



Quaddles: A multidimensional 3-D object set with parametrically controlled and customizable features

Marcus R. Watson¹ · Benjamin Voloh² · Milad Naghizadeh¹ · Thilo Womelsdorf^{1,2}

Published online: 7 August 2018
© Psychonomic Society, Inc. 2018

Abstract

Many studies of vision and cognition require novel three-dimensional object sets defined by a parametric feature space. Creating such sets and verifying that they are suitable for a given task, however, can be time-consuming and effortful. Here we present a new set of multidimensional objects, *Quaddles*, designed for studies of feature-based learning and attention, but adaptable for many research purposes. Quaddles have features that are all equally visible from any angle around the vertical axis and can be designed to be equally discriminable along feature dimensions; these objects do not show strong or consistent response biases, with a small number of quantified exceptions. They are available as two-dimensional images, rotating videos, and FBX object files suitable for use with any modern video game engine. We also provide scripts that can be used to generate hundreds of thousands of further Quaddles, as well as examples and tutorials for modifying Quaddles or creating completely new object sets from scratch, with the aim to speed up the development time of future novel-object studies.

Keywords Stimulus set · Feature space · Multidimensional objects

In many experiments in the cognitive sciences, participants must view three-dimensional (3-D) stimuli, or 2-D projections of 3-D stimuli, that they have not encountered before. Such novel object sets have been used in studies of such phenomena as object recognition and discrimination (e.g., Biederman & Gerhardstein, 1993; Bülthoff & Edelman, 1992; Chuang, Vuong, & Bülthoff, 2012; Gauthier, James, Curby, & Tarr, 2003; Harman & Humphrey, 1999; Harman, Humphrey, & Goodale, 1999; Hayward & Tarr, 1997; Richler, Wilmer, & Gauthier, 2017; Tarr, Bülthoff, Zabinski, & Blanz, 1997; Wong & Hayward, 2005), perception and attention to different object properties (Arnott, Cant, Dutton, & Goodale, 2008; Cant & Goodale, 2007), memory for objects (Humphrey & Khan, 1992; Knutson, Hopkins, & Squire, 2012; Mercer & Duffy,

2015), facial perception and recognition (e.g., Gauthier & Tarr, 1997; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Wong, Palmeri, & Gauthier, 2009), category representation (e.g., Wallraven, Bülthoff, Waterkamp, van Dam, & Gaissert, 2014; Williams, 1998), conditioned fear responses (e.g., Barry, Griffith, Vervliet, & Hermans, 2016; Bennett, Vervoort, Boddez, Hermans, & Baeyens, 2015; Scheveneels, Boddez, Bennett, & Hermans, 2017), linguistic demonstratives and gestures (Cooperrider, 2015), and emotional influences on perception (Estes, Jones, & Golonka, 2012). Some sets of novel objects have also been presented in articles specifically written to encourage their adoption by other researchers (Barry, Griffith, De Rossi, & Hermans, 2014; Buffat et al., 2014), or simply hosted online (Harris, 2015) for other researchers to use. Figure 1 shows representative exemplars of these sets.

As the use of novel objects in research has become more commonplace, there has been a parallel rise in studies in which participants have engaged and interacted with complex, continually changing virtual environments. Such dynamic tasks, presented on traditional monitors or stereoscopic displays, enable the presentation of much richer stimuli and the collection of much richer data streams than more traditional, static tasks. Dynamic tasks have been used to investigate the processes underlying phenomena such as the mechanisms of spatial navigation in humans and other animals (Bohil, Alicea, & Biocca,

✉ Marcus R. Watson
watsonmr@yorku.ca

Thilo Womelsdorf
thilo.womelsdorf@vanderbilt.edu

¹ Department of Biology, Centre for Vision Research, York University, Toronto, Ontario, Canada

² Department of Psychology, Vanderbilt University, Nashville, TN, USA

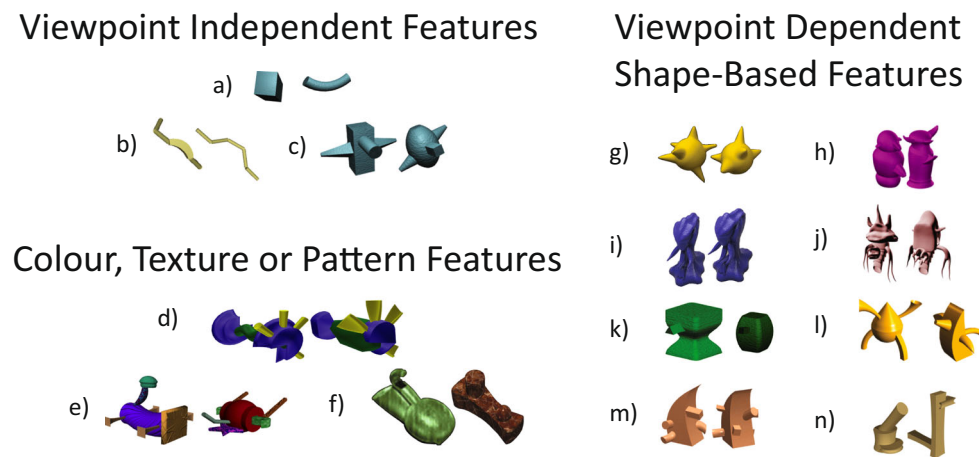


Fig. 1 Representative exemplars from a number of previously reported sets of novel objects. Most do not have features that are viewpoint-independent or that are defined by elements other than shape, both of which qualities we needed for studies with multidimensional view-invariant features. No set has both. (a) geons (Biederman & Gerhardstein, 1993; Hayward & Tarr, 1997), (b) strings (Biederman & Gerhardstein, 1993; Bühlhoff & Edelman, 1992; Tarr et al., 1997), (c) multi-geons (Biederman & Gerhardstein, 1993); (d) Yadgits (Harris, 2015), (e) Fribbles (Barry et al., 2014; Williams, 1998); (f) nonsense objects (Cant & Goodale, 2007; Humphrey & Khan, 1992); (g) amoeboids (Bühlhoff & Edelman, 1992; Wong & Hayward, 2005), (h)

Greebles (Gauthier & Tarr, 1997), (i) YUFOs (Gauthier et al., 2003), (j) Sheinbugs (Richler et al., 2017), (k) two-part objects (Hayward & Tarr, 1997), (l) pair-wise similar objects (Biederman & Gerhardstein, 1993), (m) geons with occlusion (Wong & Hayward, 2005), and (n) Ziggerins (Wong et al., 2009). Exemplar pairs g, m, and n were taken from <http://ww2.psy.cuhk.edu.hk/~mael/Stimuli.html>; a, b, c, d, e, h, i, k, and l were taken from http://wiki.cnbc.cmu.edu/Novel_Objects; j was taken from <http://gauthier.psy.vanderbilt.edu/resources/>; and f was taken from Cant and Goodale (2007). Some stimuli were edited to remove background color in the figure

2011; Ekstrom et al., 2003; Weisberg, Schinazi, Newcombe, Shipley, & Epstein, 2014); multisensory integration in the determination of one's own location (Ehrsson, 2007; Lenggenhager, Tadi, Metzinger, & Blanke, 2007); memory retrieval (Watrous, Tandon, Conner, Pieters, & Ekstrom, 2013); priority in attention, gaze, and memory (Aivar, Hayhoe, Chizk, & Mruczek, 2005; Jovancevic, Sullivan, & Hayhoe, 2006); the temporal organization of gaze in realistic tasks (Johnson, Sullivan, Hayhoe, & Ballard, 2014); subliminal cueing (Aranyi et al., 2014; Barral et al., 2014); and the development of brain–computer interfaces (Leeb et al., 2007).

Both dynamic tasks and novel 3-D object sets, then, have become standard tools in the cognitive science repertoire. We are not aware of any published work that combines the two streams, but we anticipate that this will rapidly become commonplace, as more researchers become aware of the power and flexibility these tools enable, without a corresponding sacrifice in experimental control. Our laboratory has begun running such studies, in which we examine attentional and oculomotor changes as participants learn about a novel object set in a dynamic environment (Watson, Voloh, Naghizadeh, Chen, & Womelsdorf, 2017). With so many sets of novel objects freely available (see Fig. 1), it came as a surprise that we could not find a multidimensional set that met our requirements. Instead, we had to design our own and test their suitability for our task, a much more difficult and time-consuming project than we had originally anticipated, and one that we hope to make substantially easier for future researchers.

In the present article we review this novel object set, named *Quaddles* in reference to the four feature dimensions that define the object space. In addition, we describe (and provide links to) tools that allow the creation of thousands of parametrically varied objects using preexisting features, which enable researchers to design custom features quickly and relatively easily. Finally, we present the results of a feature detection task showing how sensitivity, response bias, and detection efficiency to different feature values of a particular set of Quaddles can be quantified, establishing that, for the most part, the feature values are equally detectable and do not produce strong response biases.

Introducing Quaddles

The experimental task for which Quaddles were designed has participants moving freely using a joystick around a realistic virtual 3-D environment and choosing between objects in that environment. Their object selections are either rewarded or not, on the basis of the particular feature values of the objects, and participants have to learn through trial and error which feature values are associated with reward. Our requirements for these objects were that they have:

- An aesthetically pleasing appearance
- Multiple feature dimensions, including nonshape dimensions
- Multiple feature values along each feature dimension
- Feature values that are roughly equally perceptible and do not show strong response biases

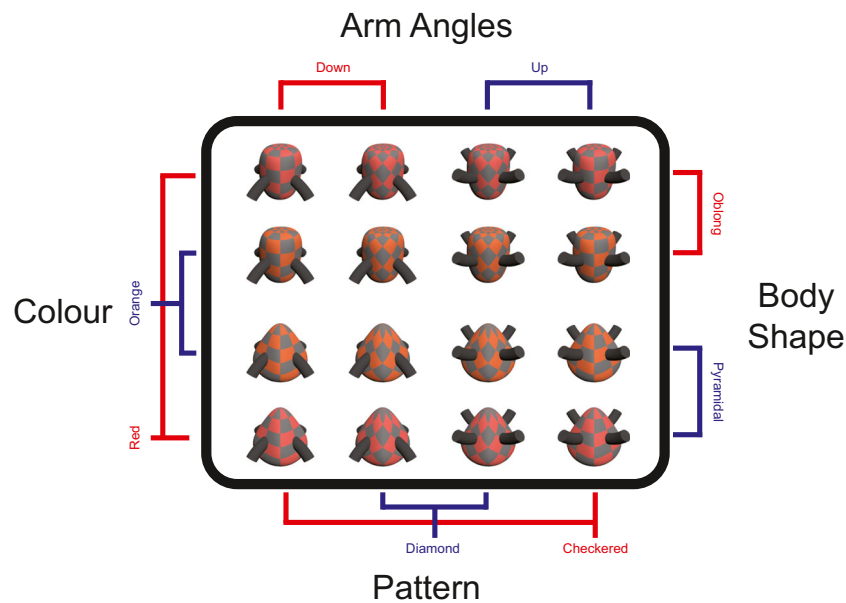


Fig. 2 The primary set of 16 Quaddles used in the feature identification study described in this article

- Vertical symmetry
- Features that can all be clearly and simultaneously viewed from any angle around the vertical axis
- The ability to be exported to any commonly used image format (PNG, GIF, JPG, etc.)
- The ability to be exported to any common video game engine (Unity 3D, Unreal Engine, etc.)

No previous set of novel objects we are aware of meets all these criteria. Very few of these sets have viewpoint-independent features that can all be viewed simultaneously from any angle around the object, and most have purely shape-based feature dimensions (see Fig. 1). Furthermore, we are only aware of one object set for which feature similarity has been quantified (Barry, De Rossi, & Hermans, 2014), and balancing perceptibility across different feature dimensions has only been attempted for much simpler stimuli (Kaldy, Blaser, & Leslie, 2006)—hence, the need for a new set of objects.

The primary set of Quaddles are defined by four feature dimensions: body shape, arm angle, surface color, and surface pattern, each of which has two possible values (Fig. 2). The set can easily be extended by adding new feature values (Figs. 3 and 4), by using intermediate (morphed) values between the existing values (Fig. 3), or by adding or removing

dimensions (Fig. 4). All Quaddles shown in these figures are freely available from our website (<http://accl.psy.vanderbilt.edu/resources/analysis-tools/>) in FBX file format. High-definition pictures of each object from a variety of perspectives are also available in JPG formats with black, white, gray, or green backgrounds, as well as in PNG format with transparent backgrounds. Videos of the objects rotating against black, white, and gray backgrounds are also available in mp4 format. Finally, the same website contains a detailed manual and links to a github script repository containing code that can generate all the objects shown here, as well as tens of thousands of others, and that can easily be modified to create custom object sets.

In the remainder of the article, we summarize the methods for creating Quaddles (more complete details can be found in the manual, hosted at <http://accl.psy.vanderbilt.edu/resources/analysis-tools/>), and present the results of a feature detection study quantifying sensitivity, response bias, and efficiency to the different feature values. In this task, participants were cued with two feature values prior to being shown a single Quaddle that contained only one of these values, and they had to report which of the two values was present. We calculated independent measures of sensitivity (d') and response bias (criterion) for each of the feature values, using standard signal detection measures (Macmillan & Creelman,

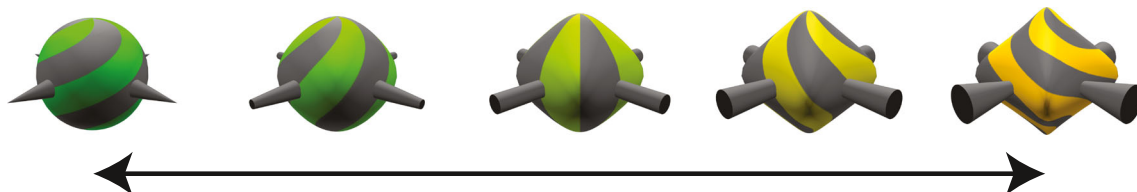


Fig. 3 Examples of morphing, using two different feature values on each of the four feature dimensions



Fig. 4 Additional objects showing some possible variations on the basic Quaddle body plan that can be easily generated via scripting. The top row shows changes in the number of arms, the addition of a “hat” feature that allows two body shapes to be stacked on top of each other, and the

addition of feature values such as the hairy surface texture, whereas the bottom two rows show objects with all the same feature dimensions, but alternative feature values

2005). Sensitivity on this task quantifies the ability to correctly report a feature as being present or absent, whereas response bias quantifies the likelihood of reporting a particular feature overall, either correctly or incorrectly. The desired results were indistinguishable sensitivities to the different feature values and bias scores that would cluster around 0, indicating no consistent biases across participants. We also calculated efficiency scores that combined accuracy and response time, controlling for speed–accuracy trade-offs (Smilek, Enns, Eastwood, & Merikle, 2006; Townsend & Ashby, 1983). Furthermore, we quantified the stability of the efficiency scores using consistency metrics, both across blocks within individual participants and between participants. To prefigure our results, this analysis demonstrated that feature value detection efficiency is fairly consistent within individual participants, but this consistency is substantially weaker across participants. Furthermore, these can be used as covariates in analyses of other tasks using Quaddles, thus controlling for any effects that differential sensitivities, strong response biases, or different efficiencies might produce on these tasks.

Method

Stimulus generation

Quaddles were generated using Autodesk 3DS Max software. The primary set, used to generate all results presented in this article, is defined by four feature dimensions (body shape, branch angularity, pattern, and color), each of which can take

on two possible feature values (e.g., body shape can be pyramidal or oblong), giving a total of eight feature values and 16 possible objects (Fig. 2). The scripts we provide online include options for incorporating further feature dimensions and values into stimulus sets, allowing the creation of hundreds of thousands more objects, but we do not describe these in detail here (full instructions are given in the manual).

Textures, which define the surface colors and patterns, are imported from PNG files created in a Matlab script. The neutral gray color is the same for all objects, whereas the other colors are chosen within the CIE $L^*c^*h^*$ space such that the L^* and c^* values (luminance and saturation, respectively) are held constant, but h^* values (hue) vary by 15° , meaning that there is a small difference in hue between the two colors, but not in the other components of the colors. Textures are applied to object surfaces using standard UV mapping options: a cylindrical wrap for pyramidal bodies and a spherical wrap for oblong ones (different wraps were chosen because they resulted in smaller artifacts at the top and bottom of objects).

Quaddle bodies are initially generated as spheres and then molded into the desired body shapes using freeform deformation (Sederberg & Parry, 1986), in which a lattice of control points is added to the object and manipulated to create the desired shape. Thus, all body shapes are morphs of each other, allowing for intermediary shapes as desired. Each Quaddle has four arms, initially generated as straight cylinders and then morphed into the desired shape, thus also supporting intermediary values. The same is true of both the hues used to define object color, and their surface patterns. This means it is easy to create objects chosen from anywhere within the feature space

defined by the four feature dimensions (Fig. 3). Given any two objects, one can also create videos of the morph between them, or even objects that morph in real time in a 3-D environment.

Object generation was automated using a 3DS Maxscript that creates and saves complete object sets. An optional function allows JPEG, PNG, or other image files to be generated of every object created from any distance, height, and rotation. Experimenters also have the option of saving videos of the objects rotating 360° from any perspective.

For illustration purposes, we generated two more feature values along each dimension, and generated partial morphs of the objects along each dimension (Fig. 3). We also generated a number of further objects with varied numbers of arms, amalgamated body shapes stacked on top of each other, and various new textures and shapes (Fig. 4). Making new Quaddles in this way is quite easy using simple modifications of our existing scripts, allowing the powerful and flexible generation of new object sets.

Experimental procedures

The York University Office of Research Ethics approved the present study as conforming to the standards of the Canadian Tri-Council Research Ethics guidelines (Certificate # 2016-214). Ten participants (mean age 28 years \pm 3.8 *SE*; six male, four female) took part in the study. One was excluded from further analyses due to chance performance. They were seated approximately 60 cm from an LED monitor with a 60-Hz refresh rate, with heads unrestrained. The entire study, including an instructional tutorial, took approximately 1 h. The task was coded in the Unity game engine.

For the duration of the experiment, participants viewed a diamond-shaped arena from one of its vertices (Fig. 5). At the start of each trial, the floor of the arena changed to one of 20

different textures, chosen at random, which were also used in the learning task for which we designed the Quaddles. The floor textures were chosen from a large, free database of textures (<https://share.allegorithmic.com>), and included a wide variety of hues, contrasts, spatial frequencies, and semantic information. After 200 ms, two cues appeared, each showing an iconic representation of one of the eight stimulus feature values (2 feature values \times 4 feature dimensions). After a further 250 ms (\pm 50-ms jitter), a single Quaddle was displayed at the center of the screen for 250 ms, subtending 3.5°–4.0° of visual angle at a 60-cm viewing distance. Participants had to quickly decide whether the single Quaddle contained the feature value of the left or right iconic image cue. A mask pattern was then flashed over the Quaddle location for 50 ms, after which both it and the Quaddle were removed. The cues remained on screen until participants had responded by pressing either the “Z” or the “/” key on a standard keyboard (indicating that the left or right cue, respectively, was accurate), or until 2 s had elapsed, whichever was quicker. If participants did not respond within 2 s, the game was paused and they were asked to respond more quickly on future trials. After response, feedback was presented for 500–600 ms in the form of a colored border around the chosen cue, with green indicating correct and red indicating incorrect. Each Quaddle had one of the two cued feature values, but not the other, and participants were tasked with reporting which of the two cues was present. After feedback the cues were removed, and the next trial started immediately (see Fig. 5).

Participants were given approximately 5 min of training on a slower version of the task prior to starting, and they were instructed to respond as quickly and accurately as possible. A single block contained 512 trials, consisting of 32 trials for each of the 16 Quaddles. Each of the four feature values

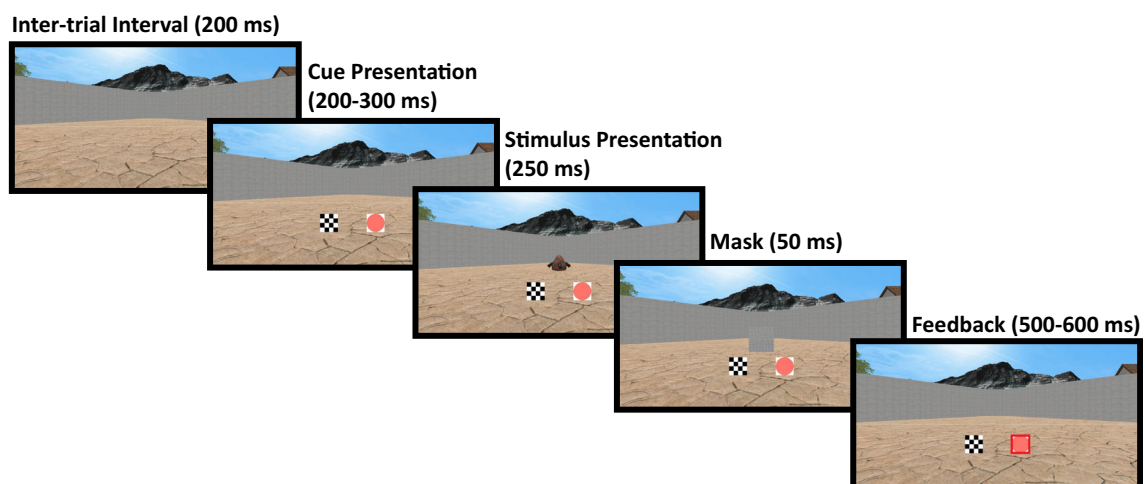


Fig. 5 Overview of a feature detection trial. On this trial, the participant incorrectly responded that the presented Quaddle had a reddish color, instead of a checkered pattern, so their incorrect choice was outlined in

red during the feedback phase. Had they chosen correctly, their choice would have been outlined in green

present on a given Quaddle was presented as a valid cue eight times, twice with each of the four feature values not found on the same Quaddle as the invalid cue, on both the left and the right side. Every 32 trials, each of the 16 Quaddles was shown twice, once with the valid cue on the right, and once on the left, but in all other respects the cues, Quaddles, and side of the valid cue were randomized. After a block, participants were given an optional break. Most participants ran through three blocks in approximately 60 min, but three participants were only able to finish two due to time constraints.

Several pilot versions of the study were run. After each, we adjusted the object feature values to try to eliminate any gross response biases. We present the results only for participants run using the final set of feature values, which had the most unbiased performance across feature dimensions.

Results

One participant was excluded from analyses due to chance accuracy. For the remaining participants, all trials in which a given feature value was a valid cue were grouped together, as were all trials in which it was an invalid cue. This resulted in 16 groups of trials, within which the accuracy and response time on correct trials were averaged. There were substantial interindividual differences, with individual accuracies for particular feature values ranging from 54% to 99%, and correct response times ranging from 500 to 1,180 ms.

For each of the eight feature values, we calculated a hit rate (the accuracy on all trials in which the target was validly cued) and a false alarm rate (1 – accuracy on all trials in which the target was invalidly cued). This effectively treated our experiment as eight independent yes–no tasks, where reporting a given feature value as present corresponded to a “yes” response and reporting any other value as present corresponded to “no,” allowing for the calculation of standard sensitivity (d') and bias (c , criterion) measures (Macmillan & Creelman, 2005), shown in Fig. 6. A c of 0 indicates no bias, negative numbers indicate a preference for a given alternative, and positive numbers indicate a preference against a given alternative (thus, Fig. 6b shows very minor biases for most features, but stronger biases for participants to report checkered patterns and not to report downward-bending arms). Individual sensitivities for particular feature values ranged from 1.01 to 4.42, and the absolute values of biases ranged from 0 to .44.

The mean sensitivity across all participants and all feature values was 2.59, which lay between Quartiles 1 and 2 for all feature values. The checkered and diamond patterns were clearly different from the other feature values, in that they had the two lowest median sensitivities, which were the only median sensitivities below the first quartiles of any of the other feature values, as well as third quartiles that were below two of the other values' median sensitivities (orange color and

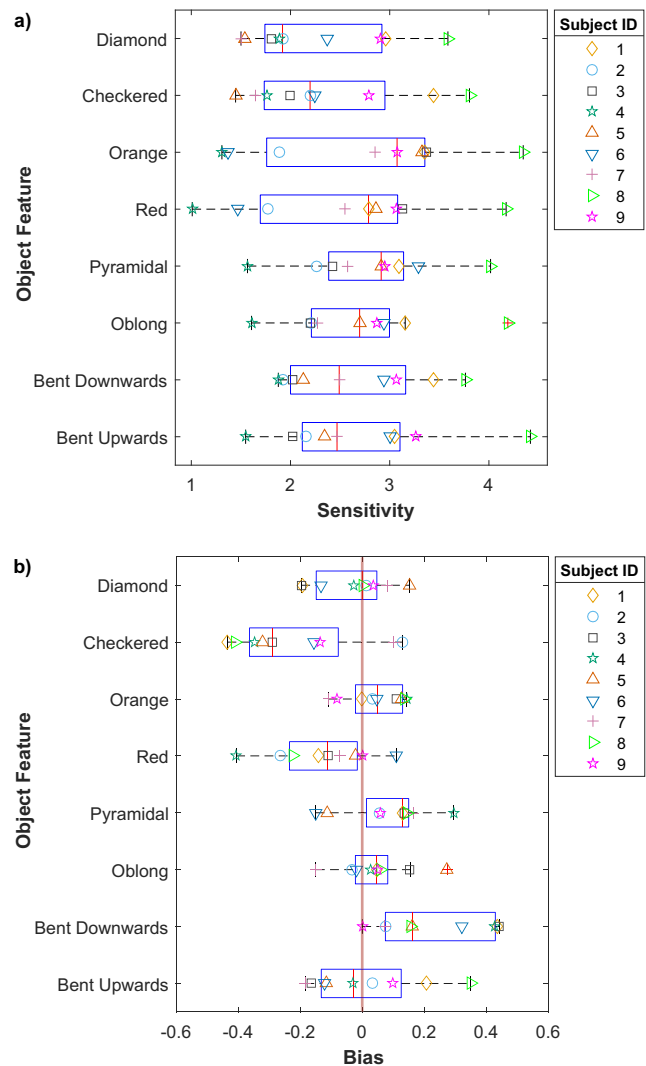


Fig. 6 Feature detection study sensitivities (a) and response biases (b) for each of the eight feature values. Higher sensitivity scores indicate more accurate detection of the feature in question; negative bias indicates a tendency to report the feature in question; positive bias shows a tendency not to report it; and 0 indicates no bias

pyramidal shape). All other median sensitivities were between the first and third quartiles for all other feature values, with a single exception (the oblong shape's third quartile was below the median sensitivity of the orange color). Thus, participants were consistently less sensitive to pattern than to the other feature dimensions. Consistent with this, trials with either valid or invalid pattern cues had the lowest median response times (not shown), indicating that the values of the pattern dimension were particularly difficult to discriminate and that the lower sensitivity was not simply due to a speed–accuracy trade-off.

The response biases were clustered around 0, with a median bias for one feature value (diamond pattern) at exactly 0, three below 0, and four above. The absolute value of all medians was less than .20, with the exception of the checkered pattern, which had a bias of –.29. All feature values produced both positive

and negative biases in different participants, with the exception of downward-bent arms, which had no bias or a positive bias for all participants. Thus, in general, feature values did not produce strong and consistent response biases, with the exceptions of checkered patterns and downward-bent arms.

Another method of quantifying performance on two-alternative forced choice tasks such as these, which has the advantage of producing a single number for each participant while controlling for a speed–accuracy trade-off, is to compute *efficiency scores* from the raw accuracies and response times (Smilek et al., 2006; Townsend & Ashby, 1983), which we used to quantify interindividual differences in performance and to examine the degree to which individuals’ detection efficiency was self-consistent across blocks and consistent between participants (Figs. 7 and 8).

There were 56 cue pairs (8×8 feature values, but the same value cannot be paired with itself), for which accuracies and response times were combined by dividing the response time on correctly answered trials in which these cues were presented by the mean accuracy on these trials (Smilek et al., 2006; Townsend & Ashby, 1983). Doing this corrects for speed–accuracy trade-offs in an intuitive way: When accuracy is perfect, efficiency will be identical to the mean response time, and as accuracy decreases, the combined score increases, making efficiency scores similar to response time in that smaller values indicate better performance. This assumes that response time and accuracy are linearly correlated, which was supported in the present data ($r = -.53, p < .001$). These scores were then transformed into z scores using the mean and

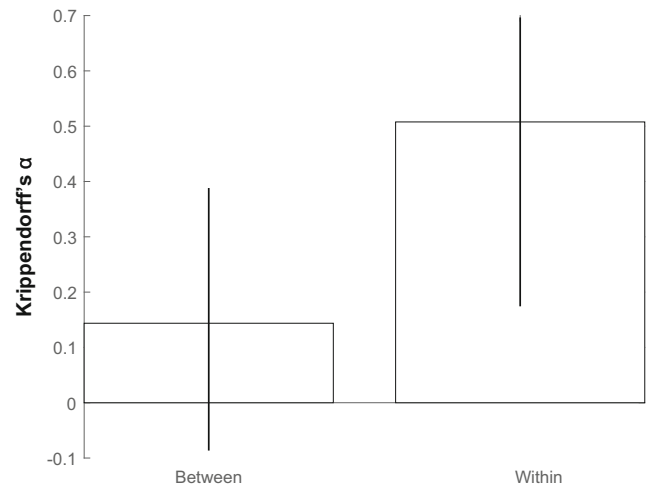


Fig. 8 Krippendorff’s alpha values for normalized efficiency score ranks for feature values between and within participants. The between-participant scores are close to 0, indicating that any given individual’s feature value preferences are close to random. Within-participant scores, on the other hand, are higher, indicating that individual preferences are stable across multiple blocks. Error bars indicate 95% confidence intervals, calculated using 10,000 bootstrap samples

standard deviation of the efficiency scores across all trials for each participant. These 56 scores were combined into 28 for each participant by averaging over the valid–invalid and invalid–valid trials for each feature value pair, which was also justified, on the basis of a strong correlation between the two sets of scores ($r = .83, p < .001$). Finally, these 28 feature value pair scores were combined into eight feature value scores, by

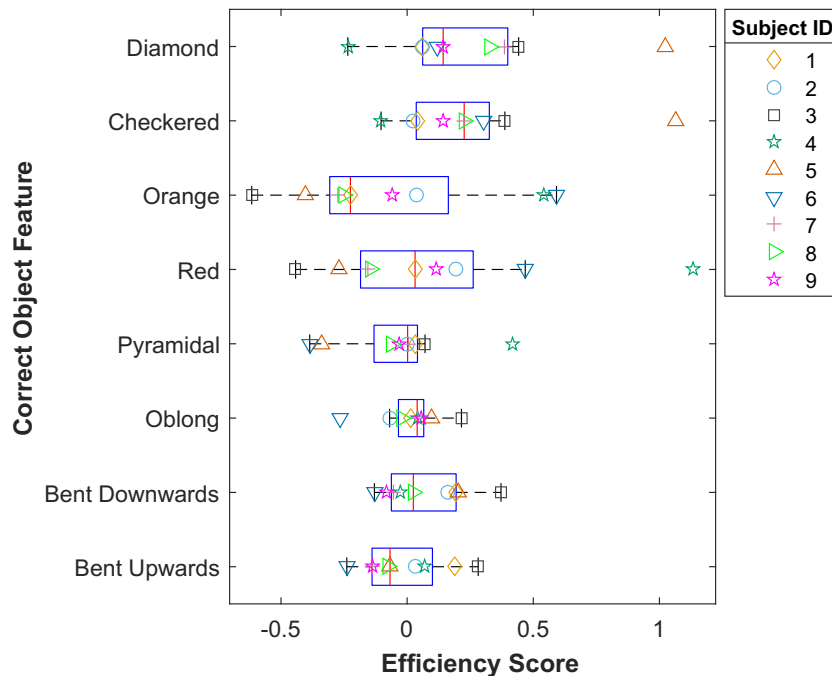


Fig. 7 Normalized efficiency scores on the feature identification task. Lower numbers indicate better performance. 0 indicates each participant’s mean efficiency across all trials, and 1 indicates a standard deviation away from this mean

averaging all trials in which a given feature value was presented as a cue with any of the other seven feature values. A score of 0 indicated the mean efficiency across all feature values for each participant, and 1 (or -1) indicated a standard deviation away from this mean, with lower scores indicating greater efficiency.

These normalized efficiency scores are shown in Fig. 7. Their median values across all participants are close to 0, with no scores being more than 0.25 *SDs* away from the mean, indicating that most of the variance in response times and accuracies on the present task was due to factors other than response biases to particular feature values. All the 95% confidence intervals for these scores included 0, indicating little, if any, difference between the efficiencies for any feature values. However, once again the two surface pattern feature values have the two highest scores (.23 and .14), consistent with the lowered sensitivity scores to these values (Fig. 6).

As well as quantifying efficiency scores for particular feature values, we wanted to quantify how consistent these scores were, both between and within participants. This was accomplished using a tool from the content analysis field, *Krippendorff's alpha*, or $K\alpha$ (Hayes & Krippendorff, 2007; Krippendorff, 2011). $K\alpha$, which indicates the reliability of multiple sets of scores for a number of items, ranges between -1 and 1, where 1 indicates perfect consistency, 0 indicates a completely random distribution of scores across sets, and -1 indicates perfectly systematic disagreement (Krippendorff, 2008). Generally speaking, $K\alpha$ is used to measure the consistency of questionnaires or other rating instruments, in which case a high value (e.g., .80 or higher) is desirable. However, in the present case, values approaching 0 indicated a *lack of* consistent bias toward particular feature values, as we desired for our object set.

To calculate $K\alpha$, the normalized efficiency scores were transformed to rank orders. For the between-participants $K\alpha$, this was done across all trials to give a single set of scores for each participant. $K\alpha$ was calculated using a freely available Matlab script (Eggink, 2012), and standard errors and confidence intervals were calculated using a bootstrap method, as recommended by Zapf, Castell, Morawietz, and Karch (2016), save that we used 10,000 samples due to the small number of participants, and used bias-corrected and accelerated confidence intervals, which provide more accurate estimates of the true interval (DiCiccio & Efron, 1996). The resulting between-participant consistency was low ($K\alpha = .14$, $SE = .17$), and its 95% confidence interval included 0. For the within-participants $K\alpha$, we calculated a separate set of efficiency scores for each block performed by each participant, found the rank-ordering of these scores, then calculated a single across-block $K\alpha$ for each participant using these rank orders and averaged these values across participants, using a 10,000-sample bootstrap to calculate standard errors and confidence intervals. This showed a substantial degree of within-

participant consistency ($K\alpha = .51$, $SE = .12$). Thus, individuals have reasonably consistent efficiencies to the different feature values over time, although across individuals these efficiencies are much closer to randomly distributed (see Fig. 8).

Discussion

This article presented the multidimensional set of Quaddle objects, suitable for use both in studies involving navigation through visually appealing virtual 3-D worlds, or for more static studies that require images or videos of multidimensional objects. The results of a simple feature discrimination task showed that participants are roughly equally sensitive to the feature values along each dimension (Fig. 6a), with the exception of the two pattern dimensions; that there are not strong and consistent response biases to most of these feature values (Fig. 6b); and that although individuals' response efficiencies for particular feature values are somewhat consistent across time, there is little consistency across participants (Figs. 7 and 8). These results suggest that Quaddles can be used "off the shelf" in a wide array of tasks that require balanced feature values, and they point to modifications that could further improve this balance (e.g., by increasing the size of individual checkers or diamonds, or increasing the contrast between their dark and light elements, to increase the discriminability of the patterns). They can even be 3-D-printed, should an experiment be truly real-world, and can generate stereoscopic images for use with virtual reality or augmented reality experiments.

Limitations of the discrimination study

Although the results of our feature detection task are encouraging, it is important to acknowledge their limitations. First, there is substantial individual variability. For a study in which it is critical that each feature value be equally discriminable by all participants (for some arbitrarily small value of "equally"), it might be necessary to produce a much larger set of objects with many intermediate feature values, and run participants on a complex adaptive staircase task (see, e.g., Anderson & Johnson, 2006; Klein & Macmillan, 2001; Kujala & Lukka, 2006; Treutwein, 1999), resulting in a personalized set of objects for each participant. Producing such intermediate objects is possible with relatively simple modifications of our existing scripts. Developing such a staircase task would require careful consideration of the specific requirements of the experiment in question.

Our study presents objects at a single distance. In a study in which object distances vary, such as any involving navigation through a 3-D world, different features will become more or less discriminable at different distances. Equating discriminability across multiple distances would make for a much longer and more complex study than was feasible in our time frame.

Finally, our study presented objects within a single arena that does not change, save for the floor, which changes drastically across trials. This is simply because the study for which we developed these objects involves a single arena with floors that change across trials. The surround of an object can have powerful effects on feature discrimination, but our study does not control for these effects, as we reasoned that given the wide variety of floors we present, their effects would be essentially random. Once again, any experimenter for whom this is a critical concern will have to run another set of feature detection studies, modifying the objects and environment as needed.

We believe it is apparent that *controlling for all factors that affect feature detection is not feasible*, because one can only measure detection in a given experimental context. We have mentioned three factors that we did not control for, which likely interact in highly non-linear ways: individual preferences, object distance, and visual background. Even if their interactions are completely linear, controlling for all of them simultaneously would require a very complex task and a very large number of participants. Furthermore, there are certainly other relevant contingencies that we have not outlined here. Indeed, preliminary results from our laboratory suggest that in a dynamic object selection task in which the rewarded feature values are learned through trial and error, Quaddle shapes may have an advantage over the other feature dimensions (data not shown). Sensitivity, response bias, and efficiency scores need to be measured in their respective task environments, consistent with the insight that multiple factors determine whether an object guides attention in visual search tasks (Wolfe & Horowitz, 2017).

Experimenters will have to determine to what degree their particular task requires controlling for different factors that might affect feature value discrimination and detection, and design their objects and tasks accordingly. Alternatively, instead of controlling for such factors, one could simply quantify their differential effects, and use sensitivity, bias, or efficiency scores as covariates in statistical analyses, to be partialled out from the main effects of the respective studies.

Possibilities for further customization

Figures 2, 3, and 4 show several ways in which the basic Quaddle feature space can be manipulated, but there are many other ways in which researchers might change Quaddles for their own purposes. For example, they might wish to remove the vertical symmetry of some, or all, feature dimensions, so that manipulating or navigating around objects would be an important part of identifying them, as is the case with most of the objects shown in Fig. 1, as well as with many, if not most, real-world objects. Similarly, it might also be of interest to systematically vary the discriminability or salience of different feature dimensions, and to quantify this variance using a feature detection task. This enables the role of feature bias to be

directly studied, as opposed to minimized as with the present object set. Such changes would require trivial modifications to the existing scripts. Aside from these examples, many other possibilities of course exist for future studies to implement additional changes.

Concluding remarks

With this article, we introduced a new object set, characterized its discriminability, and provided tools to facilitate its use in a wide range of possible future studies. This novel set of 3-D objects has normed, parametric features, suitable for a wide range of tasks; open online access to the examples and tools allows researchers to rapidly create custom object sets suitable for other studies. This pragmatic aspect resonates well with the spirit of recent toolkits for video game engines that have streamlined the development and running of dynamic experiments (Doucet, Gulli, & Martinez-Trujillo, 2016; Jangraw, Johri, Gribetz, & Sajda, 2014). The properties of Quaddles make them a suitable set of novel objects for future studies using more realistic and complex tasks, and the scripts we provide can significantly reduce the time necessary to develop different objects for such tasks.

Author note This work was supported by Grant MOP 102482 from the Canadian Institutes of Health Research (to T.W.) and by the Natural Sciences and Engineering Research Council of Canada, Brain in Action CREATE-IRTG program (to M.R.W. and T.W.). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of this article. The authors thank Hongying Wang for technical support, and Isabel Gauthier for comments on a draft version of the manuscript. The study described herein was approved by the York University Office of Research Ethics (Certificate # 2016-214).

References

- Aivar, M. P., Hayhoe, M. M., Chizk, C. L., & Mruczek, R. E. B. (2005). Spatial memory and saccadic targeting in a natural task. *Journal of Vision*, 5(3), 177–193. <https://doi.org/10.1167/5.3.3>
- Anderson, A. J., & Johnson, C. A. (2006). Comparison of the ASA, MOBS, and ZEST threshold methods. *Vision Research*, 46, 2403–2411. <https://doi.org/10.1016/j.visres.2006.01.018>
- Aranyi, G., Kouider, S., Lindsay, A., Prins, H., Ahmed, I., Jacucci, G., . . . Cavazza, M. (2014). Subliminal cueing of selection behavior in a virtual environment. *Presence: Teleoperators and Virtual Environments*, 23, 33–50. https://doi.org/10.1162/PRES_a_00167
- Amott, S. R., Cant, J. S., Dutton, G. N., & Goodale, M. (2008). Crinkling and crumpling: An auditory fMRI study of material properties. *NeuroImage*, 43, 368–378.
- Barral, O., Aranyi, G., Kouider, S., Lindsay, A., Prins, H., Ahmed, I., . . . Cavazza, M. (2014). Covert persuasive technologies: bringing subliminal cues to human–computer interaction. In A. Spagnolli, L. Chittaro, & L. Gamberini (Eds.), *Persuasive technology* (pp. 1–12). Berlin, Germany: Springer International Publishing.
- Barry, T. J., Griffith, J. W., De Rossi, S., & Hermans, D. (2014). Meet the Fribbles: Novel stimuli for use within behavioural research.

- Frontiers in Psychology*, 5, 103. <https://doi.org/10.3389/fpsyg.2014.00103>
- Barry, T. J., Griffith, J. W., Vervliet, B., & Hermans, D. (2016). The role of stimulus specificity and attention in the generalization of extinction. *Journal of Experimental Psychopathology*, 7, 143–152.
- Bennett, M., Vervoort, E., Boddez, Y., Hermans, D., & Baeyens, F. (2015). Perceptual and conceptual similarities facilitate the generalization of instructed fear. *Journal of Behavior Therapy and Experimental Psychiatry*, 48, 149–155. <https://doi.org/10.1016/j.jbtep.2015.03.011>
- Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 1162–1182. <https://doi.org/10.1037/0096-1523.19.6.1162>
- Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*, 12, 752–762. <https://doi.org/10.1038/nrn3122>
- Buffat, S., Chastres, V., Bichot, A., Rider, D., Benmussa, F., & Lorenceau, J. (2014). OB3D, a new set of 3D objects available for research: a web-based study. *Frontiers in Psychology*, 5, 1062. <https://doi.org/10.3389/fpsyg.2014.01062>
- Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences*, 89, 60–64.
- Cant, J. S., & Goodale, M. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cerebral Cortex*, 17, 713–731.
- Chuang, L. L., Vuong, Q. C., & Bülthoff, H. H. (2012). Learned non-rigid object motion is a view-invariant cue to recognizing novel objects. *Frontiers in Computational Neuroscience*, 6, 26. <https://doi.org/10.3389/fncom.2012.00026>
- Cooperrider, K. (2015). The co-organization of demonstratives and pointing gestures. *Discourse Processes*, 53, 632–656. <https://doi.org/10.1080/0163853X.2015.1094280>
- DiCiccio, T. J., & Efron, B. (1996). Bootstrap confidence intervals. *Statistical Science*, 11, 189–228.
- Doucet, G., Gulli, R. A., & Martinez-Trujillo, J. C. (2016). Cross-species 3D virtual reality toolbox for visual and cognitive experiments. *Journal of Neuroscience Methods*, 266, 84–93. <https://doi.org/10.1016/j.jneumeth.2016.03.009>
- Eggink, J. (2012). kriAlpha [Matlab script]. Retrieved August 30, 2017 from <https://www.mathworks.com/matlabcentral/fileexchange/36016-krippendorff-s-alpha>
- Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science*, 317, 1048. <https://doi.org/10.1126/science.1142175>
- Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., & Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature*, 425, 184–188. <https://doi.org/10.1038/nature01964>
- Estes, Z., Jones, L. L., & Golonka, S. (2012). Emotion affects similarity via social projection. *Social Cognition*, 30, 584–609.
- Gauthier, I., James, T. W., Curby, K. M., & Tarr, M. J. (2003). The influence of conceptual knowledge on visual discrimination. *Cognitive Neuropsychology*, 20, 507–523.
- Gauthier, I., & Tarr, M. J. (1997). Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision Research*, 37, 1673–1682. [https://doi.org/10.1016/S0042-6989\(96\)00286-6](https://doi.org/10.1016/S0042-6989(96)00286-6)
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform ‘face area’ increases with expertise in recognizing novel objects. *Nature Neuroscience*, 2, 568–573.
- Harman, K. L., & Humphrey, G. K. (1999). Encoding “regular” and “random” sequences of views of novel three-dimensional objects. *Perception*, 28, 601–615.
- Harman, K. L., Humphrey, G. K., & Goodale, M. (1999). Active manual control of object views facilitates visual recognition. *Current Biology*, 9, 1315–1318.
- Harris, J. (2015). Yadgits. Retrieved August 24, 2017, from http://wiki.cnbcmu.edu/Novel_Objects, last modified May 14, 2015
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1, 77–89.
- Hayward, W. G., & Tarr, M. J. (1997). Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1511–1521. <https://doi.org/10.1037/0096-1523.23.5.1511>
- Humphrey, G. K., & Khan, S. C. (1992). Recognizing novel views of three-dimensional objects. *Canadian Journal of Psychology*, 46, 170–190.
- Jangraw, D. C., Johri, A., Gribetz, M., & Sajda, P. (2014). NEDE: An open-source scripting suite for developing experiments in 3D virtual environments. *Journal of Neuroscience Methods*, 235, 245–251. <https://doi.org/10.1016/j.jneumeth.2014.06.033>
- Johnson, L., Sullivan, B., Hayhoe, M., & Ballard, D. (2014). Predicting human visuomotor behaviour in a driving task. *Philosophical Transactions of the Royal Society B*, 369, 20130044. <https://doi.org/10.1098/rstb.2013.0044>
- Jovancevic, J., Sullivan, B., & Hayhoe, M. (2006). Control of attention and gaze in complex environments. *Journal of Vision*, 6(12), 9: 1431–1450. <https://doi.org/10.1167/6.12.9>
- Kaldy, Z., Blaser, E. A., & Leslie, A. M. (2006). A new method for calibrating perceptual salience across dimensions in infants: the case of color vs. luminance. *Developmental Science*, 9, 482–489.
- Klein, S. A., & Macmillan, N. A. (2001). Threshold estimation: The state of the art. *Perception & Psychophysics*, 63, 1277–1278.
- Knutson, A. R., Hopkins, R. O., & Squire, L. R. (2012). Visual discrimination performance, memory, and medial temporal lobe function. *Proceedings of the National Academy of Sciences*, 109, 13106–13111. <https://doi.org/10.1073/pnas.1208876109>
- Krippendorff, K. (2008). Systematic and random disagreement and the reliability of nominal data. *Communication Methods and Measures*, 2, 323–338. <https://doi.org/10.1080/19312450802467134>
- Krippendorff, K. (2011). Computing Krippendorff’s alpha-reliability. Retrieved August 28, 2017, from http://repository.upenn.edu/asc_papers/43
- Kujala, J. V., & Lukka, T. J. (2006). Bayesian adaptive estimation: The next dimension. *Journal of Mathematical Psychology*, 50, 369–389.
- Leeb, R., Friedman, D., Müller-Putz, G. R., Scherer, R., Slater, M., & Pfurtscheller, G. (2007). Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: A case study with a tetraplegic. *Computational Intelligence and Neuroscience*, 2007, 79642. <https://doi.org/10.1155/2007/79642>
- Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum: Manipulating bodily self-consciousness. *Science*, 317, 1096–1099. <https://doi.org/10.1126/science.1143439>
- Macmillan, N. A., & Creelman, D. C. (2005). *Detection theory: A user’s guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mercer, T., & Duffy, P. (2015). The loss of residual visual memories over the passage of time. *Quarterly Journal of Experimental Psychology*, 68, 242–248. <https://doi.org/10.1080/17470218.2014.975256>
- Richler, J. J., Wilmer, J. B., & Gauthier, I. (2017). General object recognition is specific: Evidence from novel and familiar objects. *Cognition*, 166, 42–55.
- Scheveneels, S., Boddez, Y., Bennett, M. P., & Hermans, D. (2017). One for all: The effect of extinction stimulus typicality on return of fear. *Journal of Behavior Therapy and Experimental Psychiatry*, 57, 37–44.
- Sederberg, T. W., & Parry, S. R. (1986). Free-form deformation of solid geometric models. In D. C. Evans & R. J. Athay (Eds.), SIGGRAPH ’86: Proceedings of the 13th Annual Conference on Computer

- Graphics and Interactive Techniques (pp. 151–160): New York, NY: ACM.
- Smilek, D., Enns, J. T., Eastwood, J. D., & Merikle, P. M. (2006). Relax! Cognitive strategy influences visual search. *Visual Cognition, 14*, 543–564.
- Tarr, M. J., Bülthoff, H. H., Zabinski, M., & Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint. *Psychological Science, 8*, 282–289. <https://doi.org/10.1111/j.1467-9280.1997.tb00439.x>
- Townsend, J. T., & Ashby, F. G. (1983). Stochastic modeling of elementary psychological processes. New York, NY: Cambridge University Press.
- Treutwein, B. (1999). Adaptive psychophysical procedures. *Vision Research, 35*, 2503–2522.
- Wallraven, C., Bülthoff, H. H., Waterkamp, S., van Dam, L., & Gaissert, N. (2014). The eyes grasp, the hands see: Metric category knowledge transfers between vision and touch. *Psychonomic Bulletin & Review, 21*, 976–985. <https://doi.org/10.3758/s13423-013-0563-4>
- Watrous, A. J., Tandon, N., Conner, C. R., Pieters, T., & Ekstrom, A. D. (2013). Frequency-specific network connectivity increases underlie accurate spatiotemporal memory retrieval. *Nature Neuroscience, 16*, 349–356. <https://doi.org/10.1038/nn.3315>
- Watson, M. R., Voloh, B., Naghizadeh, M., Chen, S., & Womelsdorf, T. (2017). Information sampling and object selection strategies demonstrate the learning and exploitation of feature relevance. In *2017 Neuroscience Meeting Planner* (Program No. 801.30). Washington, DC: Society for Neuroscience.
- Weisberg, S. M., Schinazi, V. R., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2014). Variations in cognitive maps: Understanding individual differences in navigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 669–682. <https://doi.org/10.1037/a0035261>
- Williams, P. (1998). Representational organization of multiple exemplars of object categories (Working paper, University of Massachusetts at Boston). Retrieved August 23, 2017, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.5.8336&rep=rep1&type=pdf>
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behavior, 1*, 0058. <https://doi.org/10.1038/s41562-017-0058>
- Wong, A. C.-N., & Hayward, W. G. (2005). Constraints on view combination: Effects of self-occlusion and differences among familiar and novel views. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 110–121. <https://doi.org/10.1037/0096-1523.31.1.110>
- Wong, A. C.-N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for face-like expertise with objects: Becoming a Ziggerin expert—but which type? *Psychological Science, 20*, 1108–1117. <https://doi.org/10.1111/j.1467-9280.2009.02430.x>
- Zapf, A., Castell, S., Morawietz, L., & Karch, A. (2016). Measuring interrater reliability for nominal data—Which coefficients and confidence intervals are appropriate. *BMC Medical Research Methodology, 16*, 93. <https://doi.org/10.1186/s12874-016-0200-9>