Cognitive, Affective, & Behavioral Neuroscience (2024) 24:453–468
https://doi.org/10.3758/s13415-024-01158-z

RESEARCH ARTICLE

Investigating the effects of perceptual complexity versus conceptual meaning on the object benefit in visual working memory

Alyssa M. L. Thibeault[1] · Bobby Stojanoski[2] · Stephen M. Emrich[1]

Accepted: 6 January 2024 / Published online: 30 January 2024

Abstract

Previous research has demonstrated greater visual working memory (VWM) performance for real-world objects compared with simple features. Greater amplitudes of the contralateral delay activity (CDA)—a sustained event-related potential measured during the delay period of a VWM task—have also been noted for meaningful stimuli, despite being thought of as a neural marker of a fixed working memory capacity. The current study aimed to elucidate the factors underlying improved memory performance for real-world objects by isolating the relative contributions of perceptual complexity (i.e., number of visual features) and conceptual meaning (i.e., availability of semantic, meaningful features). Participants ($N = 22$) performed a lateralized VWM task to test their memory of intact real-world objects, scrambled real-world objects and colours. The CDA was measured during both encoding and WM retention intervals (600–1000 ms and 1300–1700 ms poststimulus onset, respectively), and behavioural performance was estimated by using *d'* (memory strength in a two-alternative forced choice task). Behavioural results revealed significantly better performance within-subjects for real-world objects relative to scrambled objects and colours, with no difference between colours and scrambled objects. The amplitude of the CDA was also largest for intact real-world objects, with no difference in magnitude for scrambled objects and colours, during working memory maintenance. However, during memory encoding, both the colours and intact real-world objects had significantly greater amplitudes than scrambled objects and were comparable in magnitude. Overall, findings suggest that conceptual meaning (semantics) supports the memory benefit for real-world objects.

Keywords Visual working memory · Contralateral delay activity · Object memory benefit · Conceptual meaning · Perceptual complexity

Introduction

With a virtually limitless amount of visual information available in our environment, the qualitative and quantitative properties of what we can remember is a major focus of memory research. Unlike visual long-term memory (LTM) in which storage capacity is immense (Brady et al., 2008), visual working memory (VWM)—a cognitive system involved in the maintenance and manipulation of visual sensory input (Baddeley, 2012)—is deemed to be capacity-limited, in that few items can be actively represented at any given time.

Although most theories of VWM agree that the system is limited in its overall capacity, evidence suggests that the amount of information that may be actively maintained differs depending on the nature of the to-be-remembered stimuli. In one study, Alvarez & Cavanagh (2004) sought to dissociate whether the capacity of VWM was limited in absolute terms by the total number of items maintained or whether limitations arose at the level of the amount of visual information (i.e., item features) being stored. Using a change detection task, subjects' WM capacity for stimuli of varying complexity (e.g., colours, polygons, Chinese characters, shaded cubes, and letters) was assessed. Not only did capacity fluctuate across stimulus type, but items with more

This work was supported by a Natural Sciences and Engineering Research Council of Canada grant to S.M.E (NSERC 2019-04865) and an Ontario Graduate Scholarship to A.M.L.T.

✉ Alyssa M. L. Thibeault
wy20rd@brocku.ca

[1] Department of Psychology, Brock University, 1812 Sir Isaac Brock Way, St. Catharines, Ontario L2S 3A1, Canada

[2] Faculty of Social Science and Humanities, Ontario Tech University, 2000 Simcoe Street North, Oshawa, Ontario L1G 0C5, Canada

visual information tended to be the least well-remembered, suggesting that increases in visual (perceptual) complexity may in fact hinder working memory performance.

In contrast, other manipulations of stimulus complexity have failed to demonstrate a reduction in working memory performance with the addition of visual features. For example, Luck & Vogel (1997) demonstrated that capacity estimates in a change detection paradigm were comparable for the maintenance of separate individual item characteristics (e.g., the colour of 3–4 items) and whole object identities (i.e., conjunction of both the colour and orientation of 3–4 items), suggesting that increases in per-item detail may not straightforwardly impair working memory performance. Similarly, using a change detection task with shaded cubes and Chinese characters, as in Alvarez & Cavanagh (2004), Barton et al. (2009) found that mnemonic resolution of a given item was unaffected by the complexity of other to-be-encoded items when set size was held constant. Based on these results, the researchers concluded that while storage of more items may impact overall VWM performance, the resolution or precision of individual VWM representations are likely unaffected by relative variations in interitem complexity.

Separately from perceptual (visual) complexity, the degree of meaningfulness of to-be-remembered stimuli also has been shown to influence VWM maintenance and storage. In a study by Ngiam et al. (2019), a change detection task was used to examine both capacity and encoding limits of VWM for letters in various fonts and alphabets. Across three experiments, the researchers found that while increases in parametric complexity—an index of the number of features of the letters—had no influence on either encoding rate or measured capacity, a familiarity effect was observed. That is, participants not only remembered more English relative to non-English letters (with English fonts and alphabets as the familiar stimulus set) but also consolidated these letters into memory more rapidly. Increases in VWM capacity as a function of familiarity was also noted by Xie & Zhang (2017a). For participants highly familiar with first-generation Pokémon stimuli, marked increases in the number of remembered first-generation characters were observed relative to recent generation, less familiar characters. Importantly, the improvement in measured capacity for the first-generation versus recent-generation stimulus sets was not observed for participants unfamiliar with both generations of stimuli, suggesting that the availability of previous knowledge of the original characters bolstered VWM capacity for the high-familiarity subjects.

Familiarity effects on memory performance have also been demonstrated for face stimuli, with significantly greater capacity estimates for the memory of famous versus comparatively unfamiliar faces, even under conditions of a high concurrent verbal memory load (Jackson & Raymond, 2008). Likewise, using "Mooney" stimuli—some of which could be perceived as face shapes, and others that were instead ambiguous, perceptually-matched nonfaces—Asp et al. (2021) observed a memory advantage when the display consisted of a greater number of meaningful stimuli (i.e., faces), both in terms of behavioural performance estimates and amplitude of the contralateral delay activity (CDA). The CDA is a sustained negativity in the hemisphere contralateral to a to-be-remembered target during the delay phase of a VWM task and is measured over occipital-parietal scalp sites using electroencephalograms (EEG; Vogel & Machizawa, 2004; McCollough et al., 2007). The CDA is often deemed an index of active VWM maintenance and storage, as the amplitude of the waveform has been shown to increase alongside increases in memory load, and plateau once the number of remembered items exceeds an individual's specific memory capacity (Vogel et al., 2005; Luck & Vogel, 2013; but see also Bays, 2014; Emrich et al., 2022).

Meaningfulness has also been examined as it relates to differences in performance between different categories of stimuli (i.e., rather than familiar and unfamiliar exemplars of the same stimulus type, such as famous versus unfamiliar faces). For example, Brady et al. (2016) noted that following a long encoding interval of either a set of real-world objects or colours, memory performance (i.e., the number of items remembered) during a two-alternative forced choice task (i.e., 2AFC) was greater for real-world objects. In contrast to this finding however, Quirk et al. (2020) and Li et al. (2020) failed to replicate the object memory benefit relative to colours, with estimated capacity for both stimulus types increasing with prolonged encoding time. It has been noted in follow-up works however that mixed results regarding improved behavioural performance for real-world objects may have been driven by the use of non-comparable test foils in studies failing to replicate the effect. For example, Brady & Störmer (2023) found that when maximally dissimilar foils were presented at test—enabling greater experimental control over semantic and visual similarities between real-world objects—the memory advantage for objects persisted compared with memory for maximally dissimilar simple features (e.g., colours separated by 180°). In this way, failure to include maximally dissimilar pairs in previous studies (Quirk et al., 2020; Li et al., 2020) may have eliminated or considerably reduced the object benefit, possibly because of increased memory and test item confusability for the more complex stimuli (e.g., real-world objects).

It has been suggested that improved online storage of real-world objects compared with simple features is supported by connections to existing knowledge, such that the recruitment of additional neural resources allows for deeper processing of item representations in VWM. This proposal aligns with the theoretical construct of conceptual short-term memory (CSTM; Potter 1976, 2012), in which the perception of meaningful stimuli rapidly activates associated information in LTM, allowing for the establishment of conceptual patterns or structures. These structures, combined with selective attention towards conceptual characteristics of a target, are what reach the level of conscious awareness and are then able to be maintained in VWM and consolidated as a whole unit. Conversely, failure to integrate information into such structures yields memory representations that are unstable or incomplete by comparison, as thus, more susceptible to forgetting. Consistent with the idea of the role of conceptual processing, Brady & Störmer (2022) found that real-world objects—but not colours—were best remembered when presented sequentially (i.e., items presented one at a time, each at one of six locations around fixation) rather than simultaneously (i.e., all six items displayed at once, each in their respective location around fixation), suggesting the formation of more elaborate representations for meaningful objects. Moreover, using functional magnetic resonance imaging (fMRI), Stojanoski et al. (2019) observed heightened activation in higher-level brain regions (e.g., ventral visual areas) during the encoding of intact, meaningful objects—but not their unrecognizable counterparts—as well as a lower guess rate and greater behavioural precision.

Although the object memory benefit has been established in some previous works, several of these studies only compared memory performance for real-world objects and colours stimuli (Brady et al., 2016; Brady & Störmer, 2023; Quirk et al., 2020; Li et al., 2020). However, given intact real-world objects differ from simple stimuli such as colours in both their perceptual complexity *and* degree of conceptual meaning, the relative contributions of each component to observed changes in memory performance become difficult to isolate. In an effort to tease apart the role of previous knowledge on performance, some studies have included a third scrambled condition for comparison with intact complex stimuli and simple stimuli. However, in some of these studies, primarily behavioural measures of memory strength were of interest (e.g., as in Brady & Störmer, 2022), non-object stimulus sets were tested (e.g., such as faces in Asp et al., 2021), or the scrambling techniques used may not have effectively preserved perceptual complexity in the manipulated images (e.g., variation of a box scramble/inversion method in Asp et al., 2021; see Stojanoski & Cusack, 2014 for overview of scrambling methods).

Current study

The motivation for the current study was twofold: 1) to provide an experimentally controlled comparison of the influence of perceptual complexity (i.e., the number of visual features an object has) versus conceptual meaning (i.e., the number of semantic, meaningful features an object has) to the object memory benefit; and 2) to examine how these controlled manipulations of perceptual complexity versus conceptual meaning differentially affect changes in the amplitude of the CDA. As such, this work uniquely contributes to the literature by integrating the methodologies and stimulus sets used in the aforementioned influential works. Namely, the current study tested participants' VWM—measured behaviourally as memory strength (d') and via comparisons of CDA amplitude—for real-world objects, scrambled real-world objects (using diffeomorphic transformations; Stojanoski & Cusack, 2014), and colours, to examine the relative contributions of complexity and meaning to memory performance and the resulting CDA.

Given there is only one visual feature by which the items may be distinguished and remembered in the colours condition and given there is no inherent meaning associated with the colours (e.g., category membership, state, behaviour, function, etc.), this stimulus set comprised the low perceptual complexity and low conceptual meaning condition. Conversely, the intact real-world objects serve as the high perceptual complexity and high conceptual meaning condition relative to either scrambled real-world objects or colours. Namely, the objects are comprised of a greater number of visual features (e.g., variability in colour, shape, etc.) relative to the colours as only a single feature-based stimulus and are recognizable entities for which conceptual associations may be established based on past knowledge or experience. For example, viewing an image of a beagle may activate related exemplars in the same basic level category (e.g., other dog breeds) or a more general superordinate category (e.g., animals).

Whereas the colours and intact real-world object conditions serve as the low complexity, low meaning and high complexity, high meaning conditions, respectively, the critical experimental manipulation was the inclusion of the scrambled real-world objects condition. In order to examine whether increased perceptual (visual) complexity or increased meaning contributes to the memory advantage for real-world objects noted in previous works, a condition which shares one—but not both—factors with each of the colours and intact objects conditions is necessary. The scrambled objects condition in the current study therefore satisfies this requirement, in that the stimuli were systematically warped to render them unrecognizable, thereby negating the higher degree of conceptual meaning otherwise available in the intact objects condition. However, given the objects were warped using the diffeomorphing method (Stojanoski & Cusack, 2014; see *Method – Stimuli* for further detail), the degree of perceptual complexity between the scrambled and intact object sets was made comparable, despite the absence of conceptual meaning in the case of the former. Unlike other common scrambling techniques—including texture, phase, and box scrambling—diffeomorphic transformations have been shown to yield simulated neural responses to low-level visual features that are comparable to the neural responses produced by intact images. As such, the higher degree of perceptual complexity apparent in the intact objects condition (relative to the colours) also was maintained in the scrambled objects condition, allowing us to assess working memory performance for a low-meaning, high-complexity condition.

Based on evidence that VWM performance is flexible and may be improved when to-be-remembered items are meaningful, we hypothesized that conceptual meaning drives the observed memory benefit for real-world objects. More specifically, we predicted that: 1) across participants, the CDA amplitude and working memory performance (d-prime, *d'*) would be greatest for real-world objects relative to scrambled objects and colours; and 2) given that both the scrambled objects and colours stimuli lack conceptual meaning—relative to the semantic information that may be extracted from intact real-world objects—the CDA amplitude and behavioural performance (d-prime, *d'*) for these conditions would not differ significantly.

Method

Pre-registration

The experimental design was pre-registered on AsPredicted (https://aspredicted.org/fh73w.pdf). Several minor modifications were made post-hoc to the pre-registered design. Namely, because of reports of participant fatigue during piloting, the block design was adjusted to 30 blocks of 30 trials each rather than three blocks of 300 trials each allowing for frequent breaks. Additionally, the order of the blocks appearing to each participant were fully randomized in the final task design (i.e., seeded to subject number) rather than counterbalanced pseudo-randomly as reported in the pre-registration. In terms of the ERP analysis, participants with greater than an average of 35% of trials excluded were removed from the analysis, rather than greater than 35% of trials in any condition (scrambled/left, scrambled/right, real-world/left, real-world/right, colours/left, colours/right). EEG also was band-pass filtered with a high pass of 0.1 Hz rather than 0.01 Hz, based on similar work examining the CDA (Salahub et al., 2019). As well, in addition to only measuring the CDA during the working memory delay (1300–1700 ms post onset of the memory array) as noted in the pre-registration, the CDA during encoding (600–1000 ms) also was examined. This post-hoc analysis was conducted given differences were observed between conditions in advance of the delay period in the current study and given other recent work has also demonstrated CDA effects during the encoding of to-be-remembered stimuli (Asp et al., 2021; see Emrich et al., 2022 for review). Finally, behavioural performance is also reported as *d'* throughout the manuscript, rather than capacity (*K*) as noted in the pre-registration. Whereas changes in *K* estimates may have reflected changes in participants' decision thresholds for the different stimulus types rather than differences in memory accuracy per se, measuring memory strength (*d'*) may instead serve as a more informative and appropriate comparison of working memory performance.[1] A scatterplot depicting the relationship between *K* and *d'* values in the current dataset is included in Fig. S1 (see *Supplemental Materials*).

Participants

The experiment was approved by the Social Science Research Ethics Board at Brock University. Participants were all aged 17–40 years, right-handed, and had normal or corrected-to-normal vision, including normal colour vision. Exclusion criteria for participant recruitment also included currently taking psychoactive medications, having braids, hair extensions, or temporary hair dye and a history of mental illness, neurological problems, or head trauma.

Participants were recruited through the Brock University SONA undergraduate participant recruitment system, flyers, and via word of mouth. The final sample was comprised of 22 participants (M_{age} = 23.64 years, range 17–34 years, 12 females).[2] This sample size was determined based on samples used in similar work (Vogel & Machizawa, 2004; Brady et al., 2016). Participants were excluded from the analysis if behavioural performance in any condition was $d' \leq 0.5$ or if greater than 35% of trials were rejected during ERP pre-processing because of eye or muscle movements. Based on this criteria, an additional 11 participants completed the study but were excluded before analysis (low *d'* values: 2 participants; excessive EEG artifacts: 9 participants). Two other participants were excluded from the final analysis because of errors with the experimental setup and data collection procedures. All participants provided informed consent to participate in the study.

Stimuli

Object stimuli were selected from a set of 120 pairs of maximally dissimilar items used by Brady & Störmer (2023). Maximally dissimilar pairs were generated by using deep convolutional neural networks sensitive to both visual and semantic features of the items (see Brady & Störmer, 2023

[1] We thank an anonymous reviewer for this useful suggestion.

[2] The pre-registered size of the sample was 20 participants. However, because of difficulties with data collection and a high exclusion rate, the rejection window was modified to omit blinks in the later half of the trial, resulting in a final sample of 22 participants in our analysis.

for details). The scrambled objects condition was comprised of the same 120 pairs of images used in the real-world objects condition; however, items were warped by using the diffeomorphing method (Stojanoski & Cusack, 2014) to 31.25% of the maximum (25/80). In doing so, objects became unrecognizable, thereby allowing for the manipulation of conceptual meaning of the items while holding perceptual complexity constant between the real-world and scrambled objects conditions. Of the two versions of each item generated from the warping procedure (i.e., contracted versus expanded) selection of which to use in the stimulus set was pseudo-random, except in instances where the generated scrambled image contained disconnected segments (in which case the least distorted of the two versions was used). For both object conditions, the five filler items in the six-item memory array were randomly selected on a per trial basis from the remaining 2160 unpaired images included in the database. Additionally, given a grey background was used in the task to reduce eye strain, all images for both object sets were modified via a filter to remove the white negative space surrounding the images. All objects (i.e., intact and scrambled) were 2.5 degrees of visual angle in size throughout the task.

For the colours condition, RGB triplets were selected from the CIE $L^*a^*b^*$ colour space (RGB triplets used by Brady & Störmer, 2023). RGB triplets in the encoding array were randomly selected, with no constraints on the distance between sample values. However, as in the 2AFC for both object conditions, the target and comparison foil presented at test for the colours were also maximally dissimilar, such that colours were 180-degrees apart on the colour wheel. Given the colours stimuli did not contain negative space as in the case of both objects sets (i.e., not variable in their shape), colours stimuli were held constant at 1.25 degrees of visual angle in size. This allowed for the perceived size of the colours stimuli to be more comparable to the size of the intact and scrambled real-world objects following the removal of negative space surrounding the items. As well, given the total number of trials for each condition (i.e., 300) either exceeded or approached the total number of items of each stimulus type (i.e., 120 pairs of object stimuli, 360 RGB triplets for the colours), some stimuli were also repeated in each condition. Sample stimuli from each condition are depicted in Fig. 1A.

Procedure

The experimental task was completed on a Windows PC. PsychoPy v2021.2.3 (Peirce, 2007) was used to run the program (and generate square shapes for the colours condition), and stimuli were presented on a grey background (win.color = [0.25, 0.25, 0.25], win.size = [800, 600]). A central fixation cross (size: 2 degrees visual angle), cue dot (during the 2AFC; size: 0.5 degrees visual angle), and central arrow cue also were generated. Viewing distance was approximately 57 cm (roughly one arm's length from the monitor). The six stimuli in the memory array were drawn such that the centre of each image was situated at one of the following positions: [−3, −3], [−4.2, 0], [−3, 3], [3, −3], [4.2, 0], [3,3]. These coordinates produced a circular arrangement of the objects around the central fixation point (0,0).

The experiment was a within-subjects design, and participants completed a total of 150 trials per cue side (left/right) per stimulus type (colours, scrambled objects, real-world objects) for a total of 900 trials (150 trials x 6 conditions: left/scrambled, right/scrambled, left/real-world, right/real-world, left/colours, right/colours).[3] Trials were separated into 30 blocks of 30 trials each. Within a given block, trials were all of the same stimulus type (i.e., real-world, scrambled, colours); however, the order of blocks and the order of images appearing within a given block were randomized across participants. Following each block was a self-paced break. Time to completion for the experimental task was approximately 1.5 hours on average, with variability in completion time attributable to differences in the duration of breaks. One block of practice trials (30 trials) also preceded the main task to familiarize participants with the paradigm.

A schematic of a sample trial is depicted in Fig. 1B. Each trial began with the presentation of an arrow cue (500 ms) to indicate which subset of items were to be remembered (those to the left or to the right of a central fixation). A fixation cross followed the arrow cue for a random interval of time between 300−500 ms. Participants were then presented with either six colours, six scrambled objects, or six real-world objects for a 1000-ms encoding period, followed by a 700-ms delay period (during which only the fixation remained onscreen). For all conditions, stimuli were presented in a circular formation around the central fixation cross during the encoding phase, with three items in each visual hemifield. Thus, the load was held constant at three to-be-remembered items for the duration of the study. Participants were instructed to maintain their gaze on the central fixation throughout the encoding and delay periods. At test, participants completed a two-alternative forced choice task (2AFC), in which they indicated which of two images had been presented at the probed location in the initial memory array. Probed positions during the 2AFC were always one of the three locations on the cued side and the two forced-choice

[3] One participant ran out of time within the testing period to complete the study, and thus their data consisted of only 840 trials total. Five subjects included in the final sample also had an uneven number of trials per each stimulus type condition because of a programming error. As well, one of these five subjects had a considerably smaller number of trials overall than the rest of the sample (i.e., 635 vs. 900). However, reconducting all statistical analyses with this subject omitted yielded no significant change to the overall findings of the behavioural, early window, or late window ERP results, and this subject did not otherwise meet exclusion criteria for the study. As such, they were still included in the final sample reported in the main analyses.

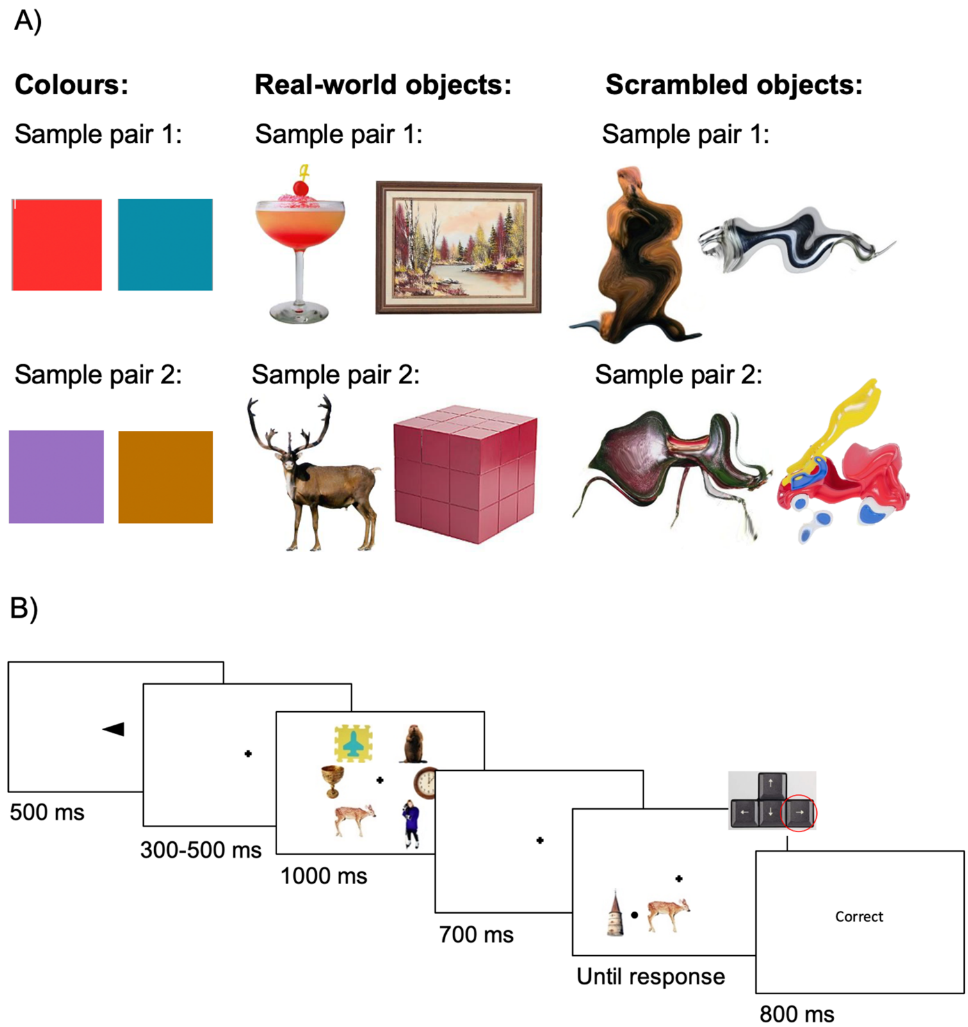

Fig. 1 **(A) Sample stimuli from each stimulus-type condition.** Examples of maximally dissimilar pairs from the colours, real-world objects, and scrambled objects conditions. Stimuli as shown are not scaled to size based on dimensions used in the task. Stimuli for all conditions were sampled from items used by Brady & Störmer (2023). **(B) Task schematic.** Trial (with correct response) for the real-world objects condition depicted. The memory array consisted of six real-world objects, six scrambled objects, or six colours. Participants completed 150 trials per condition for a total of 900 trials (150 per cue side (left, right) x 3 stimulus types (scrambled objects, real-world objects, colours))

options always consisted of the target image (i.e., the image/colour presented in the memory array) and the maximally dissimilar comparison foil. Images in the 2AFC were presented on either side of a cue dot, which was situated at the centre of the location the image had appeared in the encoding phase (coordinates under *Procedure*). Participants were instructed to select the left arrow key on the computer keyboard if they believed the leftmost item was the target and the right arrow key if they believed the rightmost item was the target. The position of the target and comparison images relative to the cue dot (i.e., left or right) was randomized on a per trial basis for all subjects. Although participants were asked to respond as quickly and as accurately as possible, the test phase lasted until a key press response was made. Following a key press response, feedback was presented at the centre of the screen ("Correct"; "Incorrect") for 800 ms before the next trial began.

Behavioural analysis

Visual working memory performance was estimated per participant per stimulus type condition using d-prime (*d'*) for a 2AFC task ($d' = [\text{zHits} - \text{zFalseAlarms}]/\sqrt{2}$). According to signal detection theory (Wickens, 2001; Schurgin et al., 2020), *d'* serves as a standardized measure of sensitivity or memory strength, wherein the degree of familiarity with studied stimuli (i.e., signal) relative to extraneous information (i.e., corruption by noise) influences whether a target item is detected at test. Thus, in the context of the current study, *d'* values reflect the extent to which participants remembered the target object or colour well enough to distinguish it from a same-category foil (i.e., a foil of the same stimulus type) at test. In other words, whether the encoded object or colour was maintained well enough in memory to generate the maximum familiarity signal at test relative to competing foil objects or colours (in which case the target item would be correctly identified in the 2AFC; i.e., a hit). Conversely, if the strength of the target memory representation was insufficient to distinguish it from a competing item, the foil object/colour may instead be selected (i.e., false alarm). Behavioural performance estimates for each condition were statistically compared by using a repeated measures analysis of variance (RM ANOVA) and paired *t*-tests.

Electrophysiological analysis

EEG activity was recorded continuously through a BioSemi Active II system (64-electrode cap with 10–20 arrangement). EEG preprocessing was done in MATLAB 2021.0 using the EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes.

The recorded signal was referenced online to the common mode sense (CMS) and driven right leg (DRL) electrodes. The horizontal electrooculogram (HEOG) was measured via a pair of external electrodes, one placed on the outer corner of each eye. The vertical electrooculogram (VEOG) was measured as the difference between activity of a pair of external electrodes placed beneath both eyes (i.e., upper middle cheeks) and Fp1.

Scalp electrodes also were re-referenced offline during data pre-processing to the average of the left and right mastoids, and all data was baseline corrected to 200 ms before the onset of the memory array. The data were also filtered by using a Butterworth filter (40-Hz low-pass and 0.1-Hz high-pass). Trials with blinks (VEOG activity > ±80 μV) or lateral eye movements (HEOG activity > ±32 μV) were removed. Artifact rejection for the HEOG was performed for individual trials on epochs starting from 200 ms before stimulus onset until the onset of the 2AFC (i.e., −200 to 1700 ms). Figure S2 (see *Supplemental Materials*) depicts group level residual HEOG traces. For the VEOG, artifact rejection was performed for the interval of 200 ms before stimulus onset until the offset of the memory array (i.e., −200 to 1000 ms). Participants with greater than an average of 35% of trials rejected were excluded from the final sample before statistical analyses. Across participants included in the final sample, an average of 14.37% (range 1.0–34.4%) of trials were rejected. Noisy channels to be included in the grand average waveform of interest (described below) also were interpolated where necessary (two participants, electrode PO4 only).

For each participant, the mean CDA amplitudes for each stimulus condition were calculated as the average activity at three posterior electrode pairs (PO3/PO4, PO7/PO8, P7/P8). This combination of posterior and occipital channels was selected based on use in similar work examining the CDA (e.g., with "Mooney" stimuli, see Asp et al., 2021; with real-world objects and colours, see Brady et al., 2016; with colours stimuli, see Salahub et al., 2019). The CDA waveforms were computed as the difference between the activity measured contralateral and ipsilateral to the to-be-remembered side of the display. The main measurement window of interest for the CDA began 300 ms after the offset of the memory array and was 400 ms in duration (i.e., until the onset of the 2AFC probe). Thus, the CDA was measured during the interval of 1300–1700 ms. Additionally, the CDA also was examined during the last 400 ms of the encoding phase, before the onset of the delay period (i.e., 600–1000 ms). This analysis was selected based on not only observed differences in the data, but also given encoding-related differences have been noted in previous works (Asp et al., 2021). Group level nonsubtracted contralateral and ipsilateral waveforms for each of the three stimulus conditions are shown in Fig. S3 (see *Supplemental Materials*). Mean amplitudes for each condition were statistically compared using RM ANOVA and paired *t*-tests.

Results

Behavioural results

Mauchly's test indicated that sphericity was violated, $\chi^2(2) = 8.48$, $p = .014$. Degrees of freedom were corrected by using the Greenhouse-Geisser estimate of sphericity, $\varepsilon = 0.743$.

Within subjects ($N = 22$), there was a significant effect of stimulus type on estimated working memory performance (d'), $F(1.49, 31.21) = 12.11$, $p < .001$, $\eta^2 = 0.366$. It was found that performance for the real-world objects condition ($M = 1.580$, $SD = 0.242$) was significantly greater than that of the scrambled objects condition ($M = 1.284$, $SD = 0.268$), $t(21) = 7.021$, 95% CI [0.208, 0.383], $p < .001$, $d = 1.497$. Performance for real-world objects was also significantly greater than for colours ($M = 1.355$, $SD = 0.376$), $t(21) = 2.969$, 95% CI [0.067, 0.382], $p = .007$, $d = 0.633$. However, no significant difference in memory performance was observed between the scrambled objects and colours conditions, $t(21) = -1.082$, 95% CI [−0.207, 0.065], $p = .292$, $d = -0.231$. Thus, behavioural results support both of our hypotheses, in that the memory benefit for real-world objects compared with scrambled objects and colours was demonstrated, and memory performance was comparable for both scrambled objects and colours (Fig. 2).

ERP results

Early window: 600–1000 ms

During encoding (i.e., 600–1000 ms), there were significant differences in the amplitude of the CDA across the various stimulus type conditions, $F(2, 42) = 6.73$, $p = .003$, $\eta^2 =$

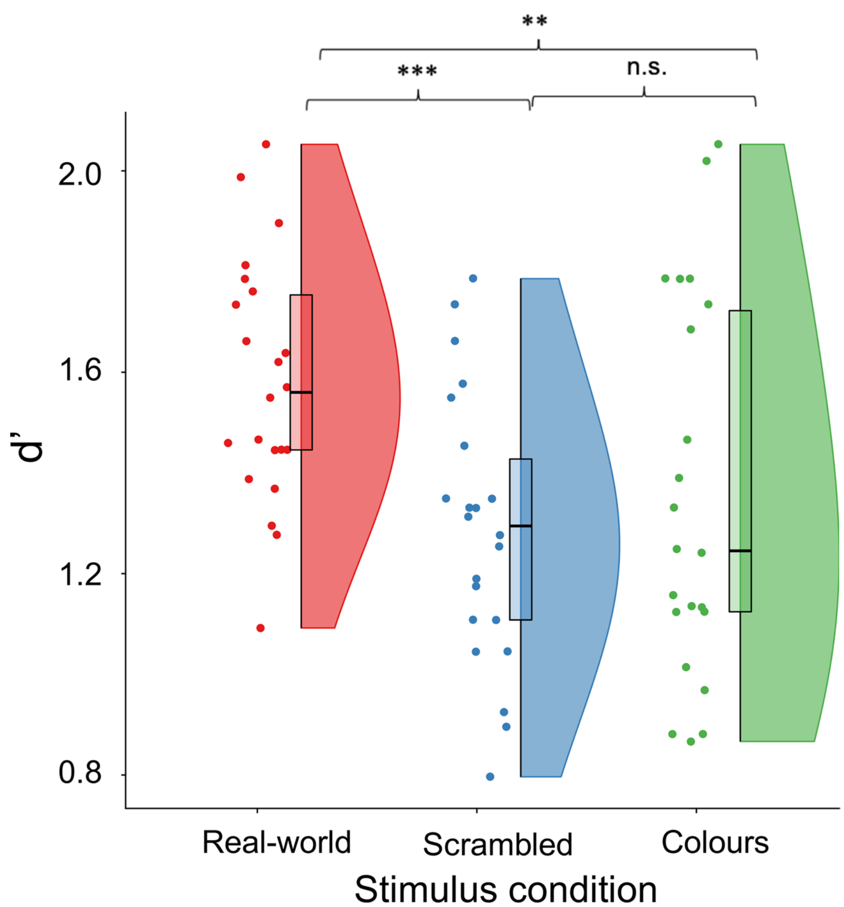

Fig. 2 Behavioural performance data. Working memory performance (mean *d'*) for the real-world objects (red), scrambled objects (blue) and colours (green) conditions. Data points reflect the performance of individual participants (N = 22). Each box represents the interquartile range, and the horizontal line within each boxplot reflects the median. $^{**}p < .01$; $^{***}p < .001$.

0.243 (Fig. 3). Pairwise comparisons[4] revealed that the CDA amplitude of the real-world objects condition ($M = -0.585$ μV, $SD = 0.685$ μV) was significantly greater than that of the scrambled objects condition ($M = -0.059$ μV, $SD = 0.676$ μV), $t(21) = -3.767$, 95% CI [−0.817, −0.236], $p = .001$, $d = -0.803$. The CDA amplitude of the colours condition ($M = -0.540$ μV, $SD = 0.720$ μV) also was significantly greater than that of the scrambled objects condition, $t(21) = -2.867$, 95% CI [−0.831, −0.132], $p = .009$, $d = -0.611$. Interestingly, the CDA amplitude of the colours condition did not differ significantly from the amplitude of the real-world objects condition during encoding, $t(21) = 0.267$, 95% CI [−0.305, 0.394], $p = .792$, $d = 0.057$. A one-sample t-test also revealed that amplitude of the scrambled objects condition during the early window did not differ significantly from zero, $t(21) = -0.406$, $p = .689$.

Late window: 1300–1700 ms

A significant effect of stimulus type on amplitude of the CDA also was observed during the working memory delay (i.e., 1300–1700 ms) as the main interval of interest, $F(2, 42) = 5.40$, $p = .008$, $\eta^2 = 0.205$ (Fig. 3). The CDA amplitude of the real-world objects condition ($M = -0.851$ μV, $SD = 0.766$ μV) was found to be significantly greater than that of

[4] Shapiro-Wilk tests revealed a deviation from normality in the scrambled objects comparisons during encoding (i.e., scrambled-colours: $W = 0.825$, $p = .001$, scrambled-real-world: $W = 0.911$, $p = .050$). However, the same overall effect across stimulus types was observed when conducting paired t-tests (detailed in text body) and Wilcoxon signed-rank tests for these comparisons (i.e., scrambled-colours: $z = 2.678$, $p = .006$, scrambled-real-world: $z = 3.295$, $p < .001$), in that the CDA amplitude for both the colours and real-world objects was significantly greater than the scrambled objects condition. Thus, results of the paired t-tests were reported as they were for the remainder of the comparisons in both the early window (i.e., encoding) and late window (i.e., WM maintenance) analyses. Given normality assessments in smaller samples need be approached with cautionary interpretation, both parametric and nonparametric assessments of normality (as described above) were considered as good practice.

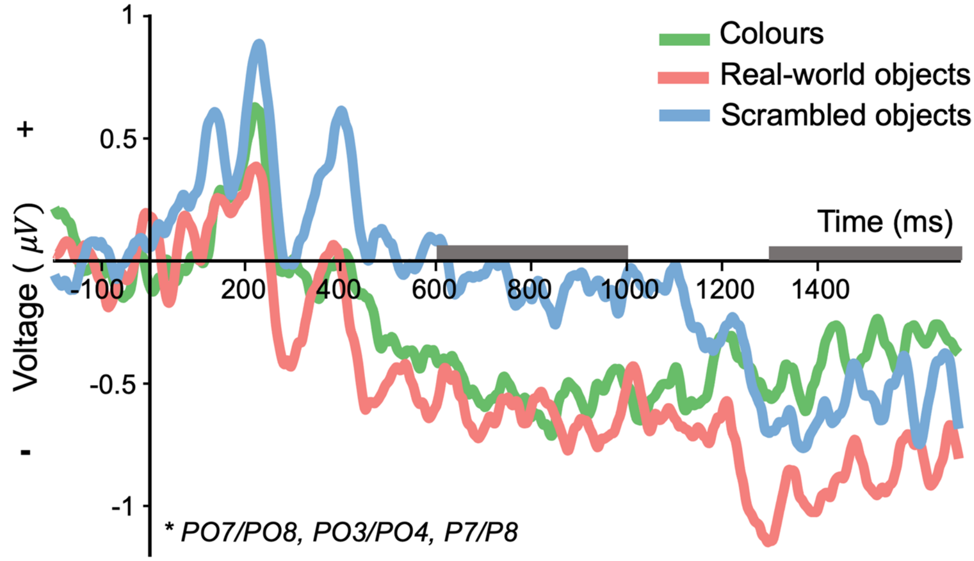

Fig. 3 Group level CDA waveform. Grand average contralateral-ipsilateral difference waveform ($N = 22$) at the average of three channel pairs (PO7/PO8, PO3/PO4, and P7/P8) for the real-world objects (red), scrambled objects (blue), and colours (green) conditions. Grey bars along the horizontal axis indicate the two time intervals examined: 600–1000 ms (CDA during encoding) and 1300–1700 ms (CDA during working memory maintenance). The horizontal axis reflects time (in milliseconds), and the vertical axis reflects voltage (in μV), with negative voltage values plotted downwards. Waveforms are filtered at 30 Hz for visualization purposes

both the scrambled objects condition ($M = -0.565$ μV, $SD = 0.617$ μV; $t(21) = -2.259$, 95% CI [−0.551, −0.023], $p = .035$, $d = -0.482$) and the colours condition ($M = -0.399$ μV, $SD = 0.588$ μV; $t(21) = -3.060$, 95% CI [−0.760, −0.145], $p = .006$, $d = -0.652$). As predicted, there was also no significant difference between the CDA amplitudes of the scrambled objects and colours conditions, $t(21) = -1.167$, 95% CI [−0.461, 0.130], $p = .256$, $d = -0.249$. The CDA amplitudes for all three stimulus type conditions (i.e., colours, real-world objects, scrambled objects) were significantly larger than zero (determined via one-sample t-tests, p's < .01). Thus, although ERP results in the late window (i.e., WM delay) converge with behavioural findings and support both predictions, a difference in the relative CDA amplitudes between the colours, real-world objects, and scrambled objects conditions was observed when comparing activity during encoding versus during working memory maintenance. Namely, whereas the CDA of the colours condition during the encoding period was comparable with the increased amplitude of the real-world objects condition, during WM maintenance, the CDA for real-world objects was comparatively greater than both colours and scrambled objects, which appeared similar in magnitude.

Comparing the CDA during the early and late windows

To further examine differences in the CDA observed in the early and late window intervals, changes in relative amplitude across time were also examined within each stimulus condition. Unlike the colours condition in which the CDA amplitude during both encoding and working memory maintenance remained consistent (e.g., early vs. late window colours CDA, $t(21) = -1.224$, 95% CI [−0.381, 0.099], $p = .234$, $d = -0.261$), marked increases in the CDA amplitudes of both the real-world objects (e.g., early vs. late window real-world objects CDA, $t(21) = 2.537$, 95% CI [0.048, 0.485], $p = .019$, $d = 0.541$) and scrambled objects conditions (early vs. late window scrambled objects CDA: $t(21) = 4.347$, 95% CI [0.264, 0.748], $p < .001$, $d = 0.927$) were noted when comparing activity in the early and late windows. Thus, differences emerging in the late window (relative to the early window) between stimulus conditions seem to be driven by fluctuations in the CDA amplitudes of the object conditions, rather than the colours condition.

Discussion

What is the relative contribution of perceptual complexity versus conceptual meaning to the object memory benefit? In other words, which property of real-world objects drives increases in the CDA amplitude and contributes to improvements in working memory performance for these items? To directly probe this question, the current study aimed to examine the behavioural and electrophysiological effects of a manipulation of the conceptual meaning of real-world objects. Namely, by employing a diffeomorphing scrambling method (Stojanoski & Cusack, 2014) to render the set of real-world objects unrecognizable, we directly assessed how the presence or absence of meaningfulness—despite a shared degree of visual complexity—differentially impacted performance.

As predicted, a memory benefit for intact real-world objects emerged, such that participants' memory was significantly better for real-world objects (i.e., higher *d'* estimates) than scrambled objects or colours, with the amplitude of the CDA waveform during WM retention being greatest for the real-world objects condition. Importantly, statistically comparable behavioural and ERP results (during WM retention, i.e., 1300–1700 ms) between scrambled objects and colours, as well as significant differences between intact real-world objects and scrambled objects variants, collectively rule out the factor of perceptual complexity as driving the memory benefit. In other words, given that performance for comparatively high and low complexity items were similar (i.e., scrambled objects and colours, respectively) with differences emerging between the object sets despite a match for featural complexity, findings suggest that it is likely not the number of visual details a real-world object has that permits improved memory for said object. Instead, measurable enhancements in behavioural performance estimates and CDA amplitude were only observed for stimuli that were recognizable, consistent with literature demonstrating that meaningfulness underlies the VWM advantage for real-world objects (Asp et al., 2021; Brady & Störmer, 2022, 2023).

Although semantics rather than object complexity appears to be driving the object benefit in the current experiment, the mechanism remains unclear. One possibility is that recognizable objects are more "nameable," thereby lending themselves to a verbal rehearsal strategy. That is, unlike scrambled objects for which participants had no prior knowledge or expectations whatsoever, intact real-world objects, and to a lesser extent colour stimuli, may have been better suited to a strategy of verbal labelling, wherein previous knowledge of category-relevant information may have facilitated encoding. Indeed, using continuous report paradigms, it has been shown that colour estimates tend to bias prototypes of colour categories rather than more ambiguous estimates near colour boundaries (Bae et al., 2015). The possibility that the object benefit can be attributed to verbal encoding appears unlikely, however, as previous studies have established that concurrent articulatory suppression has no effect on the improved working memory performance for meaningful stimuli (Brady & Störmer, 2022; Chung et al., 2023). Moreover, there is no a priori reason to believe that the use of a *verbal* encoding strategy would straightforwardly yield a larger *visual* CDA. In fact, some multisensory WM work has shown that the visual CDA is seemingly sensitive to modality-specific (i.e., visual) effects. For example, using a tactile-visual WM dual-task, Katus & Eimer (2018) found that while the visual CDA was sensitive to alterations in visual components of the task (e.g., VWM load), the waveform remained unaffected by manipulations to nonvisual (i.e., tactile) components of the task.

Whereas the CDA effects during active WM maintenance resembled behavioural findings of the object memory benefit, a separate pattern of results emerged during the encoding phase when items remained in view. Namely, the amplitudes of both the colours and real-world objects conditions were comparable and significantly greater in magnitude than the amplitude of the scrambled objects condition (which did not differ from zero). Thus, although meaningful stimuli (i.e., real-world objects) had no marked advantage during encoding relative to simple features (i.e., colours), complex yet meaningless items were disadvantaged by comparison, given the absence of a significant CDA estimate for this condition.

What might have contributed to the observed CDA effects during encoding? One possible explanation is that difficulty establishing a reliable encoding strategy for the scrambled stimuli reduced the efficiency with which items were perceptually represented and organized. Anecdotally, this aligns with comments provided by several participants during debriefing, who noted that the scrambled items were perceived as being the most difficult to maintain and required the most flexible encoding strategy on a trial-by-trial basis. Thus, it would be reasonable to infer that it took participants a comparatively greater length of time to encode the scrambled items given the nature of the stimuli, possibly yielding the comparatively longer CDA latency and reduced CDA amplitude observed in the early window for this condition relative to both the colours and real-world objects. Moreover, the CDA has also been shown to be sensitive to participants' confidence in their response, irrespective of accuracy of the response (i.e., correct vs. incorrect), with lower self-reported confidence corresponding with reductions in CDA amplitude as early as 400-ms poststimulus onset (Mayer et al., 2020). Applied to the current work, if participants were unable to encode the scrambled objects with sufficient precision—either due to added complexity and/or the absence of supplementary semantic information—their confidence in the fidelity and accuracy of these representations may have been reduced, thereby contributing to a diminished CDA amplitude compared with the other stimulus conditions. Familiarity with to-be-remembered stimuli has also been shown to affect both CDA amplitude and CDA latency. Using a change detection task with Pokémon stimuli, Xie & Zhang (2018) demonstrated that a greater CDA amplitude and shorter latency in the early window was observed for participants that were highly familiar with the first-generation stimuli than for participants more unfamiliar with the items. As it relates to the current study, it is thus possible that factors influencing encoding efficiency during the early window, including but not limited to the effects of prior experience on processing speed, may have contributed to the observed differences between stimulus conditions. For example, participants' lack of existing expectations or exposure—and thus very low familiarity—with scrambled objects may have resulted in much slower consolidation speeds compared with colours or real-world objects, thereby affecting the early window CDA (VWM consolidation is covered in more detail later in the discussion).

Although the underlying mechanism remains unclear, findings of differences in an early window CDA across conditions suggests that the waveform is sensitive to stimulus variability as early as the initial encoding phase, as opposed to solely serving as an index of the *outcome* of earlier processing stages (i.e., the number of items stored in memory). Consistent with this notion, Asp et al. (2021) observed reliable differences in the early window CDA for the presentation of one versus three objective faces in a three-item memory array, with a larger negativity emerging when more of the targets were objectively faces. This finding stands in contrast with a purely load-based account of the CDA, in that differences in amplitude were attributable to the type of stimuli in the array, independent of the number of to-be-remembered targets. That is, when all three targets were objective faces the CDA amplitude was greater than for trials in which only one objective face and two nonface targets were presented, despite set size being held constant at three items in each case. Thus, not only was the CDA amplitude seemingly sensitive to the presence of *meaningful* items in their study—rather than solely number of items—but these differences emerged in the neural data even before the WM retention interval, as in the current experiment. Additionally, as in the current study, multiple object tracking studies also have found that the CDA is sensitive to manipulations of *in-view* stimuli (rather than only indexing item load in *memory*; Drew & Vogel, 2008), with the perceptual CDA being larger in magnitude and more sustained in its duration compared with the CDA elicited during WM (Drew et al., 2011). Finally, evidence of alterations in CDA amplitude in response to the continuous and flexible allocation of memory resources has been observed, with larger negativities emerging alongside increases in item priority, despite a constant set size (e.g., 0%, 25%, 75%, or 100% likelihood that the colour of a particular shape would be probed at test; Salahub et al. 2019). In this way, findings of the current work contribute to a growing body of evidence demonstrating that the CDA may also reflect a variety of attentional and perceptual processes supplementing VWM resources (see Emrich et al., 2022 for review), rather than serving as solely a measure of a discrete VWM capacity.

Whereas the real-world objects and colours conditions had comparable CDA amplitudes during encoding, a difference emerged between the two conditions during the WM maintenance interval, such that the amplitude of the real-world objects set was significantly greater. Importantly, this difference was not attributable to a decrease in amplitude for colours (e.g., early vs. late window colours CDA) but rather by an increase in amplitude for the real-world objects (e.g., early vs. late window real-world objects CDA). This finding differs from previous findings using fMRI, where increased activity was observed in ventral visual areas for the representation of recognizable objects, but not their scrambled, unrecognizable counterparts, but only during the encoding window (Stojanoski et al., 2019). Interestingly, another study did not observe the recruitment of any additional brain areas in order to encode and maintain complex objects compared with their scrambled counterparts (Veldsman et al., 2017); the authors attributed the improved performance to greater variability in the pattern of responses (Garrett et al., 2011). Thus, the CDA and fMRI correlates of VWM maintenance appear to differ in their sensitivity to different stimulus types. Differences in the temporal resolution of the measures themselves also may be a contributing factor to cross-study discrepancies. For example, whereas CDA differences during the maintenance interval in the current work were compared within several hundred milliseconds of stimulus offset, maintenance intervals in previous fMRI works were instead only examined on the order of seconds (Stojanoski et al., 2019: maintenance period between 1–11-s long; Veldsman et al., 2017: maintenance period between 1–9-s long). Given differences across stimulus type were observed even within a 400-ms measurement window in the current study, it may be the case that probing the effects of meaningfulness requires the use of more fine-grained measures of temporal dynamics, particularly if effects are most pronounced early on in the retention interval.

Additionally, mixed results across experiments may also be attributable in part to the methodological (i.e., task-related) differences between the current work and some previous fMRI studies. For example, in the current study, continued maintenance of the semantic information encoded for intact real-world objects would have been advantageous, given the nature of the foils during the 2AFC at test. That is, the semantic information encoded for the intact objects may have provided a "prolonged" boost to performance, by enabling participants to distinguish the target more readily from a *different* within-category distractor. Conversely, in the Stojanoski et al. (2019) study, given participants were presented with only a single target during encoding, and the test phase consisted of variations of *only* the same target item, maintenance of semantic information would not have improved performance (in that there was no need to distinguish multiple item identities from one another). In fact, the continuous response format at test was designed to have participants match one of several versions of the target with that which they originally encoded—a distinction that would have been entirely perceptually driven, given each distractor item was identical in its semantic content relative to the target. As noted by Stojanoski et al. (2019), it therefore may be possible that because the task design did not require participants to use semantic content at test this activity was absent from the delay interval, despite the fact that this information still facilitated encoding for the recognizable images. Similar methodological differences exist between the current study and the fMRI study of Veldsman et al. (2017). For example, in their third experiment, although two distinct to-be-remembered items were presented during encoding, variations of only one item were presented at test (i.e., all distractor foils shared the target identity, only with varying warp levels), meaning that maintaining semantic details of the object would not assist participants in more accurately selecting the target. To better understand at the neural level the role of meaningfulness in improved WM performance for recognizable objects—a shared finding in works of Stojanoski et al. (2019), Veldsman et al. (2017), and in the current study—further research is needed to elucidate the temporal dynamics of how the brain encodes and maintains different stimuli.

Alongside consideration of the effects across stimulus type during the early versus late intervals independently, an additional (not incompatible) explanation is that there exists a difference in the VWM consolidation process. That is, how individual items from each stimulus type may be differentially transferred from a transient percept to a more durable VWM trace. For example, previous research suggests that visual LTM affects early VWM processing, with increased stimulus familiarity influencing the rate at which items may be encoded. In one study, Blalock (2015) used a backwards masking task to test participants' memory for a set of abstract polygons, some of which were learned via training tasks and others that were novel, untrained shapes. Not only was a performance advantage during a recognition memory test observed for the explicitly trained stimuli, but they were also consolidated more rapidly into VWM than the novel untrained shapes, as evidenced by the absence of an effect of the visual masks even at the earlier SOAs (stimulus-onset asynchronies). Similarly, Xie & Zhang (2017b) sought to examine the effects of prior LTM—based on participants' experience with the stimuli, rather than laboratory-based training for novel items—on performance in a Pokémon change detection task. The more familiar participants were with the first-generation Pokémon relative to the recent-generation stimuli (i.e., high-familiarity group), the faster the VWM consolidation and higher the capacity for first-generation characters compared with the recent-generation characters (for shorter stimulus onset asynchronies [SOAs]). This effect of previous LTM on consolidation speed was not observed in their low-familiarity group counterparts. Effects of stimulus familiarity were also observed by Ngiam et al. (2019). Researchers used varying SOA intervals between a memory array and visual mask in a change-detection task to estimate capacity for letter stimuli as a function of encoding time. Although an effect of prolonged encoding time was observed (i.e., gradual improvement in performance with increased SOA/greater delay of the mask), importantly, this effect did not differ as a function of parametric complexity of the fonts (i.e., number of features). As well, increasing complexity of the fonts failed to yield reductions in capacity (K) estimates. Instead, encoding rate and VWM storage capacity were only found to be affected by familiarity with the stimuli, such that more familiar fonts and alphabets were consolidated more quickly and associated with higher K estimates (at both shorter and longer presentation times/SOAs).

As it relates to results of the current work, it is therefore possible that differences between the various stimulus type conditions—both in terms of behavioural performance and CDA effects—may be attributable in part to differences in the VWM consolidation process, as opposed to solely differences in the VWM representation itself. In other words, factors that may affect *how well* and the *rate at which* information was transferred into memory in addition to considering only *what* was transferred into memory (i.e., real-world object vs. scrambled object vs. colour). For example, as noted in previous works (Ngiam et al., 2019; Xie & Zhang, 2017b), it may be the case that previous knowledge or meaning associated with real-world objects enables participants to selectively attend to or recognize only critical features of the stimuli during encoding. That is, the availability of semantic information and knowledge acts as a sort of encoding "boost" for real-world objects, relative to that which is available for colours or scrambled objects that have a lacking or absent, respectively, degree of conceptual meaning. This additional dimension of information may not only work to expedite consolidation of the real-world objects—given the amount of featural information encoded into VWM at a given time may be reduced—but also facilitate object recognition at test, by enabling the consolidated representations to be maintained in a manner that more easily distinguishes them from one another. Thus, it may be the case that for real-world objects, items were transferred from percepts to more stable representations more quickly (i.e., higher early window CDA amplitude), allowing for comparatively more information to be stored in VWM (i.e., higher late-window CDA amplitude and d' for real-world objects).

Applied to the colours condition, one could speculate that given each item differed by only a single feature, the amount of information available to be consolidated into VWM was also low (relative to say complex, scrambled objects). However, not unlike scrambled objects, colours themselves lack meaning or inherent semantic associations (relative to the real-world objects), meaning the representations at the maintenance phase may not have been as stable or as distinguished relative to one another. In this way, while the items may have been able to be encoded somewhat readily (i.e., comparable early window CDA amplitude with real-world objects), a comparatively less robust representation may have been established based on the consolidated information and/or maintenance of the encoded information may have been less successful (i.e., only separable by encoded perceptual information), ultimately impacting VWM storage (i.e., lower late window CDA amplitude and *d'* for colours than real-world objects). Maintenance of scrambled object representations may have been impacted similarly, as another condition low in conceptual meaning (as reflected by comparable *d'* and late window CDA amplitude as colours). As it relates to the reduced early window scrambled condition CDA (relative to the comparable amplitudes of the colours and real-world objects), it may be the case that high complexity in the absence of meaning significantly slowed consolidation. That is, there was not only more information available to be encoded into VWM (e.g., as opposed to colours, for which there was only one available feature), but the ability to differentiate items at test also necessitated memory for more of this information (as opposed to intact real-world objects, where reliance on a semantic "boost" to distinguish items in the 2AFC may have reduced the need to consolidate as many item-specific visual features). As such, information in the scrambled objects condition may have been encoded or consolidated into VWM less efficiently relative to information in the colours or intact objects conditions (i.e., comparatively low scrambled early window CDA amplitude), thereby affecting the stability of the corresponding traces being maintained in memory and reported on at test (e.g., late window scrambled CDA amplitude and *d'* comparable to colours). To elucidate what differences in behavioural and CDA signatures of WM across conditions may tell us about the *process* of VWM itself—as opposed to the quality or quantity of representations—future studies should probe further the relationship between *what* is encoded (i.e., the stimuli themselves) and *how well* or the *rate at which* this information is consolidated.

One critical methodological manipulation of the current work was the use of maximally dissimilar comparison foils at the test phase for all conditions. This manipulation permitted the fair comparison of memory performance across stimulus types, given previous research suggests that arbitrarily selecting complex objects at test (i.e., for the object trials) may artificially diminish working memory performance for these conditions (Brady & Störmer, 2023). Whereas maximal dissimilarity at test was controlled for, other stimulus intrinsic properties that have been shown to be associated with improved memory performance were not explicitly manipulated. For instance, recent evidence suggests that the memorability of a stimulus relative to other items in the encoding array is predictive of said item's performance at test (Torres et al., 2023; Gillies et al., 2023). As such, there may be properties beyond solely the presence or absence of conceptual information in absolute terms—as was assessed in the current study—that contribute to WM performance for real-world objects, such as the effect of memorability on improvements in maintenance efficiency or resistance to interference (Gillies et al., 2023). Future iterations of this work would benefit from the consideration of object memorability among other stimulus-intrinsic properties when considering the nature of VWM performance for real-world objects.

One factor that cannot be ruled out based on findings of the current study is the possibility that the conditions differ in how similar the stimuli are to each other. For example, are scrambled objects more or less similar to one another—in terms of the amount of overlapping perceptual information—than intact objects are with each other? Such effects of similarity within a stimulus category have been shown to affect both memory encoding and retrieval in previous works. For example, using a set of meaningless, nonverbalizable shapes in a delayed discrimination task, Jackson et al. (2015) found that differences in memory of simple and complex polygons only arose when sample and test items were similar but not when the items were dissimilar. Likewise, Awh, Barton, & Vogel (2007) noted that while change detection accuracy for cross-category changes (e.g., shaded cube to Chinese character) was comparable with detection accuracy for changes in colour stimuli, within-category changes in which the memory and test items shared a high degree of similarity (i.e., one shaded cube to another) resulted in poorer performance.

As it relates to the current manipulation, despite controls for similarity at probe (i.e., the use of maximally dissimilar foils), it is still possible that differences in relative within-category similarity between the various stimulus types contributed to performance differences. For example, inter-item similarity for the intact objects was more variable relative to the colours because of the presence of differences along multiple featural dimensions in the case of the former (e.g., variability in shape, multicoloured, etc.). Additionally, whereas diffeomorphic transformation yielded a comparable degree of perceptual complexity for both object sets, scrambled objects may have had higher featural and/or spatial similarity relative to intact objects (i.e., given the warping method applied a similar transformation to each object). However, given perceptual characteristics are intimately tied to semantics (i.e., the meaning of an object cannot be accessed if the object itself cannot first be identified), fully disentangling visual similarity from semantics would prove difficult. In fact, making the scrambled items less similar to one another would require them to be less warped, thereby negating our manipulation of testing a mismatch in conceptual meaning of the object sets. Importantly, however, some previous works that have explicitly examined stimulus similarity still seem to suggest a role for semantics or previous knowledge in improved performance. For example, Ngiam et al. (2019) noted that despite a match for both complexity *and* similarity between English and BACS fonts (i.e., characters designed to match the features of English letters), VWM capacity and encoding rate were both higher for the more familiar English font. Critically, for the current results, while item similarity has been shown to affect behavioural VWM performance, stimulus homogeneity has no demonstrable effect on CDA amplitudes (Cai et al., 2022). Thus, while our results cannot fully be accounted for by stimulus similarity, it remains possible that stimulus complexity and similarity may still be correlated to some degree in the current design. As such, consideration of within-category stimulus similarity in the context of the object benefit remains a question for future research. As a final note, it is important to acknowledge that alongside stimulus properties, there are other stimulus-specific effects that may contribute to the results observed here—e.g., conditions may differ in how much effort is required to encode the different stimuli (Xie & Zhang, 2023). Thus, while our results provide evidence of the role of meaningfulness in supporting VWM performance, further research is required to pinpoint the precise mechanism.

In sum, the memory advantage for meaningful objects noted in previous works (Brady et al., 2016; Asp et al., 2021; Brady & Störmer, 2022) was replicated in the current study, in that participants on average had significantly better memory performance (i.e., greater *d'*) for real-world objects than simple features (i.e., colours) or complex but meaningless items (i.e., scrambled objects). Additionally, greater negativity in the amplitude of the CDA during WM maintenance for real-world objects coincides with the notion that supplementary cognitive processes (i.e., semantic processing) may support the formation and active maintenance of these representations. From a theoretical perspective, the finding of significant differences in performance across stimulus types also has implications for our understanding of the nature of VWM, particularly in terms of the importance of conceptual meaning to memory performance. Whereas scrambled objects and colours were remembered similarly, participants' memory of meaningful (i.e., real-world) objects was improved by comparison, thereby demonstrating that storage limits are flexible—not fixed—and differ depending on the type of information being maintained. Results also speak to the importance of considering the methodological approach used to estimate VWM, in that memory performance may be reduced under conditions in which connections to existing knowledge and meaning are unavailable (e.g., testing VWM using only simple feature-based stimuli).

Supplementary information The online version contains supplementary material available at https://doi.org/10.3758/s13415-024-01158-z.

Acknowledgements We would like to thank Keynen Lynett, Chae Bush, and Christopher Keightley for their assistance with data collection.

Authors' contributions S.M.E. and B.S. for conceptualization. A.M.L.T. programmed the task script (with edits from S.M.E.). A.M.L.T. for data collection and data analysis. A.M.L.T. wrote the manuscript. S.M.E. and B.S. provided revisions and feedback on initial manuscript draft. A.M.L.T., S.M.E., and B.S. discussed the interpretation of results included in the final manuscript.

Funding This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC 2019-04865 (grant to S.M.E.)), and by an Ontario Graduate Scholarship awarded to A.M.L.T.

Data availability Data and materials for the study are available online at https://osf.io/sfkuz/.

Code availability The PsychoPy task script and MATLAB code used to pre-process EEG data are available online at https://osf.io/sfkuz/.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethics approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Social Science Research Ethics Board at Brock University (REB # 19-040-EMRICH).

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent to publish Participants signed informed consent regarding publishing of their anonymized data.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Asp, I. E., Störmer, V. S., & Brady, T. F. (2021). Greater visual working memory capacity for visually matched stimuli when they are perceived as meaningful. *Journal of Cognitive Neuroscience, 33*(5), 902–918. https://doi.org/10.1162/jocn_a_01693

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Bae, G. Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology, 144*(4), 744–763. https://doi.org/10.1037/xge0000076

Barton, B., Ester, E. F., & Awh, E. (2009). Discrete resource allocation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1359–1367. https://doi.org/10.1037/a0015792

Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *The Journal of Neuroscience, 34*(10), 3632–3645. https://doi.org/10.1523/JNEUROSCI.3204-13.2014

Blalock, L. D. (2015). Stimulus familiarity improves consolidation of visual working memory representations. *Attention, Perception, & Psychophysics, 77*, 1143–1158. https://doi.org/10.3758/s13414-014-0823-z

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Brady, T. F., & Störmer, V. S. (2023). Comparing memory capacity across stimuli requires maximally dissimilar foils: Using deep convolutional neural networks to understand visual working memory capacity for real-world objects. *Memory & Cognition*. https://doi.org/10.3758/s13421-023-01485-5

Brady, T. F., & Störmer, V. S. (2022). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*, 942–958. https://doi.org/10.1037/xlm0001014

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences, 113*(27), 7459–7464. https://doi.org/10.1073/pnas.1520027113

Cai, Y., Fulvio, J. M., Samaha, J., & Postle, B. R. (2022). Context binding in visual working memory is reflected in bilateral event-related potentials, but not in contralateral delay activity. *Eneuro, 9*(6), ENEURO.0207-22.2022. https://doi.org/10.1523/ENEURO.0207-22.2022

Chung, Y. H., Brady, T. F., & Störmer, V. S. (2023). No fixed limit for storing simple visual features: Realistic objects provide an efficient scaffold for holding features in mind. *Psychological Science*. https://doi.org/10.1177/09567976231171339

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2011). Delineating the neural signatures of tracking spatial position and working memory during attentive tracking. *The Journal of Neuroscience, 31*(2), 659–668. https://doi.org/10.1523/JNEUROSCI.1339-10.2011

Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *The Journal of Neuroscience, 28*(16), 4183–4191. https://doi.org/10.1523/JNEUROSCI.0556-08.2008

Emrich, S. M., Salahub, C., & Katus, T. (2022). Sensory delay activity: More than an electrophysiological index of working memory load. *Journal of Cognitive Neuroscience, 35*(1), 135–148. https://doi.org/10.1162/jocn_a_01922

Garrett, D. D., Kovacevic, N., McIntosh, A. R., & Grady, C. L. (2011). The importance of being variable. *The Journal of Neuroscience, 31*(12), 4496–4503. https://doi.org/10.1523/JNEUROSCI.5641-10.2011

Gillies, G., Park, H. G., Woo, J., Walther, D. B., Cant, J. S., & Fukuda, K. (2023). Tracing the emergence of the memorability benefit. *Cognition, 238*. https://doi.org/10.1016/j.cognition.2023.105489

Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 556–568. https://doi.org/10.1037/0096-1523.34.3.556

Jackson, M. C., Linden, D. E., Roberts, M. V., Kriegeskorte, N., & Haenschel, C. (2015). Similarity, not complexity, determines visual working memory performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(6), 1884–1892. https://doi.org/10.1037/xlm0000125

Katus, T., & Eimer, M. (2018). Independent attention mechanisms control the activation of tactile and visual working memory representations. *Journal of Cognitive Neuroscience, 30*(5), 644–655. https://doi.org/10.1162/jocn_a_01239

Li, X., Xiong, Z., Theeuwes, J., & Wang, B. (2020). Visual memory benefits from prolonged encoding time regardless of stimulus type. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(10), 1998–2005. https://doi.org/10.1037/xlm0000847

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*, 213. https://doi.org/10.3389/fnhum.2014.00213

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281. https://doi.org/10.1038/36846

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences, 17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Mayer, J. S., Korinth, S., Peters, B., & Fiebach, C. J. (2020). An electrophysiological dissociation of encoding vs. maintenance failures in visual-spatial working memory. *Frontiers in Psychology, 11*, 522. https://doi.org/10.3389/fpsyg.2020.00522

McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex, 43*(1), 77–94. https://doi.org/10.1016/S0010-9452(08)70447-7

Ngiam, W. X. Q., Khaw, K. L. C., Holcombe, A. O., & Goodbourn, P. T. (2019). Visual working memory for letters varies with familiarity but not complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(10), 1761–1775. https://doi.org/10.1037/xlm0000682

Peirce, J. W. (2007). PsychoPy - psychophysics software in python. *Journal of Neuroscience Methods, 162*(1-2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2*(5), 509–522.

Potter, M. C. (2012). Conceptual short term memory in perception and thought. *Frontiers in Psychology, 3*, 113. https://doi.org/10.3389/fpsyg.2012.00113

Quirk, C., Adam, K. C. S., & Vogel, E. K. (2020). No evidence for an object working memory capacity benefit with extended viewing time. *ENEURO*, *7*(5), ENEURO.0150- 20.2020. https://doi.org/10.1523/ENEURO.0150-20.2020

Salahub, C., Lockhart, H. A., Dube, B., Al-Aidroos, N., & Emrich, S. M. (2019). Electrophysiological correlates of the flexible allocation of visual working memory resources. *Scientific Reports, 9*(1), 19428. https://doi.org/10.1038/s41598-019-55948-4

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour, 4*(11), 1156–1172. https://doi.org/10.1038/s41562-020-00938-0

Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of Vision, 14*(12), 6–6. https://doi.org/10.1167/14.12.6

Stojanoski, B., Emrich, S. M., & Cusack, R. (2019). Representation of semantic information in ventral areas during encoding is associated with improved visual short-term memory. *Neuroscience*. https://doi.org/10.1101/2019.12.13.875542

Torres, R. E., Duprey, M., Molokwu, N., Campbell, K. L., & Emrich, S. M. (2023). Not all objects are created equal: Greater visual working memory for real-world objects is related to item memorability. *PsyArXiv*. https://doi.org/10.31234/osf.io/v2ta5

Veldsman, M., Mitchell, D. J., & Cusack, R. (2017). The neural basis of precise visual short-term memory for complex recognisable objects. *NeuroImage, 159*, 131–145. https://doi.org/10.1016/j.neuroimage.2017.07.033

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428*(6984), 748–751. https://doi.org/10.1038/nature02447

Vogel, E. K., McCollugh, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*(7067), 500–503. https://doi.org/10.1038/nature04171

Wickens, T. D. (2001). *Elementary signal detection theory*. New York: Oxford University Press, Oxford.

Xie, W., & Zhang, W. (2017a). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory & Cognition, 45*, 677–689. https://doi.org/10.3758/s13421-016-0679-7

Xie, W., & Zhang, W. (2017b). Familiarity speeds up visual short-term memory consolidation. *Journal of Experimental Psychology: Human Perception and Performance, 43*(6), 1207–1221. https://doi.org/10.1037/xhp0000355

Xie, W., & Zhang, W. (2018). Familiarity speeds up visual short-term memory consolidation: Electrophysiological evidence from contralateral delay activities. *Journal of Cognitive Neuroscience, 30*(1), 1–13. https://doi.org/10.1162/jocn_a_01188

Xie, W., & Zhang, W. (2023). Effortfulness of visual working memory: Gauged by physical exertion. *Journal of Experimental Psychology: General, 152*(7), 2074–52093. https://doi.org/10.1037/xge0001391

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open practices statement The data and materials for the experiment are available online at https://osf.io/sfkuz/ and the experiment was preregistered (https://aspredicted.org/fh73w.pdf).