stimuli (landscape or people) were divided further into static and dynamic. This partition was based on previously recorded subjective judgements that assessed the degree to which a stimulus would contain a clear sense of implied dynamism (Bara, Darda, et al., 2021). Overall, the stimuli were split into four different groups within a 2 (painting type: landscape or people) by 2 (dynamism: static vs. dynamic) design. Each image was cropped to be  $785 \times 774$  pixels in size. For a complete description of the stimuli used in Experiment 1, including the list of artworks, artists, year of production, and museum collection, please see Supplementary materials (Table S1). Copyright permitting all the art stimuli that we used are also available on our open science framework page (<https://osf.io/9q5jx/>). Example stimuli across all experiments can be seen in Figure 1.

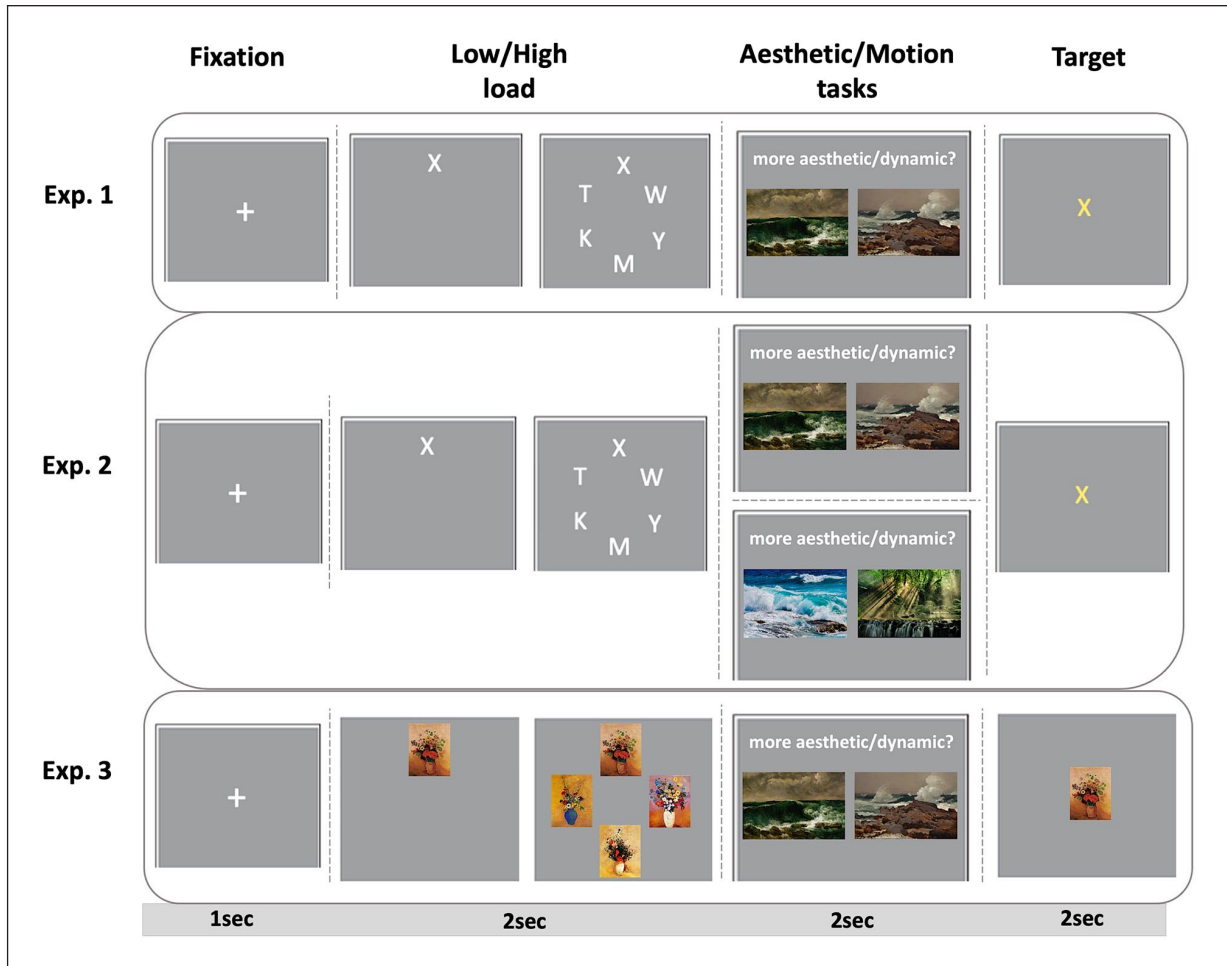
**Tasks and procedure.** The main experimental task involved completing a working memory task and a 2-alternative forced choice (2-AFC; Figure 2). The 2-AFC task consisted of the simultaneous presentation of two paintings next to each other (in the middle of the screen) and participants had to make an aesthetic judgement or an implied motion judgement. In the aesthetic judgement task, participants had to choose which of the two paintings was more aesthetically pleasing, whereas, in the motion judgement task, participants had to indicate which of the two paintings was more dynamic.

For both aesthetic judgement and motion judgement tasks, the stimuli were randomly paired from within the same category, across four categories: landscape dynamic, landscape static, people dynamic, and people static. Therefore, there were four possible pairing trial types for each judgement type. For example, an aesthetic judgement trial could consist of the pairing of two landscape dynamic paintings, two landscape static paintings, two people dynamic paintings, or two people static paintings. Individual paintings could not be paired together on the same trial, although paintings from the same category could be presented more than once, but in a different position of the screen (left vs. right). The experimental tasks were produced in PsychoPy (v2020.2.3, Peirce et al., 2019) and run online using Pavlovia and recruitment was via the Bangor University SONA system.

The concurrent working memory task involved the presentation of either one letter (low-load condition) or six letters (high-load condition) in the centre of the screen. The letters for each trial were presented in a circular arrangement and were randomly selected from a set of 10 capital letters (FHKLMTVWYX). No letters were presented twice within the same high-load trials. For the low-load trials, the space of the other missing five letters was replaced by five dots in a circular array.

Before each experimental task (aesthetic judgement and motion judgement), participants completed a practice block of 32 trials containing both the working memory task and experimental task. To avoid a familiarity effect, the art images presented in the practice block differed from the art images in the main experimental tasks. Moreover, the practice block art stimuli had the same characteristics as the art images in the main experimental block—realistic representational 19th–20th century pictorial style, divided into two main categories landscape (dynamic and static) and people (dynamic and static).





**Figure 2.** Experimental design and stimuli across Experiments 1, 2, and 3. Each experiment had the same structure. First, there was a fixation cross and then there were letters or artworks to be held in memory (one item as low load, or multiple items as high load). While memorising the working memory load content for a later probe, participants responded to 2-AFC aesthetic judgement task or motion judgement task. Following the 2-AFC aesthetic judgement or motion judgement task, a target appeared, and participants had to confirm whether the item was present or absent at the beginning of the trial.

As shown in Figure 2, each trial started with a fixation cross for 1 s followed by a memory set of either one (low load) or six (high load) letters in the middle of the screen for 2 s. Participants were instructed to memorise the presented letters to the best of their ability. Next, two paintings (0.5, 0.5—PsychoPy unit size display, where 1 unit is equal to the height of the screen on which the experiment was running) were presented concurrently at  $-0.3$  PsychoPy units to the left and  $0.3$  PsychoPy units to the right of the centre of the screen. The two paintings were presented for 2 s alongside the question “more aesthetic?” or “more dynamic?” at the top of the screen. The paintings remained on the screen for 2 s while participants were asked to make a speeded aesthetic judgement or a motion judgement by pressing down on either “j” for choosing the left painting or “k” for choosing the right painting. After the participants responded to the 2-AFC task, the memory probe letter was then presented. Participants had to press either the “e” or “d” key to indicate whether the letter was

present or absent at the beginning of each trial. The memory probe letter was displayed for 2 s while the participants had to make a response.

Overall, the experiment used a repeated measure design containing 160 trials, with 8 trial types formed by intersecting load type (high or low) with image type (landscape or people) and with judgement type (aesthetic judgement or motion judgement). The aesthetic judgement and motion judgement tasks were counterbalanced across all participants so that respondents would start with either the aesthetic judgement task or with the motion judgement task.

For exploratory purposes, at the beginning of the experiment, participants completed the Vienna Art Interest and Art Knowledge Questionnaire (VAIAK) (Specker et al., 2020) to assess participants’ art interest and art knowledge. The results are reported in Supplementary material (Table S3–S4). We also explored the relationship between the aesthetic judgement choice and image dynamism type

**Table 1.** Weakly informative priors used across all three experiments.

Prior	Class	dpar
normal (6.68, 0.5)	Intercept	
normal (5.70, 0.5)	Intercept	ndt
normal (0, 0.05)	b	
normal (0, 0.1)	sd	
normal (0, 0.1)	sd	ndt
normal (0, 0.1)	sigma	
lkj (2)	cor	

dpar: distributional parameter; ndt: non-decision time; b: population-level or fixed effects; sd: standard deviation; cor: correlation.

(dynamic vs. static). The results are reported in Supplementary material (Figure S10.A).

**Data analyses.** We pre-registered a Bayesian estimation approach to multi-level regression modelling (McElreath, 2020). The main rationale was to estimate parameters of interest in multi-level models and perform model comparison between simpler and more complex models. Therefore, when interpreting the findings, we used two approaches. First, we reported and discussed the posterior distribution of our key parameters of interest within the most complex model. Second, we performed model comparison via efficient approximate leave-one-out cross validation (LOO; Vehtari et al., 2017). LOO is a way of estimating how accurately the model can predict out-of-sample data. Therefore, we took all the models and estimated how accurate they were at predicting the out-of-sample data. In this way, we could estimate how much increasing model complexity increases model accuracy.

More specifically, we followed a recent translation of McElreath's (2020) general principles into a different set of tools (Kurz, 2020), which use the Bayesian modelling package "brms" to build multi-level models (Bürkner, 2017, 2018). Moreover, our data wrangling approach follows the "tidyverse" principles (Wickham & Grolemund, 2016) and we generate plots using the associated data plotting package "ggplot2," as well as the "tidybayes" package (Morriss-Kay, 2020). All of these analytical approaches were performed in the R programming language (R Core Team, 2020).

Given that the primary dependent variable is response time, we modelled the data using a shifted lognormal regression model, which has previously been shown to be a particularly suitable way to model response times (Haines et al., 2020). Following the "keep it maximal" approach to multi-level modelling (Barr et al., 2013), we included the maximal number of varying effects that the design permitted. As such, varying intercepts and effects of interest were estimated for participants and stimulus items when possible.

We computed nine models, which built incrementally in complexity. We first computed two intercepts-only models, just so that we could compare subsequent models that included predictors of interest with models without any predictors. Model b0 included varying intercepts for participants and stimulus items, whereas model b0.1g additionally included a varying non-decision time (ndt) parameter per participant. We then added predictors for task (b1), stimulus type (b2), and load (b3). Two-way interactions between task  $\times$  type (b4.1), task  $\times$  load (b4.2), and type  $\times$  load (b4.3) were then added in further models. Model b5 was the full model, which additionally included the three-way interaction between task, type, and load.

Factors were coded according to a deviation coding style, where factors sum to zero and the intercept can then be interpreted as the grand mean and the main effects can be interpreted similarly to a conventional analysis of variance (<http://talklab.psy.gla.ac.uk/tvw/catpred/>). As such, task, type, and load were coded as  $-0.5$  (motion/landscape/low) and  $0.5$  (aesthetic/people/high).

We set priors using a weakly informative approach (Gelman, 2006). The priors used throughout all three experiments are provided in Table 1. Weakly informative priors differ from uniform priors by placing a constrained distribution on expected results rather than leaving all results to be equally likely (i.e., uniform). They also differ from specific informative priors, which are far more precisely specified, because we currently do not have sufficient knowledge to place more specific constraints on what we expect to find. Given the relatively small effects in the field of psychology in general, as well as in reaction time studies, we centred normally distributed priors for key effects of interest on zero (i.e., no effect; see class "b" in Table 1). That means that prior to running the study, we expected effects closer to zero to be more likely than effects further away from zero. Also, by using weakly informative priors, we allow for the possibility of large effects, should they exist in the data (Gelman, 2006; Gelman et al., 2013; Gelman & Hill, 2007; Lemoine, 2019). Moreover, a further advantage of weakly informative priors is that we would not expect the choice of prior, as long as it remained only weakly informative, to matter too much because the data would dominate the structure of the posterior distribution. The formula for the full model (model 5) is specified here

$$\begin{aligned}
 \text{afc\_rtms} &\sim 1 + \text{task} \times \text{type} \times \text{load} \\
 &+ (1 + \text{task} \times \text{type} \times \text{load} \mid \text{pID}) \\
 &+ (1 + \text{task} \mid \text{item\_left}) \\
 &+ (1 + \text{task} \mid \text{item\_right}) \\
 \text{ndt} &\sim (1 \mid \text{pID})
 \end{aligned}$$

afc\_rtms=alternative forced choice response time in milliseconds; task=judgement type (motion vs. aesthetic); type=image category (landscape vs. people); load=low vs. high; pID=participant unique identifier; item\_left=image presented on the left side during alternative forced choice trials; item\_right=image presented on the right side during alternative forced choice trials; ntd=non-decision time.

Although we pre-registered an approach that built models towards the “maximal” model (Barr et al., 2013), two specific parameters were of particular interest in reference to evaluating our key hypothesis. First, we expected an overall effect of load on response time interference in the 2-AFC task, such that there would be greater interference for high than low load. This would suggest that mental operations required during the main task are relatively resource-intensive rather than resource-light, and reliant on controlled and effortful processes. Second, the task  $\times$  load interaction term was key to evaluating our main hypothesis. Evidence in favour of specialised working memory resources for aesthetic judgements would be provided by a largely positive interaction term, such that the effect of load (high > low) would be greater for aesthetic than motion judgements. In contrast, evidence in favour of largely general working memory resources for aesthetic judgements would be provided by an interaction term that largely overlaps with zero, such that the effect of load (high > low) is largely similar for aesthetic and motion judgements.

Across all three experiments, convergence across chains was carefully monitored and did not raise any concerns. The chains can be visualised in Supplementary Figure S11. For more details on the number of iterations and chains, please see our analysis code on the open science framework (<https://osf.io/9q5jx/>).

## Results

**Working memory.** Results indicated slower response time for high-load conditions compared with low-load conditions (Mean difference = 150 ms, 95% confidence interval [CI] = [130, 180]). Also, we found lower memory accuracy for high-load conditions compared with low-load conditions (Mean difference = 17.43% accuracy, 95% CI = [15.48, 19.39]). For more details, please see Supplementary Figure S3.

**2-AFC task.** Response time results for the 2-AFC task are visualised in Figure 3. Visual inspection shows longer response time on high-load conditions rather than low-load conditions, and longer response times when judging art images that contained people rather than landscapes.

Parameter estimates for the most complex model (Model 5) are shown in Figure 4 and Table 2. The posterior distribution for the main predictors indicated a largely positive response for the effect of image type (people vs.

landscape) and for the effect of load (high vs. low). These results show that response times were slower for people than landscapes and high- versus low-load conditions. As can be seen in Supplementary Figure S1, the model estimates for these effects in response times are approximately 50 ms for the effect of type and 40 ms for the effect of load. The distribution of parameter estimates for all interactions effects peaked around zero with values either side of zero emerging as the best estimate of such effects. Therefore, these interaction results provide support for similar deployment of working memory resources for both aesthetic and non-aesthetic judgements. In other words, the effect of high versus low load on response times was similar across manipulations of task type (aesthetic vs. non-aesthetic) and image type (people vs. landscape).

Model comparison analyses are visualised in Figure 5. All models with predictors performed better than the intercepts only model (Model b0), as well as the intercepts and varying effects model (Model b0.1g). Error bars for performance of the remaining models all overlapped, suggesting that they performed in a largely similar manner, in terms of out-of-sample predictive accuracy.

## Discussion

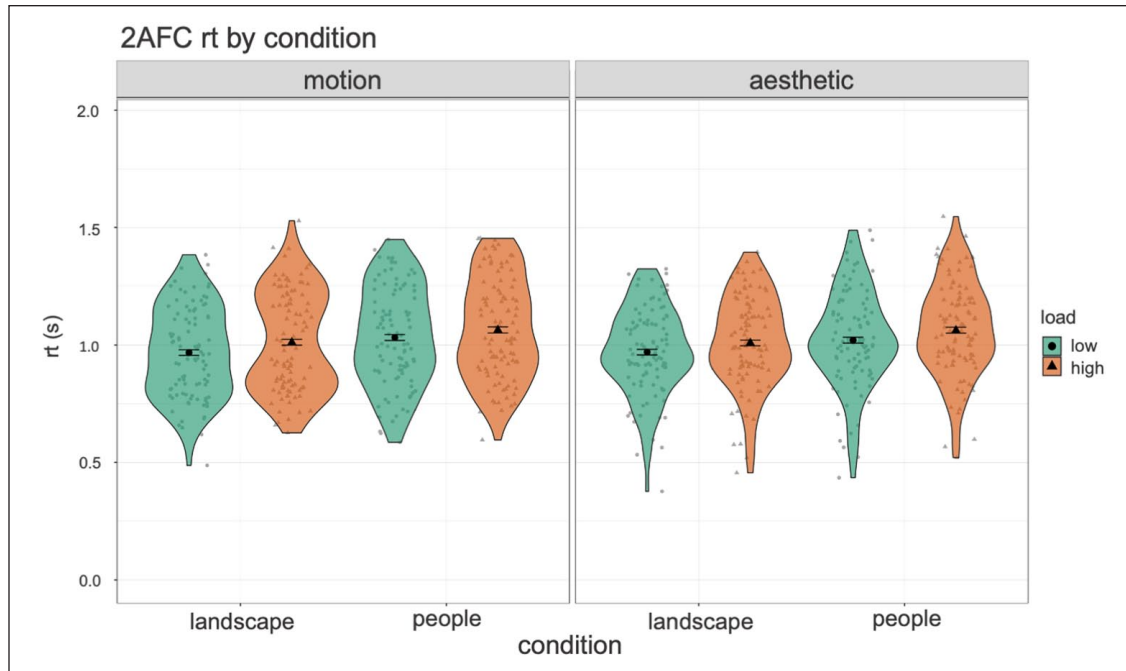
Experiment 1 demonstrated that a cognitively demanding secondary task led to indistinguishable levels of response time interference during aesthetic and implied motion judgements. In terms of our main hypothesis, therefore, we provide initial evidence to suggest that, at least in some circumstances, aesthetic and motion judgements may rely to a similar degree on operations of the working memory resources. In addition, longer response time for portraiture than landscape art suggests that the time course for aesthetic judgements is sensitive to artworks’ content. In this vein, previous work has indicated that art style and art content impact differentially the temporal course of aesthetic processing (Augustin et al., 2008; Brieber et al., 2020; Leder & Nadal, 2014).

However, before drawing firmer conclusions regarding the nature of working memory resources during aesthetic judgements, we first consider one limitation of these findings. The aesthetic and non-aesthetic judgements were restricted to art stimuli only. Given that previous work has shown that the distinction between art and non-art stimuli can become more salient when paired together (Vessel et al., 2018), it may be possible to reveal evidence for the reliance on a more distinct working memory resources by contrasting art stimuli to naturalistic photographs.

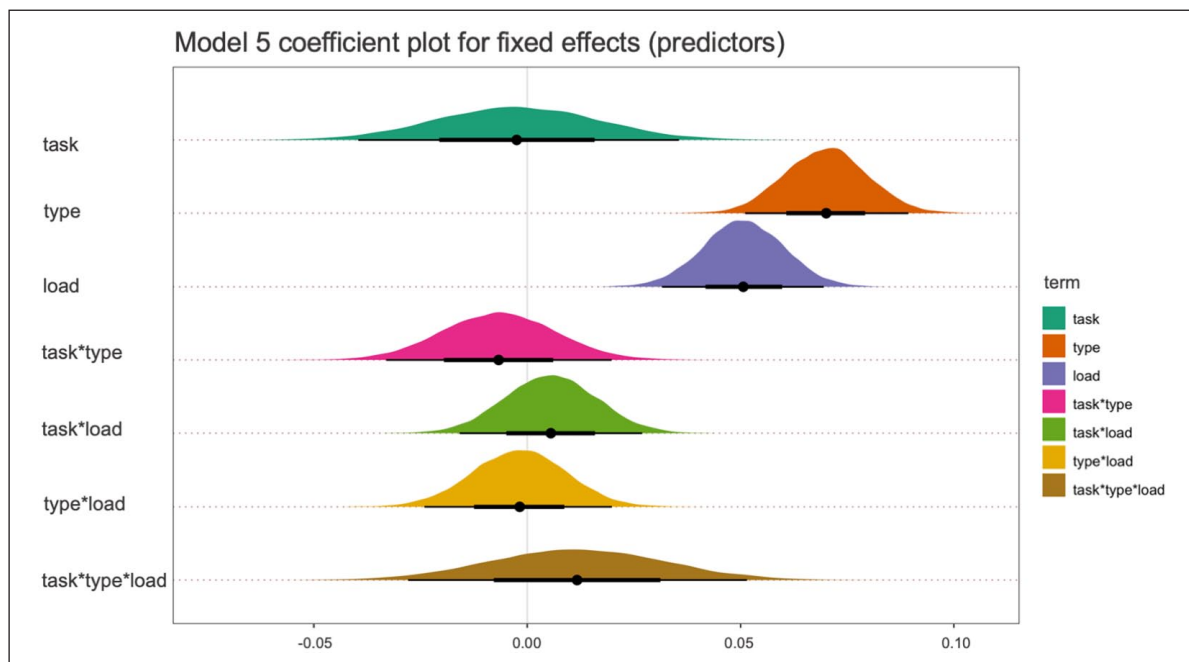
## Experiment 2

### Introduction

Experiment 2 investigated the extent to which higher working memory load produces greater response time



**Figure 3.** Results for Experiment 1—violin plots on summary data showing 2-AFC response time. Response time is reported in seconds (s). The left panel shows response times for motion judgement task on low and high load conditions for both landscape and people. The right panel shows response times for aesthetic judgement task on low and high load conditions for both landscape and people. Error bars represent 95% confidence intervals. The black markers (circles or squares) and interval estimates represent the group mean average, whereas the grey markers (circles or squares) represent the individual participants.



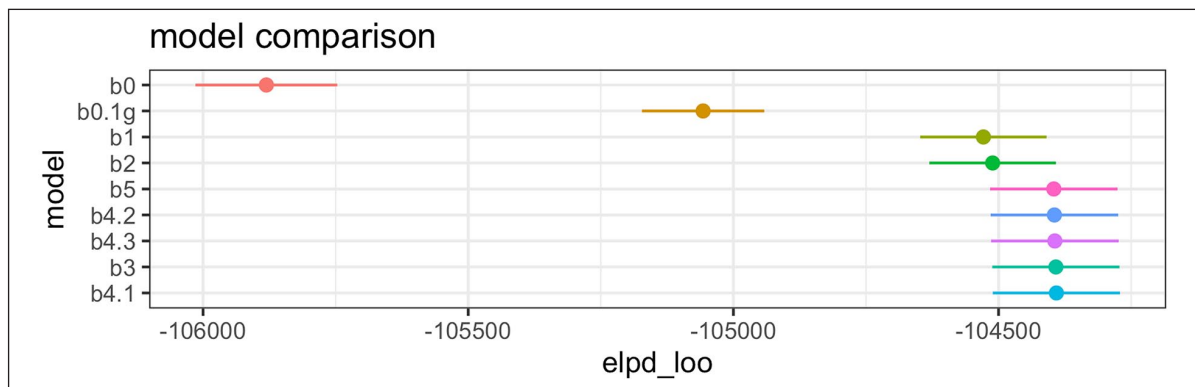
**Figure 4.** Parameter estimates for each predictor within Model 5. The main predictors that show a clear positive effect are the second and the third predictors, respectively, image type and load. The x-axis is expressed on the  $\log(\text{RT})$  scale. The direct interpretation of these parameters in terms of response times is complex as the shifted lognormal model is made of three components. To see estimates of these effects in original units (ms), please see Supplementary Figure S1. In addition, the varying effects by stimulus and by participant can be visualised in Supplementary Figure S2. task: judgement type (motion vs. aesthetic); type: image category (landscape vs. people); load: high vs. low; point estimate = median; error bars represent 66% quantile intervals (thick black lines) and 95% quantile intervals (thin black lines).



**Table 2.** Experiments 1, 2, and 3—Model b5 fixed effects.

Term	Experiment 1			Experiment 2			Experiment 3		
	value	lower	upper	value	lower	upper	value	lower	upper
Intercept	6.57	6.52	6.61	6.45	6.40	6.50	6.50	6.44	6.56
ndt_Intercept	5.16	4.98	5.34	4.98	4.78	5.18	5.06	4.85	5.27
task	-0.00	-0.04	0.04	0.01	-0.03	0.05	-0.04	-0.08	0.00
type	0.07	0.05	0.09				0.07	0.06	0.09
medium				-0.00	-0.02	0.02			
load	0.05	0.03	0.07	0.04	0.02	0.06	0.03	0.02	0.05
task*type	-0.01	-0.03	0.02				-0.00	-0.03	0.02
task*medium				0.01	-0.01	0.02			
task*load	0.01	-0.02	0.03	-0.00	-0.03	0.02	0.01	-0.02	0.04
type*load	-0.00	-0.02	0.02				0.00	-0.02	0.02
medium*load				0.00	-0.01	0.02			
task*type*load	0.01	-0.03	0.05				-0.01	-0.05	0.03
task*medium*load				-0.02	-0.05	0.02			

Note: ndt\_Intercept = non-decision time Intercept; task = motion judgment vs aesthetic judgment; type = landscape vs people; medium (Experiment 2) = photo vs painting; load = low vs high; task\*type = interaction between task (motion vs aesthetic) and type (landscape vs people); task\*medium (Experiment 2) = interaction between task (motion vs aesthetic) and medium (photo vs painting); task\*load = interaction between task (motion vs aesthetic) and load (low vs high); type\*load = interaction between type (landscape vs people) and load (low vs high); medium\*load = interaction between medium (photo vs painting) and load (low vs high); task\*type\*load = interaction between task (motion vs aesthetic), type (landscape vs people) and load (low vs high); task\*medium\*load = interaction between task (motion vs aesthetic), medium (photo vs painting) and load (low vs high). Experiment 1 and 3 had identical terms. Only Experiment 2 included the term 'medium' to describe photo vs paintings. Point estimate = median; Error bars = 95% quantile intervals.

**Figure 5.** Model comparison (1–9 models).

model b0—included varying intercepts for participants and stimuli; model b0.1—comprised a varying non-decision time parameter per participant; model b1—included predictors for task (motion vs. aesthetic); model b2—included predictors for stimulus type (photo vs. painting); model b3—included predictors for load (low vs. high); model b4.1—included interaction between task and type; model b4.2—included interaction between task and load; model b4.3—included interaction between type and load; model 5—full model which additionally included the interaction between task, type, and load; elpd\_loo: estimate of the expected log pointwise predictive density; loo: leave-one-out estimated cross validation; error bars: standard error of the mean.

interference in aesthetic judgements compared with implied motion judgements, especially while viewing artworks rather than naturalistic photographs. We reasoned that by contrasting art to non-art stimuli, we may increase the salience of the art versus non-art distinction (Vessel et al., 2018), which could make interference effects more pronounced for aesthetic than motion judgements. In addition, neuroimaging meta-analyses have demonstrated that the aesthetic response to artworks, but not naturalistic photographs, engages additional brain areas such as the amygdala

(Boccia et al., 2016) and anterior medial prefrontal cortex (Chuan-Peng et al., 2020), which suggests that more elaborate processing takes place when viewing artworks than photographs.

### Method

**Pre-registration.** We used the same design and analysis pipeline as in Experiment 1, all of which we pre-registered in advance of the experiment commencing. The

pre-registration document for Experiment 2 can be found at <https://aspredicted.org/p7gs4.pdf>.

**Participants.** One hundred participants completed this experiment for course credit (16 males,  $M_{age}=20.09$ ,  $SD_{age}=4.94$ , age range=18–44). The experiment was granted ethical approval by the Research Ethics and Governance Committee of the School of Human and Behavioural Sciences at Bangor University. All participants provided informed consent before completing the experiment. According to our pre-registered exclusion criteria, participants were excluded if their average working memory accuracy was less than 55% and trials with less than a 100 ms response time on the judgement task were also excluded. The final sample included 96 participants and a total of 29,946 trials.

### *Stimuli, task, and procedure*

**Selection and validation of non-art stimuli.** To ensure that naturalistic photographs match the standards of familiarity, aesthetic appreciation, implied motion, and evocativeness previously established for art stimuli, we conducted a separate behavioural stimuli validation experiment ( $N=43$ ; 12 males,  $M_{age}=20.63$ ,  $SD_{age}=3.44$ ). For full results of the stimuli validation experiment, please see supplementary materials (Tables S5 and S6—<https://osf.io/9q5jx/>). In brief, we wanted to investigate the extent to which implied dynamism, evocativeness, and familiarity predict aesthetic appreciation. A hierarchical regression model revealed that implied dynamism increased the predictive capacity of the model by 33%, evocativeness by 28%, and familiarity by 14%, suggesting that all predictors of interest had a unique contribution to aesthetic appreciation (Table S6—supplementary materials). Moreover, such results mirror our prior work, which assessed similar types of judgements of the art stimuli used in the current experiments (Bara, Darda, et al., 2021).

The naturalistic photos were obtained from <https://www.pexels.com/>, a free database containing a diverse range of photos and videos. The photographic stimuli dataset consisted of 80 images depicting either human bodies (40 images) or landscapes (40 images). Each group (landscape or people) was divided further into static and dynamic. Overall, photographic stimuli were divided into four different groups within a 2 (photo type: landscape or people) by 2 (dynamism: static vs. dynamic). In total, therefore, we used 160 stimuli: 80 art images from Experiment 1 and 80 naturalistic photographs. As in Experiment 1, all the stimuli were cropped to be  $785 \times 774$  pixels in size and were presented in colour and with no additional filters to original images. All the naturalistic stimuli that we used in Experiment 2 are freely available on our open science framework page (<https://osf.io/9q5jx/>). Example stimuli used in Experiment 2 are visualised in Figure 1.

The tasks used in Experiment 2 were identical to Experiment 1 with a few exceptions (Figure 2). The 2-AFC aesthetic judgement task and motion judgement task consisted of the simultaneous presentation of either two photos or two paintings next to each other in the middle of the screen. The stimuli were randomly paired from within the same category, across eight categories: photos landscape dynamic, photos landscape static, photos people dynamic, photos people static, paintings landscape dynamic, paintings landscape static, paintings people dynamic, and paintings people static. Therefore, there were eight possible pairing trial types for each judgement type. For example, an aesthetic judgement could consist of the pairing between two photos or two paintings from the “landscape dynamic” category. The same was true for the other seven categories. Individual photographic images or paintings could not be paired together on the same trial, although paintings from the same category could be presented more than once, but in a different position of the screen (left vs. right). Overall, the total number of trials in Experiment 2 increased to 320 trials per participant compared with 160 trials in Experiment 1 due to an extra experimental condition (e.g., naturalistic photo condition).

As in Experiment 1, before each experimental task, participants completed a practice block of 32 trials containing both the working memory task and experimental tasks. To avoid a familiarity effect, the images used in the practice block differed from the images used in the main experimental tasks.

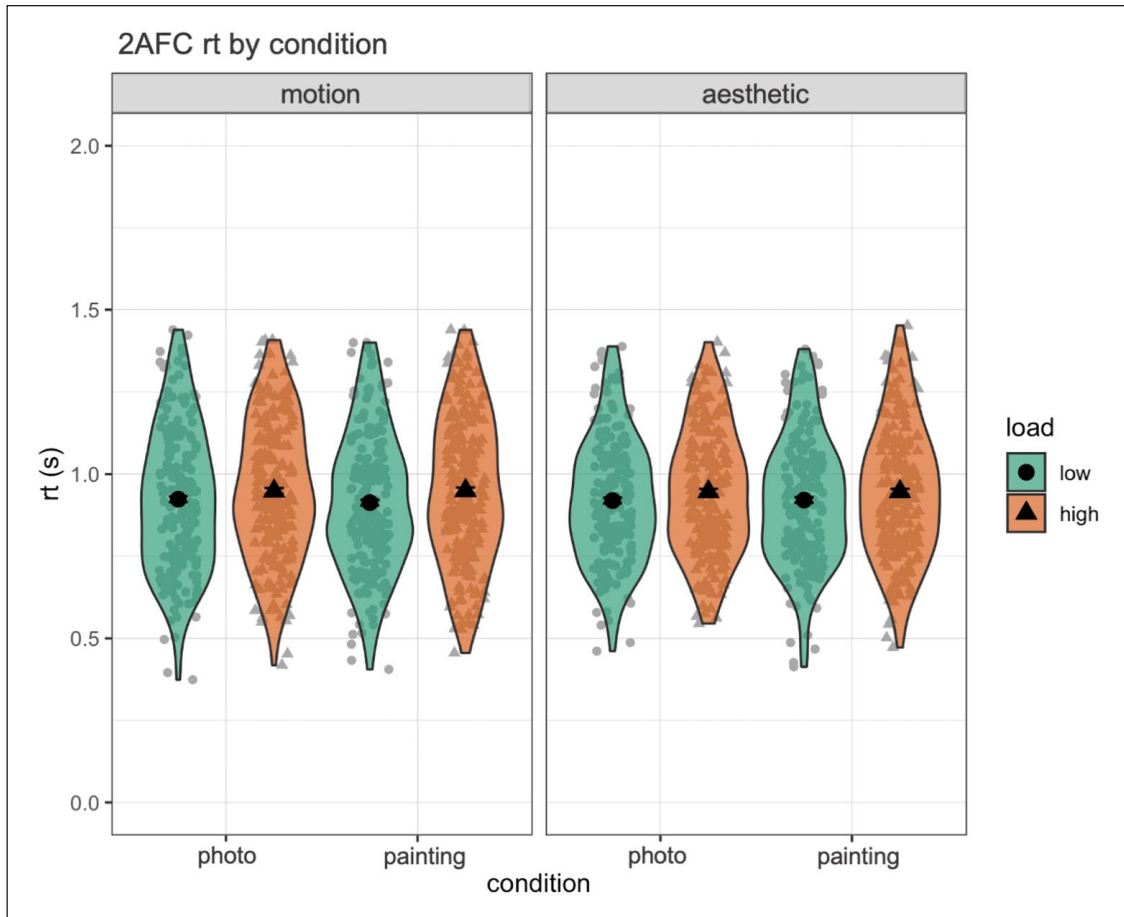
We also explored the relationship between the aesthetic judgement choice and image dynamism type (dynamic vs. static). The results are reported in Supplementary material (Figure S10.B).

**Data analyses.** We used the identical approach to data analyses as performed in Experiment 1 with one exception. Instead of modelling the type of stimulus (landscape vs. people), we modelled the type of medium (photograph vs. artwork). As such, the modelling process had the same overall structure as Experiment 1, but one factor was different.

## **Results**

**Working memory.** Results showed slower response time for high-load conditions compared with low-load conditions (Mean difference=140 ms, 95% CI=[120, 150]). Also, we found decreased memory accuracy for high-load conditions compared with low-load conditions (Mean difference=17.25% accuracy, 95% CI=[15.90, 18.59]). For more details, please see Supplementary Figure S6.

**2-AFC task.** The 2-AFC response time results are shown in Figure 6. Like Experiment 1, on average, participants took longer to respond to high-load conditions rather than low-load conditions.



**Figure 6.** Results for Experiment 2—violin plots on summary data showing 2-AFC response time. Response time is reported in seconds (s). The left panel shows response times for motion judgement task on low and high load conditions for both photos and paintings. The right panel shows response times for aesthetic judgement task on low and high load conditions for both photos and paintings. Error bars represent 95% confidence intervals. The black markers (circles or squares) and interval estimates represent the group mean average, whereas the grey markers (circles or squares) represent the individual participants.

Parameter estimates for the most complex model (Model 5) are shown in Figure 7 and Table 2. The posterior distribution for the main predictors indicated a largely positive response for the effect of load (high vs. low). This result shows that response times were slower for the high-versus low-load condition. As can be seen in Supplementary Figure S4, the model estimates in response times are between 20 and 40 ms for the effect of load. The distributions for all remaining parameters including all interaction terms showed substantial overlap with either side of zero. These interaction effect results suggest that the effect of high versus low load on response times was similar across manipulations of task type (aesthetic vs. non-aesthetic) and medium type (artwork vs. photograph).

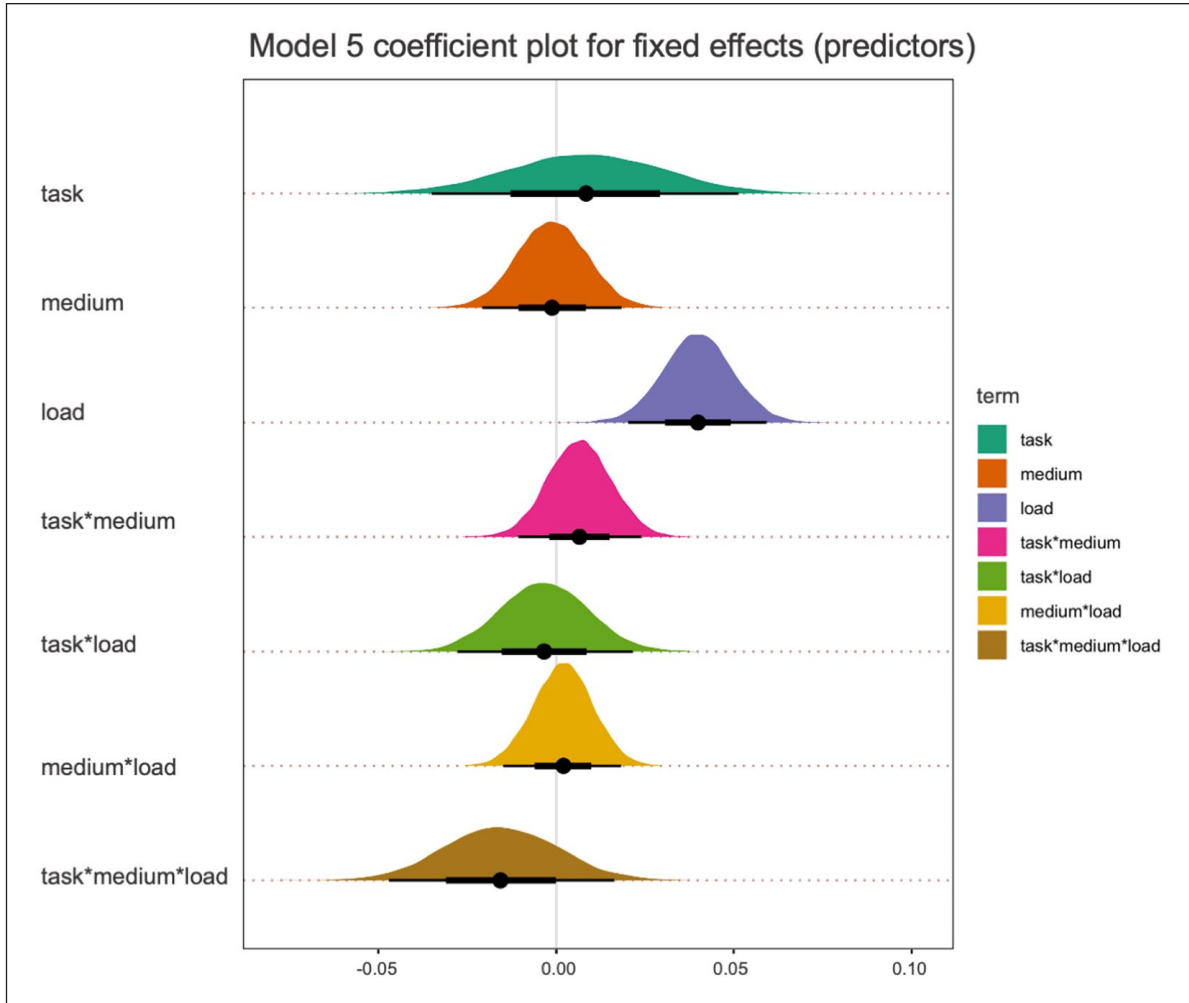
Model comparison analyses are visualised in Figure 8. All models with predictors performed better than the intercepts only model (Model b0), as well as the intercepts and varying effects model (Model b0.1g). Error bars for performance of the remaining models all overlapped,

suggesting that they performed in a largely similar manner, in terms of out-of-sample predictive accuracy.

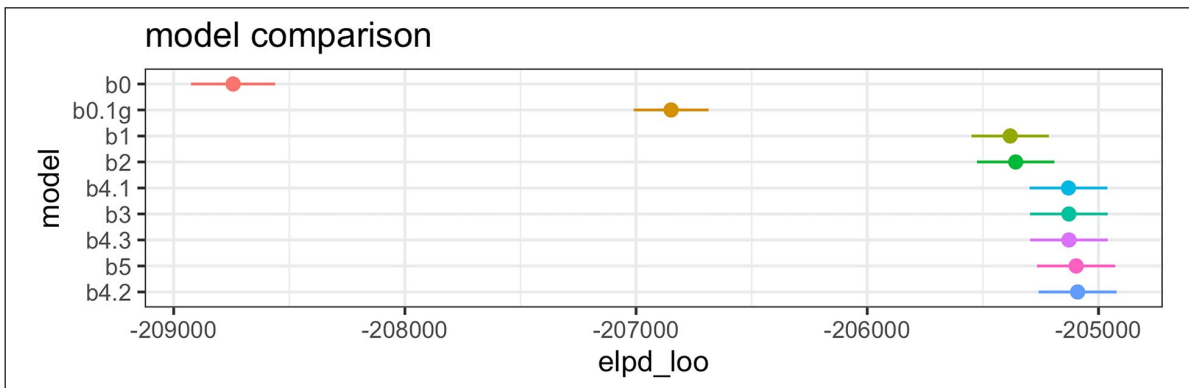
### Discussion

Experiment 2 showed that the effect of load did not vary by judgement type (aesthetic vs. motion) or image medium type (photos and paintings). These findings, therefore, provided further support for the hypothesis that the nature of working memory resources that underpin aesthetic judgements are largely similar as those deployed across a range of distinct judgement types and stimulus types.

In the next experiment, we made additional changes to the experimental procedure, to provide a further test of our general hypothesis. In Experiment 3, we modified the content of working memory load from letters to images of visual artworks. By changing the load content, we were able to probe how different aspects of working memory (from verbal in Experiments 1 and 2 to visual in Experiment



**Figure 7.** Parameter estimates for each predictor within Model 5. The main predictor that shows a clear positive effect is the load (third predictor). The x-axis is expressed on the log(RT) scale. The direct interpretation of these parameters in terms of response times is complex as the shifted lognormal model is made of three components. To see estimates of these effects in original units (ms), please see Supplementary Figure S4. In addition, the varying effects by stimulus and by participant can be visualised in Supplementary Figure S5.  
 task: judgement type (motion vs. aesthetic); medium: image type (photos vs. paintings); load: high vs. low. Point estimate=median. Error bars represent 66% quantile intervals (thick black lines) and 95% quantile intervals (thin black lines).



**Figure 8.** Model comparison. Models (b1–b5) performed better than the intercepts only model (b0) and intercepts and varying effects model (b0.1g). Labels for models (1–9) are similar to Figure 5, except the term stimulus “type” which has been replaced by “medium” = photo vs. painting.



3) impacts aesthetic judgements compared with non-aesthetic judgements. Given that image medium variation (photos vs. paintings) did not increase the sensitivity to interference effects, in Experiment 3 we used art stimuli only in the main experimental tasks.

## Experiment 3

### Introduction

In Experiment 3, we addressed the contribution of different modality-specific components of the working memory resources by changing the content of working memory load. Previous models of working memory have distinguished between verbal working memory, such as the phonological loop, which is responsible for managing speech-based information, and visual working memory, such as the visuospatial sketchpad, which is involved in maintaining and manipulating visuospatial imagery (Allen et al., 2017; Baddeley, 1992, 2012). As such, using letters as working memory content in Experiments 1 and 2 loaded verbal working memory and enabled verbal rehearsal subprocesses to occur. In contrast, in Experiment 3, we used art images as working memory content to load visual working memory and object feature-related subprocesses. The main purpose of using paintings instead of letters as load content was to increase the domain overlap between working memory load content and main tasks' stimuli content. We reasoned that increasing domain overlap in terms of art features would make it more likely that interference would be greater for aesthetic than non-aesthetic judgements.

### Method

**Pre-registration.** We used the same design and analysis pipeline as in Experiments 1 and 2, all of which we pre-registered in advance of the experiment commencing. The pre-registration document for Experiment 3 can be found at <https://aspredicted.org/mf85z.pdf>.

**Participants.** One hundred and one participants completed this experiment for course credit (20 males,  $M_{age} = 21.83$ ,  $SD_{age} = 4.80$ , age range = 18–43). The experiment was granted ethical approval by the Research Ethics and Governance Committee of the School of Human and Behavioural Sciences at Bangor University. All participants provided informed consent before completing the experiment. The pre-registered exclusion criteria were identical to the first two experiments: data files were excluded if an average working memory accuracy was less than 55% and trials with less than a 100 ms response time on judgement task were also excluded. The final sample included 97 participants and a total of 15,052 trials.

**Stimuli, task, and procedure.** The stimuli and tasks were similar to Experiment 1 with the following exception: for

the working memory load manipulation, we used still-life paintings instead of letters (see Figure 2). The high-load conditions consisted of the presentation of four still-life paintings in a circular arrangement, whereas the low-load conditions consisted of the presentation of one still-life painting. Participants were informed to memorise the still-life paintings during the retention period and then to indicate whether the memory probe still-life painting was present or absent at the beginning of each trial. The still-life paintings stimuli depicted 10 different vases of flowers by French artist, Odilon Redon (1840–1916). For a complete description of the load content stimuli, including the list of artworks, artists, year of production, and museum collection, see the Supplementary materials (Table S2). The still-life paintings stimuli that we used for the load content are available on our open science framework page (<https://osf.io/9q5jx/>).

The match between load content stimuli and main tasks' stimuli content was carefully balanced. In terms of similarities, both load content and main tasks' stimuli content were artworks described by a realistic pictorial style. However, the main difference referred to the subject matter; while the memory load content depicted still-life art—vases of flowers, the stimuli in the main tasks described landscape and people in dynamic and static postures.

For exploratory purposes, at the end of the main task, participants completed a short questionnaire about their memory strategies used during the task. The results are reported in Supplementary material (Figure S12). We also explored the relationship between the aesthetic judgement choice and image dynamism type (dynamic vs. static). The results are reported in Supplementary material (Figure S10.C).

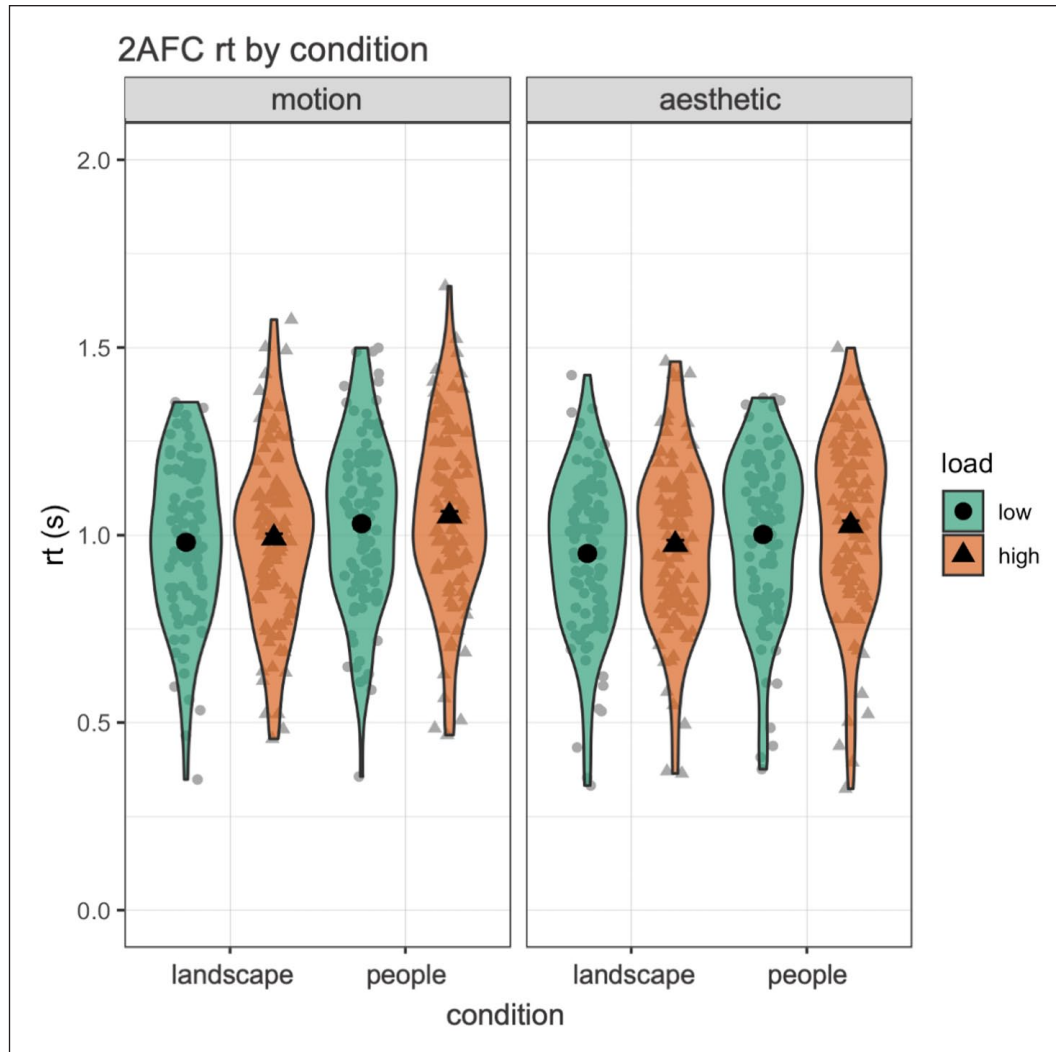
**Data analyses.** We used the identical approach to data analysis as performed in Experiment 1.

### Results

**Working memory.** Results indicated slower response time for high-load conditions compared with low-load conditions (Mean difference = 100 ms, 95% CI = [70, 102]). Also, we found lower memory accuracy for high-load conditions compared with low-load conditions (Mean difference = 21.42% accuracy, 95% CI = [19.52, 23.33]). For more details, please see Supplementary Figure S9.

**2-AFC task.** The 2-AFC response time results are illustrated in Figure 9. First, on average across participants, a greater response time was observed for high-load conditions compared with low-load conditions. Second, we see that motion judgements took longer than aesthetic judgements.

Parameter estimates for the most complex model (Model 5) are shown in Figure 10 and Table 2. The posterior distribution for the effect of judgement task showed a



**Figure 9.** Results for Experiment 3—violin plots on summary data showing 2-AFC response time. Response time is reported in seconds (s). The left panel shows response times for motion judgement task on low and high load conditions for both landscape and people. The right panel shows response times for aesthetic judgement task on low and high load conditions for both landscape and people.

Error bars represent 95% confidence intervals. The black markers (circles or squares) and interval estimates represent the group mean average, whereas the grey markers (circles or squares) represent the individual participants.

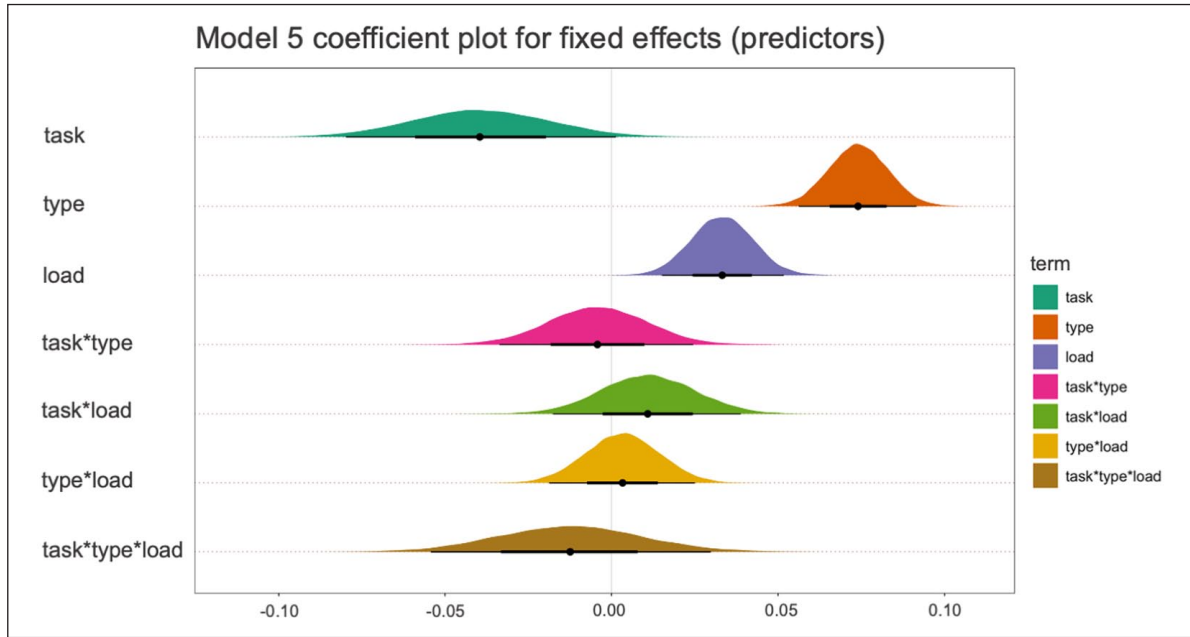
clear difference with motion judgements taking longer than aesthetic judgements. In addition, the parameter estimates indicated a largely positive response for image type (people vs. landscape) and for the effect of load (high vs. low). As can be seen in Supplementary Figure S7, the model estimates for these effects in response times are approximately 60 ms for the effect of type and between 10 and 40 ms for the effect of load. These results show that response times were slower for people than landscapes and high versus low load conditions. In addition, consistent with Experiments 1 and 2, the distributions for the interaction terms all showed substantial overlap with zero. These interaction effect results suggest that the effect of high versus low load on response times was similar across

manipulations of task type (aesthetic vs. non-aesthetic) and image type (people vs. landscape).

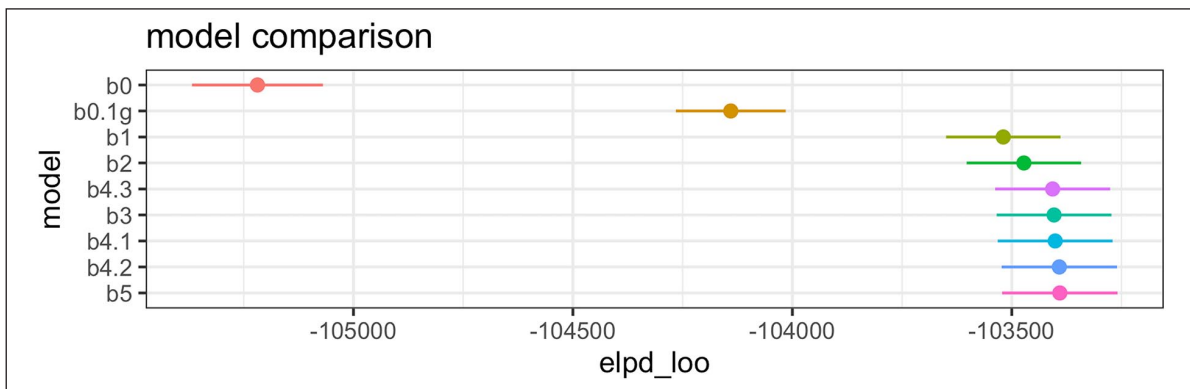
Model comparison analyses are visualised in Figure 11. All models with predictors performed better than the intercepts only model (Model b0), as well as the intercepts and varying effects model (Model b0.1g). Error bars for performance of the remaining models all overlapped, suggesting that they performed in a largely similar manner, in terms of out-of-sample predictive accuracy.

### Discussion

The results from Experiment 3 confirmed and extended the general pattern of findings from Experiments 1 and 2. The



**Figure 10.** Parameter estimates for each predictor within Model 5. The main predictors that show a clear effect are the task (motion vs. aesthetic), the type of image (people vs. landscape), and the load (high vs. low). The x-axis is expressed on the log(RT) scale. The direct interpretation of these parameters in terms of response times is complex as the shifted lognormal model is made of three components. To see the estimates of these effects in original units (ms), please see Supplementary Figure S7. In addition, the varying effects by stimulus and by participant can be visualised in Supplementary Figure S8. Point estimate=median; error bars represent 66% quantile intervals (thick black lines) and 95% quantile intervals (thin black lines).



**Figure 11.** Model comparison. Models (b1–b5) performed better than the intercepts only model (b0) and intercepts and varying effects model (b0.1g). The labels used for the models (1–9) are similar to Figure 5.

primary result showed that even when there is greater feature overlap between load content and the main task (compared with Experiments 1 and 2), there remains a similar deployment of working memory resources while making aesthetic judgements compared with non-aesthetic judgements. Much like Experiment 1, the current results reaffirmed that art images describing people required longer response time than art landscape, suggesting a different temporal course for aesthetic judgement depending on the subject matter. One possible explanation could be that

paintings describing people are perceived as visually more complex scenes than their landscape counterparts and they might need longer time for an aesthetic response. However, more research is needed to confirm this suggestion.

### General discussion

The main objective of the present study was to investigate the extent to which domain-general or domain-specific working memory resources support aesthetic judgements.

Across three pre-registered experiments, we found clear evidence that increasing working memory load produces similar response time interference on aesthetic judgements relative to non-aesthetic (motion) judgements. We also showed that this similarity in processing across aesthetic versus non-aesthetic judgements holds across variation in the form of art (people vs. landscape), medium type (artwork vs. photographs), and load content (art images vs. letters). These findings, therefore, suggest that across a range of experimental contexts, aesthetic and motion judgements rely on domain-general working memory mechanisms, rather than mechanisms that are more specifically tied to aesthetic contexts. In doing so, these findings show a pattern of results that generalises across a range of stimulus features and task conditions and shines new light on the cognitive structures that support aesthetic judgements.

### *Extension to dual-task research on aesthetics*

The current findings extend prior aesthetics research using dual-task paradigms. Prior work using dual-task paradigms addressed the role of executive control resources across aesthetic contexts only (Briellmann & Pelli, 2017; Che et al., 2021; Mullennix et al., 2013, 2016). In contrast, here we use a dual-task paradigm to compare between aesthetic and non-aesthetic categories of judgement. By finding a similarly-sized effect of load on interference across aesthetic and non-aesthetic judgements, it can be inferred that the degree to which resource-intensive compared with resource-light cognitive processes are deployed are largely the same across aesthetic and non-aesthetic contexts. Taken together, we can see that although variations in aesthetic tasks and stimuli can differentially engage executive resources (Briellmann & Pelli, 2017; Che et al., 2021), our results nonetheless suggest that when contrasting aesthetic to non-aesthetic judgement, such processing may still reflect the operations of a largely general set of executive systems.

### *Theoretical impact: specialised versus generalist accounts of aesthetic experience*

Understanding the form and structure of executive control that is deployed during aesthetic judgements has theoretical impact for cognitive models of aesthetic information processing, as well as our understanding of cognition more generally. Reliance on domain-general executive control mechanisms in both aesthetic and non-aesthetic contexts provides empirical evidence for the proposal that the underlying cognitive mechanisms that support aesthetic appraisal are comparable to those that support general-purpose behaviour (Bara, Binney, et al., 2021). More generally, these findings provide support for broader theoretical models from social and cognitive neuroscience,

which emphasise the role played by domain-general executive systems in information processing (Barrett, 2012; Binney & Ramsey, 2020; Duncan, 2010; Ramsey & Ward, 2020; Spunt & Adolphs, 2017). In contrast, we provide no support for accounts of aesthetic information processing that propose roles for partly distinct mechanisms between aesthetic and non-aesthetic contexts (Goldman, 2001; Goyer, 2005). This, of course, does not imply that there are no aesthetic contexts where specialised forms of executive control may be relied upon. Instead, we simply show a series of different task contexts and stimulus features, which rely on generalised forms of working memory resources.

### *Limitations and constraints on generality*

Due to the nature of the 2-AFC judgement task that we used, which does not include a “correct” answer but instead reflects a personal judgement, it can be difficult to verify the degree to which each button-press accurately corresponds to a true aesthetic or non-aesthetic judgement. In Experiments 1 and 2, for example, the type of task (aesthetic vs. non-aesthetic) had no overall impact on response times. It is possible, therefore, that participants were not actually making a meaningful judgement in the aesthetic versus motion task, but instead just pressing buttons at an appropriate time. However, in Experiment 3, there was a difference in response time between tasks, which suggests that distinctive judgements were being made, and yet the primary results remained the same as Experiments 1 and 2. This provides greater confidence that a different judgement was being made, but that it relied on a common form of working memory resources. Moreover, we have used the same stimuli in previous research, and they led to distinctive judgements (Bara, Darda, et al., 2021). In addition, the observed levels of accuracy on the load task demonstrate that participants were paying close attention to other aspects of the task. On balance, therefore, we feel that we have sufficient evidence to suggest that it is likely that participants were making distinctive judgements between task conditions.

As previously suggested by Simons et al. (2017), it is also important to recognise relevant constraints on the generality of our findings. Even though we find evidence for a generalised form of working memory in the current experiments that operates across a range of stimulus features and task conditions, we cannot rule out that there are distinct forms of working memory resources deployed in other aesthetic contexts. We can only assert that as tested in the current work, there is no evidence for specialised processing. In addition, we acknowledge that working memory resources might operate differently across art experts or in naturalistic contexts, such as art galleries. Therefore, of particular interest for future work would be to test how working memory resources operate in aesthetic and non-aesthetic judgements across real-world environments.



Furthermore, in this current work, we conceptualise working memory load according to Lavie's framework (Lavie et al., 2004, 2005, 2010). However, we acknowledge that other approaches exist (e.g., Musslick & Cohen, 2021), and future research may consider competing frameworks to conceptualise and investigate different working memory load predictions across aesthetic and non-aesthetic contexts. For example, in the current work, we have primarily focused on the maintenance function of the working memory, and we cannot rule out the possibility of increased cognitive interference in concurrent tasks targeting the manipulation function of working memory. Therefore, investigating how different working memory functions might operate under art and non-art stimuli and on aesthetic and non-aesthetic judgements would represent a valuable research avenue.

### Acknowledgements

This research was performed as part of an all-Wales Doctoral Training Centre PhD studentship (awarded to R.R. and I.B., PhD student: I.B.). We thank Andrew Wildman for help with developing the PsychoPy task.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Economic and Social Research Council (ESRC)

### ORCID iD

Inela Bara  <https://orcid.org/0000-0001-8442-5089>

### Data accessibility statement



The data and materials from the present experiment are publicly available at the Open Science Framework website: <https://osf.io/9q5jx/>.

### Supplementary material

The supplementary material is available at [qjep.sagepub.com](http://qjep.sagepub.com).

### References

- Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2017). Executive and perceptual distraction in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 43(9), 1677–1693. <https://doi.org/10.1037/xhp0000413>
- Augustin, M. D., Leder, H., Hutzler, F., & Carbon, C. C. (2008). Style follows content: On the microgenesis of art perception. *Acta Psychologica*, 128(1), 127–138. <https://doi.org/10.1016/j.actpsy.2007.11.006>
- Augustin, M. D., Wagemans, J., & Carbon, C. C. (2012). All is beautiful? Generality vs. specificity of word usage in visual aesthetics. *Acta Psychologica*, 139(1), 187–201. <https://doi.org/10.1016/j.actpsy.2011.10.004>
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556–559. <https://doi.org/10.1126/science.1736359>
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Bara, I., Binney, R. J., Ward, R., & Ramsey, R. (2021). A generalised semantic cognition account of aesthetic experience. Pre-print. <https://doi.org/10.31234/osf.io/wbeqx>
- Bara, I., Darda, K. M., Kurz, A. S., & Ramsey, R. (2021). Functional specificity and neural integration in the aesthetic appreciation of artworks with implied motion. *European Journal of Neuroscience*, 54(9), 7231–7259. <https://doi.org/10.1111/EJN.15479>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Barrett, H. C. (2012). A hierarchical model of the evolution of human brain specializations. *Proceedings of the National Academy of Sciences*, 109(Suppl. 1), 10733–10740. <https://doi.org/10.1073/pnas.1201898109>
- Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, 34(1), 149–159. [https://doi.org/10.1016/S0896-6273\(02\)00642-6](https://doi.org/10.1016/S0896-6273(02)00642-6)
- Berlyne, D. E. (1971). *Aesthetics and psychobiology*. Appleton-Century-Crofts.
- Binney, R. J., & Ramsey, R. (2020). Social semantics: The role of conceptual knowledge and cognitive control in a neurobiological model of the social brain. *Neuroscience and Biobehavioral Reviews*, 112, 28–38. <https://doi.org/10.1016/j.neubiorev.2020.01.030>
- Boccia, M., Barbetti, S., Piccardi, L., Guariglia, C., Ferlazzo, F., Giannini, A. M., & Zaidel, D. W. (2016). Where does brain neural activation in aesthetic responses to visual art occur? Meta-analytic evidence from neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 60, 65–71. <https://doi.org/10.1016/j.neubiorev.2015.09.009>
- Brieber, D., Forster, M., & Leder, H. (2020). On the mutual relation between art experience and viewing time. *Psychology of Aesthetics, Creativity, and the Arts*, 14(2), 197–208. <https://doi.org/10.1037/aca0000210>
- Briellmann, A. A., & Pelli, D. G. (2017). Beauty requires thought. *Current Biology*, 27(10), 1506–1513.e3. <https://doi.org/10.1016/j.cub.2017.04.018>
- Brown, T. A., & Barlow, D. H. (2009). A proposal for a dimensional classification system based on the shared features of the DSM-IV anxiety and mood disorders: Implications for assessment and treatment. *Psychological Assessment*, 21(3), 256–271. <https://doi.org/10.1037/A0016608>
- Bürkner, P. C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>

- Bürkner, P. C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *R Journal*, *10*(1), 395–411. <https://doi.org/10.32614/rj-2018-017>
- Cattaneo, Z. (2019). Noninvasive brain stimulation: An overview of available approaches for research in neuroaesthetics. *Empirical Studies of the Arts*, *37*(2), 153–171. <https://doi.org/10.1177/0276237418818637>
- Chatterjee, A. (2003). Prospects for a cognitive neuroscience of visual aesthetics. *Bulletin of Psychology and the Arts*, *8*(2), 55–60. <https://doi.org/10.1037/e514602010-003>
- Chatterjee, A., & Vartanian, O. (2014). Neuroaesthetics. *Trends in Cognitive Sciences*, *18*, 370–375. <https://doi.org/10.1016/j.tics.2014.03.003>
- Che, J., Sun, X., Skov, M., Vartanian, O., Rosselló, J., & Nadal, M. (2021). The role of working memory capacity in evaluative judgments of liking and beauty. *Cognition and Emotion*, *35*, 1407–1415. <https://doi.org/10.1080/02699931.2021.1947781>
- Chuan-Peng, H., Huang, Y., Eickhoff, S. B., Peng, K., & Sui, J. (2020). Seeking the “beauty center” in the brain: A meta-analysis of fMRI studies of beautiful human faces and visual art. *Cognitive, Affective and Behavioral Neuroscience*, *20*(6), 1200–1215. <https://doi.org/10.3758/s13415-020-00827-z>
- Cross, E. S., & Ramsey, R. (2021). Mind meets machine: Towards a cognitive science of human–machine interactions. *Trends in Cognitive Sciences*, *25*, 200–212. <https://doi.org/10.1016/j.tics.2020.11.009>
- Davies, S. (2012). *The artful species: Aesthetics, art, and evolution*. Oxford University Press.
- Dissanayake, E. (1995). *Homo aestheticus: Where art comes from and why*. University of Washington Press.
- Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*, 172–179. <https://doi.org/10.1016/j.tics.2010.01.004>
- Dutton, D. (2009). *The art instinct. Beauty, pleasure, and human evolution*. Bloomsbury Press.
- Fechner, G. T. (1876). *Vorschule der Ästhetik (Propaedeutics of Aesthetics)*. Breitkopfand Härtel.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). *Bayesian Analysis*, *1*, 515–534. <https://doi.org/10.1214/06-BA117A>
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). <https://doi.org/10.1201/b16018>
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Goldman, A. (2001). *The Routledge companion to aesthetics* (B. Gaut & D. Mcleaver, Eds.). Routledge.
- Graf, L. K. M., & Landwehr, J. R. (2015). A dual-process perspective on fluency-based aesthetics: The pleasure-interest model of aesthetic liking. *Personality and Social Psychology Review*, *19*(4), 395–410. <https://doi.org/10.1177/1088868315574978>
- Guyer, P. (2005). *Values of beauty: Historical essays in aesthetics*. Cambridge University Press.
- Haines, N., Kvam, P.D., Irving, L.H., Smith, C., Beauchaine, T.P., Pitt, M.A., Ahn, W.Y. and Turner, B.M. (2020). Theoretically informed generative models can advance the psychological and brain sciences: Lessons from the reliability paradox. *Pre-print*. <https://psyarxiv.com/xr7y3/>
- Hoffmann, D. L., Standish, C. D., Garcia-Diez, M., Pettitt, P. B., Milton, J. A., Zilhao, J., & Pike, A. W. G. (2018). U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art. *Science*, *359*, 912–915. <https://doi.org/10.1126/science.aap7778>
- Iigaya, K., O’Doherty, J. P., & Starr, G. G. (2020). Progress and promise in neuroaesthetics. *Neuron*, *108*(4), 594–596. <https://doi.org/10.1016/j.neuron.2020.10.022>
- Jacobsen, T. (2006). Bridging the arts and sciences: A framework for the psychology of aesthetics. *Leonardo*, *39*(2), 155–162. <https://doi.org/10.1162/leon.2006.39.2.155>
- Jacobsen, T., & Höfel, L. (2003). Descriptive and evaluative judgment processes: Behavioral and electrophysiological indices of processing symmetry and aesthetics. *Cognitive, Affective and Behavioral Neuroscience*, *3*(4), 289–299. <https://doi.org/10.3758/CABN.3.4.289>
- Jacobsen, T., Schubotz, R. I., Höfel, L., & Cramon, D. Y. V. (2006). Brain correlates of aesthetic judgment of beauty. *NeuroImage*, *29*(1), 276–285. <https://doi.org/10.1016/j.neuroimage.2005.07.010>
- Kirsch, L. P., Urgesi, C., & Cross, E. S. (2016). Shaping and reshaping the aesthetic brain: Emerging perspectives on the neurobiology of embodied aesthetics. *Neuroscience and Biobehavioral Reviews*, *62*, 56–68. <https://doi.org/10.1016/j.neubiorev.2015.12.005>
- Kurz, A. S. (2020). *Statistical Rethinking with brms, ggplot2, and the tidyverse* (2nd ed.). <https://doi.org/10.5281/zenodo.4302550>
- Lamarque, P. (1999). The aesthetic and the universal. *Source: The Journal of Aesthetic Education*, *33*, 1–17. <https://about.jstor.org/terms>
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*, 75–82. <http://dx.doi.org/10.1016/j.tics.2004.12.004>
- Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*, 143–148. <http://dx.doi.org/10.1177/0963721410370295>
- Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. <https://doi.org/10.1037/0096-3445.133.3.339>
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*, 489–508.
- Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode—Developments and challenges in empirical aesthetics. *British Journal of Psychology*, *105*(4), 443–446. <https://doi.org/10.1111/bjop.12084>
- Lemoine, N. P. (2019). Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, *128*(7), 912–928. <https://doi.org/10.1111/oik.05985>
- Locher, P., Krupinski, E. A., Mello-Thoms, C., & Nodine, C. F. (2007). Visual interest in pictorial art during an aesthetic experience. *Spatial Vision*, *21*, 55–77.
- Locher, P., Overbeeke, K., & Wensveen, S. (2010). Aesthetic interaction: A framework. *Design Issues*, *26*(2), 70–79. [https://doi.org/10.1162/DESI\\_a\\_00017](https://doi.org/10.1162/DESI_a_00017)

- Marshack, A. (1996). A Middle Paleolithic symbolic composition from the Golan Heights: The earliest known depictive image. *Current Anthropology*, 37(2), 357–365.
- Martindale, C., Locher, P., Petrov, V. M., & Berleant, A. (2019). *Evolutionary and neurocognitive approaches to aesthetics, creativity and the arts*. Routledge.
- Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings of the Royal Society B: Biological Sciences*, 249, 149–155. <https://royal-societypublishing.org/>
- Maunsell, J. H. R., & Van Essen, D. C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, 3(12), 2563–2586. <https://doi.org/10.1523/jneurosci.03-12-02563.1983>
- McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2nd ed.). <https://doi.org/10.4324/9780429029608>
- Morriss-Kay, G. M. (2010). The evolution of human artistic creativity. *Journal of Anatomy*, 216(2), 158–176. <https://doi.org/10.1111/j.1469-7580.2009.01160.x>
- Mullennix, J. W., Foytik, L. R., Chan, C. H., Dragun, B. R., Maloney, M., & Polaski, L. (2013). Automaticity and the processing of artistic photographs. *Empirical Studies of the Arts*, 31(2), 145–171. <https://doi.org/10.2190/EM.31.2.c>
- Mullennix, J. W., Varnecky, A., Chan, C. H., Korenoski, S., Mickey, Z., & Polaski-Hoffman, L. (2016). The effect of color on automaticity of aesthetic judgments. *Empirical Studies of the Arts*, 34(1), 8–34. <https://doi.org/10.1177/0276237415621183>
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., . . . Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1, Article 0021. <https://doi.org/10.1038/s41562-016-0021>
- Musslick, S., & Cohen, J. D. (2021). Rationalizing constraints on the capacity for cognitive control. *Trends in Cognitive Sciences*, 25(9), 757–775. <https://doi.org/10.1016/j.tics.2021.06.001>
- Nadal, M., & Chatterjee, A. (2019). Neuroaesthetics and art's diversity and universality. *Wiley Interdisciplinary Reviews: Cognitive Science*, 10(3), Article e1487. <https://doi.org/10.1002/wcs.1487>
- Nadal, M., & Skov, M. (2015). Neuroaesthetics. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 656–663). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.56027-9>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087–11092. <https://doi.org/10.1073/PNAS.0805664105>
- Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology*, 64(1), 77–107. <https://doi.org/10.1146/annurev-psych-120710-100504>
- Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics: The cognitive neuroscience of aesthetic experience. *Perspectives on Psychological Science*, 11(2), 265–279. <https://doi.org/10.1177/1745691615621274>
- Pearce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Pelowski, M., & Akiba, F. (2011). A model of art perception, evaluation and emotion in transformative aesthetic experience. *New Ideas in Psychology*, 29(2), 80–97. <https://doi.org/10.1016/j.newideapsych.2010.04.001>
- Ramsey, R., & Ward, R. (2020). Putting the nonsocial into social neuroscience: A role for domain-general priority maps during social interactions. *Perspectives on Psychological Science*, 15, 1076–1094. <https://doi.org/10.1177/1745691620904972>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Redies, C. (2015). Combining universal beauty and cultural context in a unifying model of visual aesthetic experience. *Frontiers in Human Neuroscience*, 9, Article 218. <https://doi.org/10.3389/fnhum.2015.00218>
- Satpute, A. B., & Lieberman, M. D. (2006). Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Research*, 1079(1), 86–97. <https://doi.org/10.1016/j.brainres.2006.01.005>
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, 12(6), 1123–1128. <https://doi.org/10.1177/1745691617708630>
- Specker, E., Forster, M., Brinkmann, H., Boddy, J., Pelowski, M., Rosenberg, R., & Leder, H. (2020). The Vienna Art Interest and Art Knowledge Questionnaire (VAIAK): A unified and validated measure of art interest and art knowledge. *Psychology of Aesthetics, Creativity, and the Arts*, 14(2), 172–185. <https://doi.org/10.1037/aca0000205.supp>
- Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: Insights from social neuroscience. *Nature Reviews Neuroscience*, 18, 559–567. <https://doi.org/10.1038/nrn.2017.76>
- Van de Cruys, S., & Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. *I-perception*, 2(9), 1035–1062. <https://doi.org/10.1068/i0466aap>
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Vessel, E. A., Maurer, N., Denker, A. H., & Starr, G. G. (2018). Stronger shared taste for natural aesthetic domains than for artifacts of human culture. *Cognition*, 179, 121–131. <https://doi.org/10.1016/j.cognition.2018.06.009>
- White, R., Mensan, R., Bourrillon, R., Cretin, C., Higham, T. F. G., Clark, A. E., & Chiotti, L. (2012). Context and dating of Aurignacian vulvar representations from Abri Castanet, France. *Proceedings of the National Academy of Sciences*, 109, 8450–8455. <https://doi.org/10.1073/pnas.1119663109>
- Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language*, 29(6), 716–732. [https://doi.org/10.1016/0749-596X\(90\)90045-2](https://doi.org/10.1016/0749-596X(90)90045-2)
- Wickham, H., & Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data*. O'Reilly Media.
- Zeki, S. (1999). Art and the brain. *Journal of Consciousness Studies*, 6(6–7), 76–96.