OXFORD

Article

Technical and conceptual considerations for using animated stimuli in studies of animal behavior

Laura Chouinard-Thuly[a,*,†], Stefanie Gierszewski[b,*,†], Gil G. Rosenthal[c,d], Simon M. Reader[a], Guillaume Rieucau[e], Kevin L. Woo[f], Robert Gerlai[g], Cynthia Tedore[h], Spencer J. Ingley[i], John R. Stowers[j,k], Joachim G. Frommen[l], Francine L. Dolins[m] and Klaudia Witte[b]

[a]Department of Biology, McGill University, 1205 Docteur Penfield, Montréal, Quebec, Canada H3A 1B1, [b]Research Group of Ecology and Behavioral Biology, Institute of Biology, University of Siegen, Adolf-Reichwein Str. 2, Siegen 57068, Germany, [c]Ecology & Evolutionary Biology, Texas A&M University, 3258 TAMU College Station, TX 77843, USA, [d]Centro de Investigaciones Científicas de las Huastecas "Aguazarca", Calnali, Hidalgo, México, [e]Department of Biological Sciences, Florida International University, 3000 Northeast 151 Street, North Miami, FL 33181, USA, [f]SUNY Empire State College, Metropolitan Center, 325 Hudson Street, New York, NY 10013-1005, USA, [g]Department of Psychology, University of Toronto Mississauga, 3359 Mississauga Road, Mississauga, Ontario, Canada L5L 1C6, [h]Lund Vision Group, Department of Biology, Lund University, Sölvegatan 35, Lund 22362, Sweden, [i]Department of Biology, University of North Carolina at Chapel Hill, CB#3280, Coker Hall, Chapel Hill, NC 27599, USA, [j]Research Institute of Molecular Pathology IMP, Vienna Biocenter VBC, Dr. Bohr-Gasse 7, Vienna 1030, Austria, [k]loopbio gmbh, Hauptstrasse 93, Kritzendorf 3420, Austria, [l]Department of Behavioural Ecology, Institute of Ecology and Evolution, University of Bern, Wohlenstrasse 50a, Hinterkappelen 3032, Switzerland, and [m]Department of Behavioral Sciences, University of Michigan-Dearborn, 4901 Evergreen Road, Dearborn, MI 48128, USA

*Address correspondence to Laura Chouinard-Thuly, E-mail: laura.thuly@gmail.com and Stefanie Gierszewski, E-mail: gierszewski@chemie-bio.uni-siegen.de.

[†]The authors equally contributed to this work.

Abstract

Rapid technical advances in the field of computer animation (CA) and virtual reality (VR) have opened new avenues in animal behavior research. Animated stimuli are powerful tools as they offer standardization, repeatability, and complete control over the stimulus presented, thereby "reducing" and "replacing" the animals used, and "refining" the experimental design in line with the 3Rs. However, appropriate use of these technologies raises conceptual and technical questions. In this review, we offer guidelines for common technical and conceptual considerations related to the use of animated stimuli in animal behavior research. Following the steps required to create an animated stimulus, we discuss (I) the creation, (II) the presentation, and (III) the validation of CAs and VRs. Although our review is geared toward computer-graphically designed stimuli, considerations on presentation and validation also apply to video playbacks. CA and VR allow both new behavioral questions to be addressed and existing questions to be addressed in new ways, thus we expect a rich future for these methods in both ultimate and proximate studies of animal behavior.

Key words: animal behavior, animated stimulus, computer animation, experimental design, virtual reality, visual communication.

Introduction

Recent advances in the technical development of **computer animations (CAs)** and **virtual reality (VR)** systems in computer sciences and film have primed the technologies for adoption by behavioral researchers with a variety of interests. CAs are computer graphically generated stimuli which, in contrast to video playback, allow full control of stimulus attributes and can be **pre-rendered** or **rendered in real-time**. VRs are also computer-generated stimuli but are rendered in real-time and display perspective-correct views of a 3D scene, in response to the behavior of the observer. CA and VR are powerful alternatives to live or real-world stimuli because they allow a broader range of visual stimuli together with standardization and repeatability of stimulus presentation. They afford numerous opportunities for testing the role of visual stimuli in many research fields where manipulation of visual signals and cues is a common and fruitful approach (Espmark et al. 2000). Historically, researchers have tried to isolate the specific cues [e.g., "key" or "sign" stimuli; Tinbergen (1948); see overview in Gierszewski et al. (2017)] that trigger a behavioral response, but this is difficult to accomplish with live stimuli or in an uncontrolled environment (Rowland 1999). In contrast, researchers can use CA and VR to create visual stimuli with single traits or trait combinations that are difficult or impossible to achieve using live animals without surgical or other manipulations [compare Basolo (1990) with Rosenthal and Evans (1998)]. It is also possible to present phenotypes of animals that are encountered only very rarely in the wild (e.g., Schlupp et al. 1999), to present novel phenotypes (e.g., Witte and Klink 2005), or to vary group composition or behavior (e.g., Ioannou et al. 2012; Gerlai 2017). With CA and VR, researchers can further allow stimuli to be shaped by evolutionary algorithms and create entire virtual environments (Ioannou et al. 2012, Dolins et al. 2014, Thurley et al. 2014; Dolins et al. in preparation; Thurley and Ayaz 2017). Finally, CA and VR allow "replacement" and "reduction" of animals used for experimentation, as well as "refinement" of experimental design, which is important for both practical and ethical reasons and thus addressing the requirements of the "3Rs" (Russel and Burch 1959; ASAB 2014). Yet, despite the demonstrated achievements of these techniques over the last few decades and their promise for behavioral research, relatively few researchers have adopted these methods (see Supplementary Table S1). This may be due to technical and methodological hurdles in using computer graphics in behavioral research. Here, we aim to address these difficulties and to discuss technical and conceptual considerations for the use of animated stimuli to study animal behavior.

Stimulated by the symposium on "Virtual Reality" at the Behaviour 2015 conference in Cairns, Australia, and inspired by a workshop consensus article on considerations for video playback design (Oliveira et al. 2000), here we bring together researchers with varied backgrounds in animal behavior, experimental psychology, and animal visual systems to discuss and build consensus on the design and presentation of CA and VR. We also offer recommendations for avoiding pitfalls, as well as some future research opportunities that these techniques provide (see also Rosenthal 2000; Baldauf et al. 2008). Even though we focus on CA and VR, many considerations discussed below also apply to the use of video playback in animal behavior experiments. With the abundance of novel conceptual and technical applications in this fast developing field, we re-consider limitations and constraints of using CA and VR, and discuss the utility of these methods, and the type of questions they may be able to address in animal behavior and related

disciplines (see also Powell and Rosenthal 2017). This review is divided in 3 sections: (I) how to create animated stimuli, (II) how to present them to nonhuman animals, and (III) how to validate the use of CA and VR regarding the perception of test subjects. A flowchart outlining the most important conceptual and technical questions can be found in Figure 1. We indicate in bold important technical and conceptual terms used in the text, and provide definitions in the glossary (Table 1).

Creation of an Animated Stimulus

In this section, we discuss the creation of animated stimuli. We suggest that the animal's visual system, if known (otherwise see "validation" section), its biology, as well as the research question must drive the decisions about the technological components and the type of animation (**2D**, **3D**, or **VR**) needed (see Figure 1). Although it might be sufficient for certain studies to present a simple, moving 2D shape to elicit a response, in other contexts a highly realistic animation may be required. The term "realistic" is itself ambiguous and limited by what humans can measure. Realism can be considered the sum of many visual cues including; "photo realism",
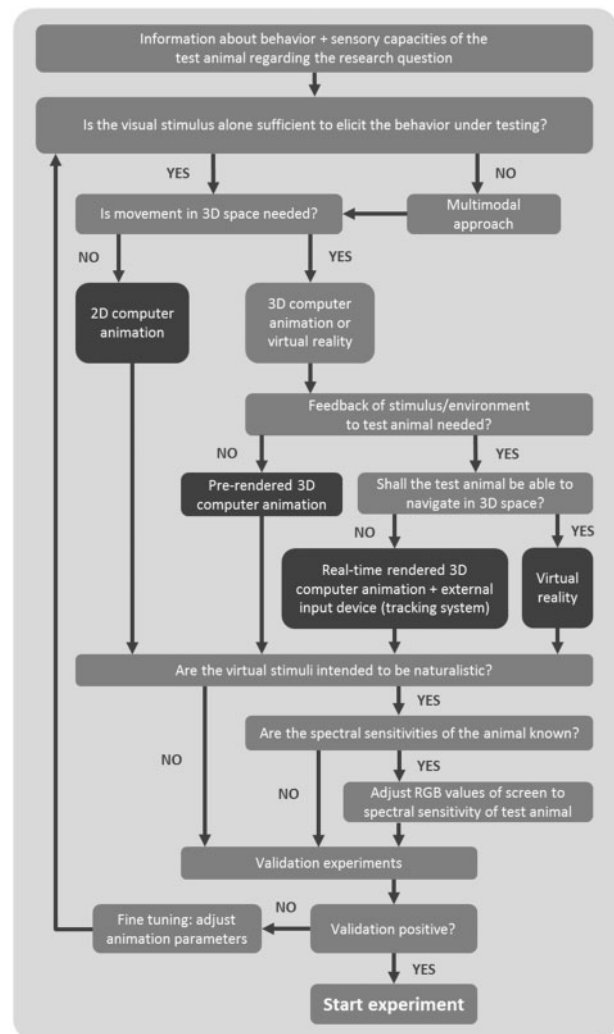

Figure 1. Simplified workflow with the most important conceptual and technical questions that have to be raised when creating and using CA or VR.

Table 1. Glossary: definitions for technical and conceptual terms used in the text, ordered alphabetically

Glossary

2D animation/2D animated stimulus: 2-dimensional animated stimulus

3D animation/3D animated stimulus: 3-dimensional animated stimulus

Computer animation (CA): visual presentation of a moving computer-graphically generated stimulus, presented on a screen to an observer. The stimulus is either animated in 2D (x-, y-axis) or 3D (x-, y-, z-axis) virtual space. CA is usually open-loop and pre-rendered. Viewing perspective on the animated stimuli is not necessarily correct for a moving observer.

CFF: critical flicker-fusion frequency (in Hz). Lowest frequency at which a flashing light is perceived as constantly glowing. Important parameter to consider when using CRT monitors for stimulus presentation.

Closed-loop: the visual stimulus responds to specific actions (movement or behavior) of the observer (c.f. open-loop where the visual stimulus is independent of the actions of the observer).

CRT monitor: cathode ray tube monitor. No longer in general production.

Frame rate/frames per second (fps): commonly refers to image frequency in CAs and video; describes the number of single images (frames) that are displayed in 1 s (fps). Perception of fluent animations depends on the capabilities of the observer's visual system as well as lighting conditions. Frame rate is also frequently called IPR.

Game engine: software framework used to develop video games. Typically provides 2D and 3D graphics rendering in real-time and incorporates options for interaction (e.g., input from video game controller).

Gamut (color): the range of colors that can be presented using a given display device. A display with a large color gamut can accurately show more colors than a display with a narrow gamut.

Geometric morphometrics: a method for analyzing shape that uses Cartesian geometric coordinates rather than linear, areal, or volumetric variables. Points can be used to represent morphological landmarks, curves, outlines, or surfaces.

Interpolation (in keyframing animation): process that automatically calculates and fills in frames between 2 set keyframes to generate continuous movement.

Keyframing: saving different x, y, (z) positions/postures/actions of a stimulus to specific times in the animation and letting the software generate the in-between frames to gain a smooth change of positions.

Latency (lag): response time;
   Latency (display): describes the difference in time between the input of a signal and the time needed to present this signal on screen.
   Latency, closed-loop (in VR): time delay taken from registering a change in the position of the observer, and that change being reflected on the display to ensure viewpoint correct perspective.

LCD monitor: liquid crystal display monitor.

Mesh/polygon mesh: the representation of an object's 3D shape made of vertices, edges, and faces. The mesh provides the base for a 3D model.

Open-loop: see "closed-loop".

Plasma display: a type of flat panel display that uses small cells of electrically charged gas called plasmas. No longer in general production.

Pseudoreplication: " [...] defined as the use of inferential statistics to test for treatment effects with data from experiments where either treatments are not replicated (though samples may be) or replicates are not statistically independent." (Hurlbert 1984). In terms of CAs, this problem arises when measurements of test animals are gained by presenting the identical stimulus (or a pair of stimuli) over and over again, while neglecting natural variation of, for example, a presented trait or cue (see McGregor 2000). Responses toward the presentation of such stimuli cannot be considered as statistically independent replicates.

Rendering: final process of transferring a designed raw template or raw model into the final 3D graphic object by the animation software.
   Pre-rendered animation: rendering of the animated scene or object conducted prior to an experiment and the final output was saved as a movie file for presentation.
   Real-time rendered animation: rendering of the animated scene or object is conducted continuously during the experiment in real-time as response given to input from an external device like a video game controller or subject position data provided by a tracking system. Real-time rendering needs considerably more sophisticated hardware to process constant data flow.

RGB (red, green, blue) color model: color space in which every color is simulated by different proportions of red, green, and blue primaries. Fundamental for color presentation on screens and devices, with each pixel of an image composed of 3 RGB values.

Rig: a mesh can be rigged with a virtual skeleton (a rig of bones and joints) to provide realistic movement for animating the object. Through the process of "skinning" the rig is interconnected to the mesh, providing deformation of the 3D shape while moving certain bones.

Rotoscoping: an animation technique in which animators trace over video footage, frame by frame, to create a realistic animation sequence. In the context of behavioral studies, rotoscoping could be used to create realistic behavioral sequences (e.g., mating display) based on real-life videos of behaviors.

Texture: the visualized surface of a 3D model. Animators can "map" a texture onto a 3D polygon mesh to create a realistic object. See also "UV map".

Uncanny valley: After a hypothetical modulation by Mori (1970) that was original developed for robots, the uncanny valley predicts that acceptance of an artificial stimulus increases with increased degree of realism until this graph suddenly declines very steeply (into the uncanny valley) when the stimulus reaches a point of almost, but not perfectly, realistic appearance. The uncanny valley then results in rejection of the artificial stimulus.

UV map: a flat plane generated by "unfolding" the 3D model into single parts of connecting polygons, like unfolding a cube. This 2D plane is described by specific "U" and "V" coordinates (called UV, because X, Y, and Z are used to describe the axes of the original 3D object). In some cases, UV maps are created automatically within the animation software, while in other cases they can be manually created according to specific needs. UV maps are used to assign textures to a model.

Virtual animal/stimulus: a CA of an animal/stimulus designed to simulate an artificial counterpart (hetero/conspecific, rival, predator) toward a live test animal.

Virtual reality (VR): CAs of stimuli and/or environments that are rendered in real-time in response to the behavior of the observer. The real-time responsiveness of VR may include behavioral responses to specific actions as well as perspective-correct adjustments of viewpoint, changes in viewing angle of the stimulus while the observer is moving. The first allows for true communication between the observer and the virtual stimulus, and the second means that the observer and the virtual stimulus share the same space. VR hence simulates physical presence of the observer in the virtual environment.

"realistic movement patterns", "realistic depth queues" through perspective correctness, and other visual features the experimenter believes salient for their animal. Information on evaluating this correctness is found in the "validation" section. Information on software to create stimuli can be found in the Supplementary Materials (Supplementary Table S2).

Depending on the aims of the study, the animated stimulus can be created as a CA in 2D, 3D, or as a VR. In VR, but not CA, a correct perspective is maintained even when the subject animal moves. While both **2D** and **3D animation** can simulate movement of **virtual animals**, the visual perspective of the simulation as seen by the watching animal (the observer), will not necessarily be correct. Indeed, a pre-rendered animation is only perspective-correct for one observer position and will appear incorrect to the observer when at other positions. However, the degree of this difference, and whether it is of consequence to the behavior being studied, depends on the animal, the question, and the testing setup. For example, the difference in perspective (between the actual observer position and the position for which the animation was rendered) is small when the distance between the virtual and observing animal is large.

A **3D animation** would be particularly useful, although not necessarily required, when presenting spatially complex visual displays involving movement in the third dimension (front–back axis), such as courtship behavior (Künzler and Bakker 1998), or when questions about spatial features are addressed (Peters and Evans 2007). However, complex 3D animations might not be required in all cases. Indeed, **2D animations** might be sufficient for testing conspecific mate preference (Fischer et al. 2014), conspecific grouping (Qin et al. 2014), or social responses to morphological parameters (e.g., fin size; Baldauf et al. 2010). VRs can be used to track an animal's movements in this environment (Peckmezian and Taylor 2015) and are thus particularly useful for investigating spatial cognition (Dolins et al. in preparation; Thurley and Ayaz 2017).

CAs can be **pre-rendered** or **real-time rendered**. In a **pre-rendered** animation, each frame of the animation is exported from the software before the experiment, and joined to produce an animated video file that can be played during experiments. Hence, motion and behavior of the virtual animal are pre-defined and non-interactive. **Real-time rendered** animation allows the motion and behavior of the virtual animal to be determined in real-time during the experiment, by receiving feedback from a tracking software or, input given by controllers or sensors manipulated by the test animal or the experimenter. Real-time rendering is one requirement for VR, as the viewpoint of the test animal depends on changes in head and/or body position (Thurley et al. 2014; Peckmezian and Taylor 2015) or by input given to a joystick (Dolins et al. 2014).

2D techniques

2D animations present a virtual animal stimulus that only moves up and down (y-axis), and/or left and right (x-axis) on the screen. Such animations are limited in how well they can simulate motion of the virtual animal between the front and back of the scene (z-axis), or correctly simulate the animated animals' rotation. 2D animations might be sufficient when motion is absent or confined to a 2D plane and orientation or perspective can be neglected. 2D animations are less complex to create, and are particularly appropriate when the stimulus being simulated is in itself spatially and temporally simple. Simple 2D animations can be created using digitally modified photographs of animals assembled to construct movement via defined movement paths. In a number of recent studies, test animals were presented left-right/up-down swimming movements of 2D fish

stimuli animated from images in Microsoft PowerPoint (Amcoff et al. 2013; Fischer et al. 2014; Levy et al. 2014; Balzarini et al. 2017), or other image editing software such as GIMP (Baldauf et al. 2011; see Supplementary Table S2). **Keyframed** animation, which has been applied to the study of courtship and agonistic behavior of jumping spiders *Lyssomanes viridis*, can be created using the program Adobe After Effects (Tedore and Johnsen 2013, 2015). Details on image editing software are available in Supplementary Table S2.

3D techniques

Producing **3D animations** requires more sophisticated software, but offers the flexibility to create stimuli that move freely in 3 dimensions (x-, y-, and z-axis). Movement patterns in 3D may appear more realistic to subjects than with only 2 dimensions. Moreover, even though 3D animations can be drawn with some 2D editing programs, 3D animation software offers special graphical interfaces to deal with the 3D geometry of objects, making it easier to present different angles of a stimulus and postural changes than in 2D software (note that most 3D software can also produce 2D animations; see more details in Supplementary Table S2). With 3D animations, it might also be possible to simulate interactions with animated objects or between several animated stimuli (such as animated animal groups) more realistically than in 2D animations, especially if these interactions involve postural changes in depth. Animations in 3D are thus particularly useful to portray complex physical movement patterns of animals (Künzler and Bakker 1998; Watanabe and Troje 2006; Wong and Rosenthal 2006; Parr et al. 2008; Van Dyk and Evans 2008; Campbell et al. 2009; Nelson et al. 2010; Woo and Rieucau 2012, 2015; Gierszewski et al. 2017; Müller et al. 2017). Similar to 2D animation, 3D techniques vary in complexity. All 3D animations require the stimulus to be created as a **mesh** in 3 dimensions, featuring a **rig** and **texture** (see the "shape" section). Various software is available to create a 3D stimulus (see Supplementary Table S2 and Box S1). Once the 3D stimulus is created, it has to be animated, that is, movement patterns have to be assigned to the stimulus, to finalize the animation (discussed in the "motion" section).

VR techniques

VR is the presentation of visual stimuli that are perspective-correct from the observer's body orientation and position, and which continuously adjust to changes in body position in real-time. This continuous rendering allows more realistic depth cues and movement in 3D. Usually, VRs are used to simulate virtual environments and can facilitate investigations of spatial cognition, for example, navigation (Dolins et al. 2014; Thurley et al. 2014; Peckmezian and Taylor 2015; Dolins et al. in preparation; Thurley and Ayaz 2017). In these contexts, VR also allows for greater ecological validity than some traditional experimental methods, allowing natural environments to be realistically mimicked while also allowing experimental manipulation.

Components of an animated stimulus

An animated stimulus is assembled by 3 components: its shape, its texture, and its motion. Additionally, the scene in which it is presented is important as well. These principles apply equally to the creation of animated animals to be used in CA or in VR.

Shape

The shape is defined by the outer margins of the animated stimulus. This is the contour in 2D animations. For 3D animations, there are 3 distinct ways to create the virtual animal's shape, which is made of a mesh and subsequently rigged with a skeleton. The first, simplest, and most popular method is to shape the mesh based on high-quality digital photographs of the subject taken from different angles. These pictures are imported into the animation software and the mesh is built by "re-drawing" the body shape in 3D. The second method is to use landmark-based **geometric morphometric** measurements taken on a picture to define and shape the virtual animal model. This requires slightly more skill and sophisticated software. For fish it is relatively easy to do in software such as *anyFish* 2.0 (Veen et al. 2013; Ingley et al. 2015; see Supplementary Box S1 and Table S2). A third and probably most accurate method is to construct an exact replica by using computed tomography (CT) scans or 3D laser scans of anesthetized or preserved animals, or digitized scans of sections of animals. For example, Woo (2007) used a high-quality 3D laser scanner to construct a precise virtual model of a taxidermic jacky dragon lizard *Amphibolurus muricatus*, while Künzler and Bakker (1998) used digitized scans of sections of the three-spined stickleback *Gasterosteus aculeatus*. Although this approach has the advantage of creating an extremely realistic virtual animal, it requires either a scanner of high quality or technical abilities for sectioning. This method allows precise reconstruction of existing individuals, and it is not clear whether recreating the exact shape to such an extent is essential for eliciting a proper behavioral response (Nakayasu and Watanabe 2014). For research questions where displaying a specific individual is of interest, such as whether chimpanzees *Pan troglodytes* yawn contagiously more to familiar individuals (Campbell et al. 2009), 3D laser scans and CT scans could be particularly beneficial.

Texture

After determining the general shape of the virtual animal, its appearance has to be finalized by adding a **texture** or surface pattern to the object. Strictly speaking, the texture refers to the structure and coloring given to a 3D mesh. The most common and simplest technique is to use a high-quality photograph of an exemplar animal as a texture. Therefore, texturing describes the process of mapping a 2D photograph onto a 3D surface. For texturing, the 3D model is represented by a flat plane, referred to as a **UV map**, which is generated by "unfolding" the 3D model into 2D single parts of connecting polygons described by specific coordinates, like unfolding a cube. The UV maps can be either created automatically or manually according to specific needs. The 2D photograph is then placed onto the map, for example, by using image editing programs, to match the coordinates given in the UV map. The applied texture will then appear on the 3D model after rendering. With this technique, a body can be textured as a whole, or single body parts can be texture-mapped separately, which is more accurate and typically most appropriate for more complex morphologies. Additionally, certain body parts can be defined to be more rigid than others or altered in transparency or surface specifications to resemble muscles, fins, feathers, hair, or other biological structures.

Textures can be manipulated in a variety of ways using common image editing software (e.g., Adobe Photoshop, GIMP). Human-specific measures of color (hue/saturation) of the whole animal or of only a single body part can be altered [e.g., throat color in three-spined sticklebacks in Hiermes et al. (2016); darkness of operculum stripes of *Neolamprologus pulcher* in Balzarini et al. (2017)]. It is also possible to paint different colored markings directly onto the virtual animal's mesh by hand using a pen tool (e.g., egg spot in *Pseudocrenilabrus multicolor* in Egger et al. 2011), or to relocate and transfer specific markings or body parts from one texture onto another. For example, the "clone stamp" tool in Photoshop was used to test for the relevance of vertical bars for mate choice *Xiphophorus cortezi* (Robinson and Morris 2010). Finally, textures from body parts of different individuals can be used to create combinations of traits and phenotypes that would not be found in nature.

Locomotion and behavior

A major advantage of using animated stimuli in animal behavior experiments is the ability to manipulate movement and behavior in a standardized way, which is only marginally possible in live or videotaped animals. Evidence is mounting that spatiotemporal patterns are often important for visual recognition. In fishes, attention toward animated stimuli relies greatly on the movement of the stimulus (Baldauf et al. 2009; Egger et al. 2011). For instance, movement that closely mimicked an animal's natural behavior, referred to as "biological motion" [see Johansson (1973) for an analysis model], elicited a closer association to focal fish (Abaid et al. 2012; Nakayasu and Watanabe 2014). Biologically relevant movement has been found to increase association time of the test animal with animated stimuli regardless of the shape of the stimulus presented in reptiles and fish (Woo and Rieucau 2015; Gierszewski et al. 2017). In addition to the movement of the animated stimulus through the active space, the correct syntax of behavior is important for signal recognition in the jacky dragon (Woo and Rieucau 2015). Motion of the objects includes patterns related to physical movement and displacement in the scene. For example, a bird's movement when flying would involve parameters related to wing flapping, the displacement of the bird through the space, and some breathing motion of the body. We suggest looping some of these patterns, for example the wings flapping, throughout the animation.

Regardless of whether one uses video playbacks, 2D, 3D animation, or VR, the illusion of movement is created by presenting a series of still images of the stimulus at different positions. The number of **frames per second** of an animation needs to be adjusted depending on the tested species' abilities for motion perception, to ensure that the focal animal does not perceive single frames but a smooth motion. Hence, an animation displayed at a low **frame rate** (lower than the species-specific minimum fps needed for continuous motion perception) will be perceived as juddering or in the worst case as discontinuous series of single images. The higher the frame rate and the smaller the distance, the smoother a certain movement or behavioral pattern can be displayed; hence, fast motion needs more fps than slower motion. For humans, an approximate minimum frame rate of 15 fps is sufficient for continuous motion perception, and in cinematic movies 24 fps are common (Baldauf et al. 2008). Typical fps are derived from industry standards encoding formats such as NTSC or PAL, and a widely used and validated frame rate for fish is 30 fps, although species and individuals vary within each family. Higher frame rates might be needed for other animals that possess highly sensitive motion perception, such as some birds (see the "monitor parameter" section for additional information). Indeed, Ware et al. (2015) recently showed that male pigeons *Columba livia* significantly increased courtship duration toward a video of the opposite sex when the frame rate increased from 15 to 30, and from 30 to 60 fps. Moreover, the tested animals only tailored the courtship behavior to the individual presented in the 60 fps condition,

suggesting that with each frame rate different behavioral cues may have been available to be assessed by the tested pigeons. The frame rate has to be determined before creating the motion of the animal, as an animation containing 120 frames will last 2 s when rendered at 60 fps and last 4 s when rendered at 30 fps, and speed of the moving animal will be twice as fast in the former than in the later.

The first and perhaps the technically simplest way to encode motion is to **keyframe** the position and posture of the object every few frames and let the software **interpolate** the position of the object between keyframes. Most 3D animation software provide the possibility to assign a virtual skeleton to the 3D model. Moving the skeleton results in naturalistic postural changes of the different body parts that can be tweaked by the experimenter (see Müller et al. 2017). Generally, the higher the adjusted frame rate the smoother the interpolated movement between 2 set keyframes.

A second way for the experimenter to control movement of an animated animal through a space is the "video game" method, in which the experimenter controls the movement through the use of a game remote controller (Gierszewski et al. 2017; Müller et al. 2017). In this case, some behavioral patterns could also be looped or defined by rules (turn tail when turning left). This method requires the use of a **game engine** (Supplementary Table S2 for examples of game engines) or custom-made software. The movement can be recorded for later display, but this method could also be used for real-time rendering if the experimenter steers the animated animal during the trial.

The third method presented is similar to the "video game" method and only applies to real-time rendering. It is rarely used, but offers the most opportunities for future research. The position or behavior of the animated stimulus is determined based on input from software tracking the test animal (Butkowski et al. 2011; Müller et al. 2016). With this approach, real-time rendered (following the live subject's position) and pre-rendered (showing a specific action, e.g., fin raising) animations can also be combined in a way such that the test animal's actions trigger sequences taken from a library of pre-defined behavioral patterns. Here, as well as with the "video game" method, the virtual animal's movement through the virtual space is realized using algorithms comprised of defined movement rules that account for, for example, object collision and movement speed. For example, Smielik et al. (2015) analyzed fish movement from videos using polynomial interpolation to develop an algorithm which transfers the movement onto a virtual 3D fish skeleton (Gierszewski et al. 2017; Müller et al. 2017). Such algorithms can also be used to let software determine a movement path independent from external input (Abaid et al. 2012).

A fourth way to specify an animal's motion is through a **rotoscoping** method, where the movement of a real animal is assigned to a virtual one (Rosenthal and Ryan 2005). This method allows the investigation of specific behavior, and could be useful for investigating individual differences. For movement through space, the path can be extrapolated from tracking the movement of a live animal, or from a video. This can be automated for some animals using tracking software. If multiple points are consistently tracked over time (Nakayasu and Watanabe 2014), their position can be used to map the path and postural changes onto the animated object. Similarly, a live animal's movement can be recorded using optical motion capture where sensors are directly placed onto a behaving real animal's body and movement patterns are captured in 3D (Watanabe and Troje 2006).

Any of these techniques can produce moving stimuli that can potentially be used in behavioral experiments. To verify whether the

movement generated in the animation corresponds to the live animal on which it is based, optic flow analyses (analysis of image motion) can be performed, as described by Woo and Rieucau (2008) and New and Peters (2010). Optic flow analyses validate generated movements by comparing motion characteristics of the animation, particularly characteristics of velocity and acceleration, to movement patterns gained from videos taken from a behaving live animal. This method is particularly useful when used to verify animal visual displays.

The scene: background and light

In addition to the virtual animal, computer animators must create the surrounding environment referred to as the scene. In the case of 2D animations, most commonly a single background color is used or stimuli are animated on a background image taken from the natural environment (e.g., coral reef in Levy et al. 2014). It is advisable to test the effect of the background color on the test animal, as some colors might produce behavioral changes (e.g., sailfin mollies *Poecilia latipinna* avoid screens with a black background; Witte K, personal communication). Whether the animation is produced in 2D or 3D on a computer screen, both animation styles are represented on a 2D surface. However, there are possibilities to make the 3-dimensional effect more obvious (Zeil 2000) such as creating an environment with reference objects for size and for depth perception (e.g., plants: Baldauf et al. 2009; pictures of artificial structures known by the test animals: Künzler and Bakker 1998; Zbinden et al. 2004; Mehlis et al. 2008). Depth and size cues in an animation might be provided by illusionary effects (e.g., occlusion of objects and texture gradients), since various animals have been shown to respond to visual illusions in a manner similar to humans (Nieder 2002). All standard animation software provides different options for light sources that can be placed to illuminate the virtual environment. Usually, there are options to change number, angle, position, filtering, color, and intensity of the light source so it might be possible to simulate illumination as found in natural environments (e.g., the flickering of underwater light, diffuse scattering of light, or reflection). Illuminating a scene is also a prerequisite for adding realistic shadows to improve the illusion of 3D space (see Gierszewski et al. 2017), a feature also implemented in *anyFish* 2.0 (Ingley et al. 2015; see Supplementary Box S1 and Table S2).

Combined traits and multimodal stimuli

CAs enable controlled testing of combinations of traits and their effect on behavior in numerous contexts. For example, Künzler and Bakker (2001) studied mate choice in sticklebacks and presented live females with virtual males differing in throat coloration, courtship intensity, body size, or a combination of these traits. Live females showed a stronger preference when more traits were available to assess the quality of virtual males. Tedore and Johnsen (2013) created different 2D stimuli of jumping spiders that were comprised of combinations of face and leg textures taken from different sexes or different species to investigate visual recognition in the jumping spider *L. viridis*. To widen the scope of research applications even further, we emphasize that it is also possible to present a visual stimulus together with a cue derived from a different modality to investigate interactions of multimodal stimuli (see also Figure 1). It is possible to add auditory signals (Partan et al. 2005), olfactory cues (Tedore and Johnsen 2013), or even tactile information such as vibrations used by some spiders for communication (Uetz et al. 2015; Kozak and Uetz 2016), or lower frequency vibrations detected by

the lateral line of fish (Blaxter 1987). These cues can be altered and presented either in accordance with, or in contrast to, the visual input (see e.g., Kozak and Uetz 2016). Hence, the effect of multimodal signals and priority levels of ornaments and cues for decision making can be tested in a more standardized and controlled way than would be possible with stimuli coming from live subjects. For example, water containing olfactory cues has been employed during the presentation of virtual fish to investigate kin discrimination and kin selection in three-spined sticklebacks (Mehlis et al. 2008) and in the cichlid *Pelvicachromis taeniatus* (Thünken et al. 2014), while Tedore and Johnsen (2013) investigated male spiders' responses to virtual females with or without the presence of female pheromones in *L. viridis*.

Experimenters have to carefully consider the spatial arrangement and temporal order of presentation of stimuli if multiple cues are combined for testing, as the synchronicity of different cues can greatly affect the perception of and response to such stimuli. Kozak and Uetz (2016) combined video playbacks of male *Schizocosa ocreata* spider courtship behavior with corresponding vibratory signals to test cross-modal integration of multimodal courtship signals. They varied spatial location and temporal synchrony of both signal components and found that females responded to signals that were spatially separated by >90° as if they originated from 2 different sources. Furthermore, females responded more to male signals if visual and tactile information was presented in temporal synchrony rather than asynchronously.

Creating VR

The term "virtual reality" was first applied to a specific set of computer graphics techniques and hardware developed and popularized in the 1980s. These early systems modulated projection of the environment based on body movements of the video game player (Krueger 1991). Today there are many types of VR systems available, some for example requiring the user to wear head mounted display goggles, while others use projectors and tracking to allow the user to move freely. In order to create an immersive experience, all VR systems share 2 common criteria with the original; the display responds to the behavior of the observer (so-called **closed-loop**), and the visual stimulus presented is perspective-correct.

Subsequently, creating a VR requires the support of software to generate perspective-correct representations of 3D objects on a 2D display (projector screen or monitor), and the hardware support for tracking a moving observer in the virtual environment using custom-made or commercial tracking systems (Supplementary Table S2). The necessity of an immediate update of the simulated environment in response to the behavior of the observer makes VR systems more technically challenging than **open-loop** CA.

Most VR setups feature display screens that cover a substantial part of the test animal's field of view (e.g., panoramic or hemispherical) on which the computer generated virtual environment is presented, often using a projector. VR setups usually need an apparatus to mount the live animal in front of the screen, and a tracking system to relay the changes in position and orientation to the system (Fry et al. 2008; Stowers et al. 2014). Animals may be immobilized (Gray et al. 2002) or partly restricted, but should be able to move or change their orientation and position (Thurley et al. 2014; Peckmezian and Taylor 2015). There are multiple techniques available, such as treadmills, to track the movement of animals through a virtual space, and the efficiency of such techniques may vary depending upon the species tested (see Supplementary Table S1 for a list of examples). Creating virtual environments can be done using

common 3D modeling software (Supplementary Table S2), but their integration into a complete VR setup can be complex. Therefore, we only briefly discuss this topic to highlight its significance when using virtual stimuli. Further details and discussion may be found in Stowers et al. (2014) who review different VR systems for freely moving and unrestrained animals of several species, in Thurley and Ayaz (2017) who review the use of VR with rodents, as well as Dolins et al. in preparation for a review on VR use with nonhuman primates.

Pseudoreplication

A frequent concern about auditory or visual playback studies (video, 2D/3D animation, VR) is **pseudoreplication**. As proposed by McGregor (2000), using many variations of a stimulus or of motion paths to cover a wide range of phenotypic variation is the most reliable way to solve this problem. Unfortunately, designing various animated stimuli can be time-consuming. Furthermore, it may not be possible to determine how much variation in artificial stimuli and on which phenotypic trait is needed to address the issue of pseudoreplication. Nevertheless, by presenting identical copies of a stimulus exhibiting variation only in the trait of interest (which is exactly the power of CAs), we can clarify that a difference in behavior most probably results from this exact variation of the tested trait (Mazzi et al. 2003). Instead of creating a single replicate of one individual (exemplar-based animation), many researchers design a stimulus that represents an average phenotype based on population data (parameter-based animation; Rosenthal 2000). This can be achieved by displaying a virtual animal with mean values for size measured from several individuals. The software *anyFish*, for example, incorporates the option to use a consensus file of a study population to create a virtual fish on the basis of geometric morphometrics obtained from several individuals (Ingley et al. 2015). Therefore, it is a straightforward extension to create consensus shapes based on different sets of individuals in order to create a variety of stimuli. It is worth noting that the presentation of an averaged individual still measures the response to a single stimulus and could produce strange artifacts when this stimulus is then used to create groups of individuals, as to the test animal it may be very unusual to have a group composed of several identical individuals. The ideal solution is therefore to create at random different models whose parameters fit within the range seen in natural populations to create such groups.

Displaying the Animated Stimulus

When it comes to presenting the animated stimulus, we are confronted with the issue that readily available display technologies are specifically designed for the human visual system, which may differ considerably from animal visual systems. Hence, there are some important considerations that we have to address to ensure that test animals perceive animated stimuli in a manner needed to test a certain hypothesis.

Animal visual systems and their implications for the presentation of animated stimuli

The visual systems of animals rely on the reception of light by different classes of photoreceptors that are each activated by a limited range of wavelengths. Color is encoded by the nervous system as the ratio of stimulation of different photoreceptor classes. Since color is encoded by the relative magnitude of 3 data points for trichromats like humans, a wide range of colors can be simulated in the eyes of

humans by combining various intensities of 3 spectrally distinct (red, green, and blue) phosphors in screens, the **RGB color model**. The range of colors (**gamut**) that can be displayed with accuracy by screens increases with each phosphor's specificity to the cone class it is designed to stimulate. Therefore, the spectral peak of each phosphor is very narrow, and tailored to stimulate human photoreceptors as independently as possible. The default RGB color model of human devices, monitors, cameras, and computers, is not well suited for many animals. While we discuss color more thoroughly below, it is important to note that in non-human animals the number of photoreceptor classes most commonly varies from 2 (dichromats) to 4 (tetrachromats), but that some animals such as the mantis shrimp can have up to 12 (Thoen et al. 2014), and the number of photoreceptors might further be sex-specific (Jacobs et al. 1996). It is also notable that the characteristics of an animal's photoreceptor classes, specifically their peak wavelength sensitivity and the range of wavelengths they respond to, sometimes differ within species (e.g., butterflies; Arikawa et al. 2005) and also differ from those of the human receptors for which our hardware is built.

Color is probably the most contentious aspect of CAs, because many animals have different numbers of photoreceptor classes spanning both the UV and visible spectrum, with spectral sensitivities centered differently to humans. This means that RGB phosphors will often not be aligned with another animal's photoreceptor spectral sensitivity curves in a way such that their different photoreceptor classes can be independently stimulated, and it might not be possible to accurately simulate color even for non-human trichromats. This constraint has led to a broad literature on color representation and solutions for this issue in video playbacks and images, which are also applicable to CA and VR studies (see Fleishman et al. 1998; Fleishman and Endler 2000; Tedore and Johnsen 2017).

Accurately representing color is thus difficult and testing hypotheses on color even more so, especially if the test animal's spectral sensitivities are unknown. In the case where the visual system of the tested species is not well described, one option is to render the stimulus in grayscale. The caveat is that gray rendered by an RGB system may look like different brightness of a certain color for some species, and stimuli might appear in a "redscale", for example, instead of a grayscale. This may become problematic if the perceived color has any relevance for the specific animal. When the animal's spectral sensitivity is unknown and cannot be assumed from related species, and when the research question is not specifically about the effect of color, for simplicity we suggest adjusting RGB values to look as natural as possible to the human eye. In the case where the spectral sensitivity of the organism is well described, or can be estimated from related species, it is sometimes possible to simulate accurate color representation for the study species. Tedore and Johnsen (2017) provide a user-friendly tool that calculates the best-fit RGB values for the background and every specified color patch shown in an animation, for presentation to di-, tri-, or tetrachromats. If the experimenter wishes to test hypotheses on coloration by using stimulus presentation on any display device, then these calculations are essential for the relevance and interpretation of the experimental results.

We also recommend calibrating screens for color every 2 weeks, not only for those who do specifically test for the effects of color, but also for those who are not manipulating color or testing for its effects. This is due to the fact that RGB colors may drift naturally within just 2 weeks. Calibration is important to make sure that the color presented by the screen does not change. Some monitors and operation software come with a built-in colorimeter (3-filtered light measurement device) and calibration software, but this is rare.

Purely software- or web-based calibrations, which the user conducts by eye, are available, but will not produce identical results across calibrations within, and especially not between, display devices. Proper monitor calibration requires a device containing a colorimeter which takes quantitative and repeatable measurements, such as those manufactured by X-Rite or Datacolor. Such devices are relatively inexpensive, with the highest-end models currently costing less than 250 USD.

In humans, most of the light below 400 nm (UV) is blocked by the ocular media (cornea, lens, vitreous fluid; Boettner and Wolter 1962). In contrast, a large number of both vertebrates and invertebrates have ocular media that are transparent in the UV portion of the spectrum, and have photoreceptor spectral sensitivities peaking well below 400 nm (Marshall et al. 1999; Briscoe and Chittka 2001; Hart and Vorobyev 2005). It is still generally impossible to simulate strongly reflective UV patterns using an RGB screen since RGB phosphors do not emit UV light. While this does not necessarily invalidate results, one should keep this limitation in mind when interpreting responses to live animals versus CAs. UV light plays an important role in visual communication in many species (e.g., Lim et al. 2007; Siebeck 2014). Therefore, it is likely that some information will be lost when this spectral channel is excluded which could confound the subjects' responses.

Many animal visual systems have polarization sensitivity (Wehner 2001). This is particularly common in invertebrates, which in terrestrial habitats, use polarized light for celestial navigation or to localize water sources, and in underwater habitats, to detect open water or enhance contrast between the background and unpolarized targets. There is evidence that some vertebrates have polarization sensitivity as well (e.g., anchovies: Novales Flamarique and Hawryshyn 1998; Novales Flamarique and Hárosi 2002). Moreover, some animals have polarized patterns on the body that may be used for the detection or recognition of potential mates, or possibly in mate choice (e.g., stomatopod crustaceans, Chiou et al. 2008; cuttlefish, Shashar et al. 1996). The extent to which polarized patterns of reflectance on the body have evolved as signals to communicate information to receivers in the animal kingdom is poorly known. Like UV light, this cue is difficult to control and manipulate in CAs or VR. Unfortunately, the most common type of display on the market, the **LCD** display, is highly polarized. Below, under Display parameters, we discuss alternatives to LCD displays if polarized light is a concern for the species under study. Please also see the section "emission of polarized light".

Display parameters

Since animal visual systems are extremely variable, decisions on monitor parameters must be determined depending on species-specific priorities. We can use various types of display devices (e.g., TV screens, computer monitors, notebooks, tablet PCs, smartphones, projectors) to present animated stimuli. Projectors can be used for the presentation of animations (Harland and Jackson 2002) and they are usually used for the presentation of virtual stimuli and environments in VR setups (Thurley and Ayaz 2017).

Display technologies are rapidly evolving and numerous characteristics must be considered and balanced in choosing an appropriate display method. Display characteristics to consider include temporal resolution and flicker (e.g., refresh rate, response time, backlight flicker), spatial resolution, color representation and calibration, brightness, display size, viewing angle, screen polarization, screen reflectance, active versus passive displays, and compatibility with different computer interfaces, as well as practical

considerations such as cost, weight, robustness, and continued availability. Trade-offs between display characteristics are common (see Baldauf et al. 2008). For example, high temporal resolution may mean compromises in terms of spatial resolution or color representation. Moreover, commercially available displays and software designed for humans may not be optimized for accuracy of representation but instead other characteristics, such as reducing user fatigue or enhancing object visibility.

For presentation, we highly recommend presenting a life-sized replica of the animated stimulus to enhance realism for the test animal. The choice for a particular device might be influenced by this consideration. For example, a 72-inch monitor was required to present chimpanzees a realistic image size (Dolins et al. 2014) while Harland and Jackson (2002) used an array of different lenses and filters to ensure a life-sized (± 0.1 mm) rear projection of a small 3D jumping spider *Jacksonoides queenslandicus*.

Display devices typically derive from 2 distinct designs, **CRT** or **LCD** screens. Although authors (e.g., Baldauf et al. 2008) favored in older articles the use of CRTs over LCDs, we reevaluate this preference in light of the advancement of LCD technology and the decreasing availability of CRT screens. Although CRTs were preferred for color display, viewing angle properties, and interpolation, many LCD screens are now built with IPS (in plane switching) that at least decreases viewpoint dependencies regarding luminance and color. The IPS technology produces low deformation of image color with shifting viewing angle, which may be important if the test animal is likely to move during the presentation. However, this will be less important in the case of a short-duration display to a stationary test animal, such as chimpanzees watching videos of yawning (Campbell et al. 2009). **Plasma displays** are also favored when distortion with viewing angle might be a problem (Stewart et al. 2015), although these screens are decreasingly manufactured. Although most monitors are designed to display roughly the same range of color, there are several **LCD monitors** on the market with wide gamut specifications. A wide gamut monitor is able to display colors outside the standard color space of commercial displays. Such a monitor may be useful for experimental tests involving highly saturated colors.

Emission of polarized light

It is important to note that LCD screens and projectors emit polarized light, which may interfere with the test animals' response if the species is sensitive to polarized light. However, this may be neglected if the animal's polarization sensitivity is restricted to dorsally directed photoreceptors, as these receptors are not being stimulated by the RGB display. If polarization sensitivity is a concern, a polarization scattering film can be applied to the display to minimize the issue. Otherwise, plasma screens, CRT monitors, and Digital Light Processing or CRT projectors emit little polarized light. If still unsure what monitor might be best suitable for presentation, it might be worth considering directly comparing different monitor types in an experiment to see if the species under question shows a more reliable response to a certain monitor type (Gierszewski et al. 2017). For each monitor, the temporal and spatial resolution parameters should be examined.

Temporal resolution

The important values that describe temporal resolution are the monitor's refresh rate (in Hz), the animation's frame rate, which is determined at the animation's design stage, and the **latency**.

In earlier studies using CRT monitors, a critical parameter was the monitor's refresh rate. On a CRT screen each pixel is ON only for a short time and then is turned OFF while the cathode ray is serving the other pixels, resulting in flickering of the screen. Flickering was a particular issue with animals whose critical flicker-fusion (**CFF frequency**) values exceeded that of humans or the refresh rate of standard CRT screens (e.g., Railton et al. 2010). An animal's CFF is the threshold frequency at which a blinking light will switch from being perceived as flickering to being perceived as continuous, that is, non-flickering. CFFs are highly variable among species (for a list see Woo et al. 2009 and Healy et al. 2013). Although a high CFF would not affect motion perceptions *per se*, a visible screen flicker may inhibit the perception of smooth motion by masking movement of a stimulus by variations in illumination, like movement seen under strobe light (Ware et al. 2015). The **frame rate** (see the "Locomotion and behavior" section), also called image presentation rate (IPR), is crucial to simulate continuous movement and should be adjusted to exceed the test animal's CSF value (critical sampling frequency; see Watson et al. 1986), the rate needed to render sampled and continuous moving images indistinguishable. CSF is dynamic and may vary with stimulus characteristics and the visual environment (e.g., lighting conditions) during stimulus presentation (Watson et al. 1986; Ware et al. 2015).

In LCD screens, pixels are constantly glowing (although the backlight may flicker at rates of about 150–250 Hz), and therefore the refresh rate-related flickering issue is absent. LCD screens still have a refresh rate that refers to the rate at which it samples the information to be displayed, and is generally around 60 Hz. This means that a screen set at 60 Hz displaying an animation rendered at 30 fps shows each image in the animation twice. As much as possible, hardware and software should be aligned in their temporal characteristics.

Display latency or display lag describe the difference in time between the input of a signal to the display and the time needed for this signal to be shown on the screen. This is particularly important for **closed-loop** applications like VR, as the time between the VR being rendered and it being seen by the animal should be as low as possible. Manufacturers do not always provide accurate information on a display's latency but there are ways to calculate it (see the "measuring display latency" section).

Spatial resolution

Understanding the display properties of screens is important as the trend to use newer yet less standardized devices such as tablet PCs or smartphones for behavioral research increases. The important measures that describe spatial resolution in screen specifications are: screen resolution, pixel density, and pixel spacing. Screen resolution refers to the total number of pixels that are displayed (width × height, e.g., full HD resolution of 1920 × 1080 pixels). Since screens differ in size, pixel density and spacing must also be considered. Pixel (or screen) density describes the number of pixels or dots per linear inch (ppi, dpi), and equals screen width (or height) in pixels divided by screen width (or height) in inches. Low-density screens have fewer ppi than high-density screens, and hence objects displayed on low-density screens appear physically larger than when displayed on high-density screens. Pixel spacing (in mm) describes the distance between neighboring pixels, which is low in high-density screens. The problem of pixelation affects animals when their visual acuity greatly exceeds that of humans. Animals with higher visual acuity than the resolution of the display device will view the stimulus as composed of many square pixels or even

individual red, green, and blue phosphors rather than organically flowing lines and spots. However, even for animals with lower or average visual acuity, pixelation may occur when the subject is positioned too close to the screen. For animals with high visual acuity, or in situations where they are positioned close to the screen (e.g., jumping spiders, Tedore and Johnsen 2013), we recommend the use of high-density devices with the smallest possible pixel spacing. Fleishman and Endler (2000) demonstrated how to calculate the minimum distance needed between the experimental animal and the screen to ensure that the animal cannot resolve individual pixels.

Interactivity of the animated stimulus

Currently, the most effective form of interaction between a live animal and a virtual stimulus so far has been the implementation of VR systems in research that enables real-time interaction between the animal and a virtual environment (Dolins et al. 2014, in preparation; Stowers et al. 2014; Thurley and Ayaz 2017).

CA stimuli rarely enable interaction between the animated stimulus and the experimental test subject. Typically, the behavior of an animated animal is predefined and it will not respond to the test animal, which may greatly reduce ethological relevance of the stimulus and thus the validity and interpretability of the experiment and its results. In contrast, VR or real-time rendered animated stimuli enable interaction and are considered a promising advantage for the future. However, they require more complex software or the creation of a user interface that would allow the experimenter to change the animated animal's behavior. Tracking software (see Supplementary Table S2) can provide real-time information on the position of the live test animal in 2D or 3D space, which can then be used to determine the position of the animated stimuli accordingly, based on predetermined rules. Recently, feedback-based interactive approaches were successfully used with robotic fish (Landgraf et al. 2014, 2016), but as far as we know, there have only been a few studies that implemented some degree of interaction in video playback and CA studies. Ord and Evans (2002) used an interactive algorithm for the presentation of different video sequences showing aggressive and appeasement displays of male jacky dragons, depending on behavior of an observing lizard. Sequence presentation was not completely automatic as the experimenter had to indicate the beginning of a display of the live animal by a key press. Butkowski et al. (2011) combined video game technology and the BIOBSERVE tracking software (see Supplementary Table S2) to enable a rule-based interaction between live female swordtail fish *Xiphophorus birchmanni* and an animated male conspecific and heterospecific fish. Animated males were programmed to automatically track the horizontal position of live females and to raise their dorsal fins, an aggressive signal to rival males, depending on the proximity of the female. Müller et al. (2016) developed an easy to handle method for fully automatic real-time 3D tracking of fish, the sailfin molly *P. latipinna*. The system enables interaction, with the virtual fish stimulus following the live test fish and performing courtship behavior, and it was already successfully tested in practice. Ware et al. (2017) also successfully manipulated social interaction in courtship of pigeons.

Real-time rendered animations and VR-specific considerations

To date, the total system **latency** (sometimes called lag) has been identified as a critical measurement of VR performance (see Supplementary Table S3). This would also apply for real-time rendered animations, if the experiment and the test animal require a timely response. In humans this latency has been stated as one of the most important factors limiting the effectiveness of VR, and should be less than 50 ms, with 20 ms as ideal, and < 8 ms being likely imperceptible (MacKenzie and Ware 1993; Ellis et al. 1997; Miller and Bishop 2002).

Measuring display latency

One common and thorough approach to quantify total closed-loop latency is to use the VR apparatus tracking system to measure the position of an object in the real world, and then project a virtual representation of that same object at the previously measured object position. If the real-world object moves at a known and constant velocity then the difference in position between the virtual and real object can be used to estimate the total closed-loop latency (Liang et al. 1991; Swindells et al. 2000). A simplified estimate (lacking the contribution of the tracking system to the total time) can be achieved by having the user manually change the virtual world and filming how long it takes the result to appear (see http://renderingpipeline.com/2013/09/measuring-input-latency/, on how to measure latency with video footage). Both estimates require using an external high speed imaging system, and can therefore sometimes be difficult. Other approaches attempt to use the system to estimate its own latency, by showing specific patterns on the display and recognizing those patterns using the same tracking cameras used to estimate the object position (Swindells et al. 2000).

Commercial display devices such as monitors and projectors will often be the single largest contributor to total closed-loop latency, contributing between 15 and 50 ms in an average system. When selecting a display device for VR systems, it is important to be able to measure display device latency directly, so one may purchase the best performing display device within one's budget. Display latency can be measured using custom hardware, for example, with a video signal input lag tester (www.leobodnar.com/shop). Comparative display latency can be measured using a combination of hardware and software (see websites for information on how to measure latency: tftcentral.co.uk, tft.vanity.dk). It is important to always measure display latency as the numbers provided by the manufacturer may only refer to the lag of the display panel itself, and may not include the additional signal or image processing induced lag, which takes place in the display or projector electronics. Any processing done by the monitor or projector, such as built in color correction, contrast enhancement, or scaling of the image should be disabled, as such processing takes time and thus increases display latency. If unsure or unable to measure, displays will often have a "game mode" which should generally have the lowest latency.

Validating the Animated Stimulus and VR

After creating the animated stimulus and deciding how to display it, it should be validated thoroughly for at least every test species since perception and recognition might be species and even individual specific (see also Powell and Rosenthal 2017).

A widely used first validation test is to compare the attention given to the CA and an empty counterpart (e.g., blank background image) by presenting them simultaneously. This effectively tests whether the animation is generally perceived by the test animal and whether it attracts attention, but does not determine whether the subjects perceive the animation as intended. In several studies using a 2-choice paradigm, poeciliid fish were attracted to an animated fish and preferred to spend time with the animated fish over the empty background (Morris et al. 2003; Culumber and Rosenthal

2013; Gierszewski et al. 2017). Zebrafish *Danio rerio* significantly reduced their distance to the screen when presented with an animated fish shoal versus a blank background (Pather and Gerlai 2009).

To validate that the subjects perceive the animation similarly to real animals, comparing behavior of test animals when presented with a CA, live animal, and video playback is a crucial step in the validation of a stimulus, and is becoming a gold standard (Clark and Stephenson 1999; Qin et al. 2014; Gierszewski et al. 2017). Fischer et al. (2014) describe a detailed validation of a CA to test visual communication in the cichlid *N. pulcher*.

Another approach to validate animated stimuli is to perform a classical conditioning experiment, which can reveal if test animals are able to perceive and discriminate between specific features of the animated stimuli. This approach might particularly be useful to test a preference for a certain trait that is represented by only subtle changes in morphology or differences in texture. Here, test animals are provided with the opportunity to learn to associate, for example, food with an animated stimulus during a learning phase. Afterwards, the test animals have to discriminate between the learned stimulus and another stimulus (differing in the expression of a trait or the texture) in a binary choice experiment. Such a conditioning experiment was performed successfully to investigate whether sailfin molly females perceive an artificial yellow sword attached to live males on videos presented on TV monitors in mate choice experiments (Witte and Klink 2005).

Although conducting the above tests is generally sufficient to validate an animation, one could additionally compare the performance of CAs generated with different methods (2D, 3D, VR) or containing different characteristics. As the knowledge of which visual characteristics are essential for recognition does not exist a priori for many species, testing which visual features can be simplified by comparing animations with different levels of complexity would provide a more detailed understanding of the communication and recognition systems.

Depending on the research question, control tests confirming different discrimination abilities should follow to complete the validation process. It might be required to show that test animals successfully discriminate between animated conspecifics differing in size, age, sex, familiarity, etc., and that they distinguish between animated heterospecifics or also predators. Using this method, Gerlai et al. (2009) confirmed that the animated image of a predator could be used to elicit a significant stress response in zebrafish. Fischer et al. (2014) demonstrated that cichlids were able to gain information from presented animated conspecifics, heterospecifics, and predators by adjusting their aggression and aversive behavior accordingly.

To validate the significance of results obtained with CA methods, one can perform identical experiments with animated and live animals. Early studies in guppies found that video playbacks of the stimulus and the use of live stimuli yielded similar responses while the former increased efficiency by removing temporal variation (Kodric-Brown and Nicoletto 1997). Amcoff et al. (2013) trained female swordtail characins *Corynopoma riisei* on red and green food items to induce a preference for the same colored male ornament. This preference was demonstrated for live and 2D computer animated fish. However, since in many cases CAs are employed to specifically investigate variation in traits or behavior that are hard or impossible to reproduce with live animals, this approach might not be practical for many studies. Therefore, following the above described validation tests of comparing an animated stimulus to a blank counterpart, live animals, and video playbacks should be considered sufficient to validate usage of CAs. Depending on the research question, the ability to discriminate between sex and/or species, and/or discrimination of different sizes (e.g., of animal size, trait size) should additionally be investigated prior to testing.

Validation of VR

Once the VR system is developed, it is possible to validate its performance against its real-world equivalent by testing if the test animals respond to the virtual environment as if it were real. This necessarily requires finding a strong behavioral response in the real world that can be recreated in VR. Once such a behavior is found, one approach to validation is, for example, to parametrically manipulate certain aspects of the VR system, including the system latency, thus presumably changing the system's realism and the subject's responset. As the latency approaches zero (real-world), the difference in the behavior under question elicited in the experimental subject in the VR versus the real-world context should approach zero. If the difference between VR and real-world at minimum latency is already sufficiently small then one can argue that, by this behavioral definition, the VR system is accurately simulating the real world. Relatedly, the tracking method and suitability of the VR setup should be carefully investigated and fine-tuned to the focal species [see Bohil et al. (2011) for more details on VR]. Especially as some tracking methods may require partial immobilization of animals (Thurley et al. 2014). Beyond ethical issues, such manipulations can lead to decreased ecological validity from the test animal, and decreased realism regarding the VR. Whether a particular tracking method is appropriate for the species tested should therefore also be validated (examples for species already used in Supplementary Table S1).

When interpreting validation results for VR it should be kept in mind that the immersion into highly realistic, yet imperfect, virtual environments might frighten or alarm non-human animals. Here, it is also important to consider the **uncanny valley** phenomenon, described in humans by Seyama and Nagayama (2007). For non-human animals, this phenomenon has so far only been described in long-tailed macaques *Macaca fascicularis*. Steckenfinger and Ghazanfar (2009) found that long-tailed macaques preferred to look at unrealistic synthetic monkey faces as well as real monkey faces, when compared with realistic synthetic monkey faces. Implications of the uncanny valley for other non-human animals can currently only be guessed at and should be the subject of future research (see also Alicea 2015). In general, considerations regarding the uncanny valley phenomenon can be transferred to any artificial and hence virtual stimulus (CA and VR) that is designed to be highly realistic.

If a validation is negative, for example, that the behavior of a test animal is not congruent to that found in nature, every parameter that was prior set to a CA or VR has to be evaluated and if needed to be adjusted until the validation leads to a positive result (see Figure 1). This might especially be the case when a species is tested with CA or VR for the first time and nothing is known on how and if the animal will respond to a newly created virtual stimulus.

Conclusion and Future Directions

CA and VR are useful and promising methods for studying animal behavior. That said, regardless of their potential, virtual stimuli may not be the ideal choice for all research questions involving visual stimuli. Even with external resources and support, creating CAs and VRs can be extremely time-consuming. This investment will be

particularly worthwhile if CAs and VRs enable a set of studies, if methods and stimuli can be reused by multiples researchers or laboratory, or if the required stimulus is hard to obtain using live animals. Moreover, the results obtained using CAs and VRs may not be generalizable to real-world situations, as they typically only present the visual modality (but see the "multimodal stimuli" section). For example, one cichlid species' response to mirrors may not be indicative of aggressiveness in all species (Desjardins and Fernald 2010; Balzarini et al. 2014), a consideration that might apply to CA and VR as well. We would, thus, advocate care in the interpretation of findings until further ecological validations are conducted. Their implementation, the degree of realism required, and the choice between CA and VR, will depend on both technical and conceptual considerations. Systematic experimental analysis of animal behavior will be required to determine whether stimuli are ethologically relevant, and which method of presentation is required in the context of the research question. For many research questions, relatively simple stimuli and setups may be all that are needed, but this remains an empirical question.

CAs have, up to now, been used more often than VRs in animal behavior studies (see Supplementary Table S1), partly because of the higher technical demands for implementing a VR setup, and the cost of implementing the movement tracking systems. This preference also reflects the idea that it may not be necessary to employ a VR for all questions. CAs have mostly been used to investigate questions of perception and recognition, as well as aspects of visual communication and signaling, notably the manipulation of individual traits to assess their role in mate choice. In contrast, VR has primarily been used to investigate cognitive mechanisms, especially regarding spatial navigation (see Supplementary Table S1). VR offers valuable opportunities to study how environmental cues are used in navigation, and how navigation is affected by surgical or pharmacological manipulation of neural substrates. VR systems hence represent a promising technique for future neuroscientific research, and questions of navigation, but CAs seem appropriate to answer most questions of communication and signaling. Animated stimuli have been used in all major taxonomic animal groups that rely on visual communication (see Supplementary Table S1), but fish are the group most often tested using CAs, while VRs most often are used with insects or mammals. This may be explained partly by the investment in VR systems for biomedical research in rodents, which further increases the technical knowledge and tools available for implementation.

For the use of CA, future directions should address the issue of non-interactivity as this still represents one of the major limitations when using animated animals. Ongoing improvement in tracking systems, that also function in 3D (e.g., Müller et al. 2014; Straw et al. 2010), may help to create interactive animated stimuli in the future (Müller et al. 2016). So far, animated stimuli have predominantly been used in choice experiments and their possible use in other popular testing paradigms has mostly been neglected. And yet, animated stimuli are also very well suited to be observers, bystanders, or demonstrators in experiments that investigate higher-order aspects of the social interactions of a species (e.g., Witte and Ueding 2003; Makowicz et al. 2010). Regarding VR, the majority of current systems necessitate partial or complete immobilization of the tested animal and this might limit the use of these systems much more than the complexity of the programs needed for implementation, as subjects might not be able to show their full behavioral repertoire. Future directions should hence promote the development of free-ranging VR systems that do not restrict natural behavior.

Even if CA and VR have not yet reached their peak of innovation and accessibility, current technical advances already provide opportunities for sophisticated design and presentation of animated stimuli. Software applications for both beginning and advanced users can be found and the increase of professional freeware (see Supplementary Table S2) also facilitates an inexpensive implementation of animated stimuli in research. Numerous possibilities for creating animated stimuli with varying complexity can be used to address questions concerning visual communication and spatial cognition. Further technical advances are expected, following the increasing popularity of VR in mobile gaming applications, and its use in robotic and remote surgery. Insofar as this affects animal VR one should expect to see market pressures encouraging the sale of low latency display devices. The trends highlighted by the current use of CA and VR in animal behavior research, and the prospect of technical advances imply that a major barrier for increased use of VRs and CAs may reside in the technical hurdles of building and validating a new system. As such, the creation of shareable systems (e.g., *anyFish 2.0*, see Box 1 in the Supplementary Material), open-source or freeware, how-to guides, etc. to assist in building the systems would be invaluable in improving the accessibility to virtual research techniques in the future.

We hope that our review will inspire future research and the continuous development of more advanced techniques that hence lead to novel insights into animal behavior.

Supplementary material

Supplementary material can be found at http://www.cz.oxfordjournals.org/.

References

Abaid N, Spinello C, Laut J, Porfiri M, 2012. Zebrafish *Danio rerio* responds to images animated by mathematical models of animal grouping. *Behav Brain Res* **232**:406–410.

Alicea B, 2015. Animal-oriented virtual environments: illusion, dilation, and discovery [version 2; referees: 2 approved with reservations]. *F1000 Research* **3**:202.

Amcoff M, Lindqvist C, Kolm N, 2013. Sensory exploitation and plasticity in female mate choice in the swordtail characin. *Anim Behav* **85**:891–898.

Arikawa K, Wakakuwa M, Qiu M, Kurasawa M, Stavenga DG, 2005. Sexual dimorphism of short-wavelength photoreceptors in the small white butterfly *Pieris rapae crucivora*. *J Neurosci* **25**:5935–5942.

ASAB, 2014. Guidelines for the treatment of animals in behavioural research and teaching. *Anim Behav* **87**:I–IX.

Baldauf SA, Kullmann H, Bakker TCM, 2008. Technical restrictions of computer-manipulated visual stimuli and display units for studying animal behaviour. *Ethology* 114:737–751.

Baldauf SA, Kullmann H, Thünken T, Winter S, Bakker TCM, 2009. Computer animation as a tool to study preferences in the cichlid *Pelvicachromis taeniatus*. *J Fish Biol* 75:738–746.

Baldauf SA, Bakker TCM, Herder F, Kullmann H, Thunken T, 2010. Male mate choice scales female ornament allometry in a cichlid fish. *BMC Evol Biol* 10:301.

Baldauf SA, Bakker TCM, Kullmann H, Thünken T, 2011. Female nuptial coloration and its adaptive significance in a mutual mate choice system. *Behav Ecol* 22:478–485.

Balzarini V, Taborsky M, Villa F, Frommen JG, 2017. Computer animations of colour markings reveal the function of visual threat signals in *Neolamprologus pulcher*. *Curr Zool* 63:45–54.

Balzarini V, Taborsky M, Wanner S, Koch F, Frommen JG, 2014. Mirror, mirror on the wall: the predictive value of mirror tests for measuring aggression in fish. *Behav Ecol Sociobiol* 68:871–878.

Basolo AL, 1990. Female preference for male sword length in the green swordtail *Xiphophorus helleri* (Pisces: Poeciliidae). *Anim Behav* 40:332–338.

Blaxter JHS, 1987. Structure and development of the lateral line. *Biol Rev* 62:471–514.

Boettner EA, Wolter JR, 1962. Transmission of the ocular media. *Invest Ophthalmol* 1:776–783.

Bohil CJ, Alicea B, Biocca FA, 2011. Virtual reality in neuroscience research and therapy. *Nat Rev Neurosci* 12:752–762.

Briscoe AD, Chittka L, 2001. The evolution of color vision in insects. *Annu Rev Entomol* 46:471–510.

Butkowski T, Yan W, Gray AM, Cui RF, Verzijden MN, 2011. Automated interactive video playback for studies of animal communication. *J Vis Exp* 48:UNSP e2374.

Campbell MW, Carter JD, Proctor D, Eisenberg ML, de Waal FBM, 2009. Computer animations stimulate contagious yawning in chimpanzees. *Proc R Soc B* 276:4255–4259.

Chiou TH, Kleinlogel S, Cronin T, Caldwell R, Loeffler B et al., 2008. Circular polarization vision in a stomatopod crustacean. *Curr Biol* 18:429–434.

Clark DL, Stephenson KR, 1999. Response to video and computer-animated images by the tiger barb. *Environ Biol Fish* 56:317–324.

Culumber ZW, Rosenthal GG, 2013. Mating preferences do not maintain the tailspot polymorphism in the platyfish *Xiphophorus variatus*. *Behav Ecol* 24:1286–1291.

Desjardins JK, Fernald RD, 2010. What do fish make of mirror images?. *Biol Lett* 6:744–747.

Dolins FL, Klimowicz C, Kelley J, Menzel CR, 2014. Using virtual reality to investigate comparative spatial cognitive abilities in chimpanzees and humans. *Am J Primatol* 76:496–513.

Egger B, Klaefiger Y, Theis A, Salzburger W, 2011. A sensory bias has triggered the evolution of egg-spots in cichlid fishes. *PLoS One* 6:e25601.

Ellis SR, Bréant F, Menges B, Jacoby R, Adelstein BD, 1997. Factors influencing operator interaction with virtual objects viewed via head-mounted see-through displays: viewing conditions and rendering latency. Proceedings of the IEEE 1997 Virtual Reality Annual International Symposium. Los Alamitos (CA): IEEE Computer Society Press, 138–145.

Espmark Y, Amundsen T, Rosenqvist G, 2000. *Animal Signals: Signalling and Signal Design in Animal Communication*. Trondheim, Norway: Tapir Academic Press.

Fischer S, Taborsky B, Burlaud R, Fernandez AA, Hess S et al., 2014. Animated images as a tool to study visual communication: a case study in a cooperatively breeding cichlid. *Behaviour* 151:1921–1942.

Fleishman LJ, Endler JA, 2000. Some comments on visual perception and the use of video playback in animal behavior studies. *Acta Ethol* 3:15–27.

Fleishman LJ, McClintock WJ, D'eath RB, Brainard DH, Endler JA, 1998. Colour perception and the use of video playback experiments in animal behaviour. *Anim Behav* 56:1035–1040.

Fry SN, Rohrseitz N, Straw AD, Dickinson MH, 2008. TrackFly: virtual reality for a behavioral system analysis in free-flying fruit flies. *J Neurosci Methods* 171:110–117.

Gerlai R, 2017. Animated images in the analysis of zebrafish behavior. *Curr Zool* 63:35–44.

Gerlai R, Fernandes Y, Pereira T, 2009. Zebrafish *Danio rerio* responds to the animated image of a predator: towards the development of an automated aversive task. *Behav Brain Res* 201:318–324.

Gierszewski S, Müller K, Smielik I, Hütwohl J, Kuhnert K-D et al., 2017. The virtual lover: variable and easily-guided 3D-fish animations as an innovative tool in mate-choice experiments with sailfin mollies. II. Validation. *Curr Zool* 63:65–74.

Gray JR, Pawlowski V, Willis MA, 2002. A method for recording behavior and multineuronal CNS activity from tethered insects flying in virtual space. *J Neurosci Methods* 120:211–223.

Harland DP, Jackson RR, 2002. Influence of cues from the anterior medial eyes of virtual prey on *Portia fimbriata*, an araneophagic jumping spider. *J Exp Biol* 205:1861–1868.

Hart NS, Vorobyev M, 2005. Modelling oil droplet absorption spectra and spectral sensitivities of bird cone photoreceptors. *J Comp Physiol A* 191:381–392.

Healy K, McNally L, Ruxton GD, Cooper N, Jackson A, 2013. Metabolic rate and body size are linked with perception of temporal information. *Anim Behav* 86:685–696.

Hiermes M, Rick IP, Mehlis M, Bakker TCM, 2016. The dynamics of color signals in male threespine sticklebacks *Gasterosteus aculeatus*. *Curr Zool* 62:23–32.

Hurlbert SH, 1984. Pseudoreplication and the design of ecological field experiments. *Ecol Monogr* 54:187–211.

Ingley SJ, Rahmani Asl M, Wu C, Cui R, Gadelhak M et al., 2015. anyFish 2.0: An open-source software platform to generate and share animated fish models to study behavior. *SoftwareX* 1–9.

Ioannou CC, Guttal V, Couzin ID, 2012. Predatory fish select for coordinated collective motion in virtual prey. *Science* 337:1212–1215.

Jacobs GH, Neitz M, Deegan JF, Neitz J, 1996. Trichromatic colour vision in New World monkeys. *Nature* 382:156–158.

Johansson G, 1973. Visual perception of biological motion and a model for its analysis. *Percept Psychophys* 14:201–211.

Kodric-Brown A, Nicoletto PF, 1997. Repeatability of female choice in the guppy: response to live and videotaped males. *Anim Behav* 54:369–376.

Kozak EC, Uetz GW, 2016. Cross-modal integration of multimodal courtship signals in a wolf spider. *Anim Cogn*. doi:10.1007/s10071-016-1025-y.

Krueger MW, 1991. *Artificial Reality*. 2nd edn. Reading, MA: Addison-Wesley.

Künzler R, Bakker TCM, 1998. Computer animations as a tool in the study of mating preferences. *Behaviour* 135:1137–1159.

Künzler R, Bakker TCM, 2001. Female preferences for single and combined traits in computer animated stickleback males. *Behav Ecol* 12:681–685.

Landgraf T, Nguyen H, Schröer J, Szengel A, Clément RJG et al., 2014. Blending in with the shoal: robotic fish swarms for investigating strategies of group formation in guppies. In: Duff A, Lepora NF, Mura A, Prescott TJ, Verschure PFMJ, editors. *Biomimetic and Biohybrid Systems*. Switzerland: Springer, 178–189.

Landgraf T, Bierbach D, Nguyen H, Muggelberg N, Romanczuk P et al., 2016. RoboFish: increased acceptance of interactive robotic fish with realistic eyes and natural motion patterns by live Trinidadian guppies. *Bioinspir Biomim* 11:15001.

Levy K, Lerner A, Shashar N, 2014. Mate choice and body pattern variations in the crown butterfly fish *Chaetodon paucifasciatus* (Chaetodontidae). *Biol Open* 1:1–7.

Liang J, Shaw C, Green M, 1991. On temporal-spatial realism in the virtual reality environment. Proceedings of the 4th Annual ACM Symposium on User Interface Software and Technology (UIST '91). New York (NY): ACM, 19–25.

Lim MLM, Land MF, Li D, 2007. Sex-specific UV and fluorescence signals in jumping spiders. *Science* 315:481.

MacKenzie IS, Ware C, 1993. Lag as a determinant of human performance in interactive systems. Proceedings of the ACM Conference on Human

Factors in Computing Systems—INTERCHI '93. New York (NY): ACM, 488–493

Makowicz AM, Plath M, Schlupp I, 2010. Using video playback to study the effect of an audience on male mating behavior in the sailfin molly *Poecilia latipinna*. *Behav Proc* **85**:36–41.

Marshall J, Kent J, Cronin T, 1999. Visual adaptations in crustaceans: spectral sensitivity in diverse habitats. In: Archer SN, editor. *Adaptive Mechanisms in the Ecology of Vision*. Dordrecht: Kluwer, 285–327.

Mazzi D, Künzler R, Bakker TCM, 2003. Female preference for symmetry in computer-animated three-spined sticklebacks *Gasterosteus aculeatus*. *Behav Ecol Sociobiol* **54**:156–161.

McGregor PK, 2000. Playback experiments: design and analysis. *Acta Ethol* **3**:3–8.

Mehlis M, Bakker TCM, Frommen JG, 2008. Smells like sib spirit: kin recognition in three-spined sticklebacks *Gasterosteus aculeatus* is mediated by olfactory cues. *Anim Cogn* **11**:643–650.

Miller D, Bishop G, 2002. Latency meter: a device end-to-end latency of VE systems. Proceedings of SPIE—Stereoscopic Displays and Virtual Reality Systems IX, 458–464. doi:10.1117/12.468062.

Mori M, 1970. The uncanny valley. *Energy* **7**:33–35 (in Japanese).

Morris MR, Nicoletto PF, Hesselman E, 2003. A polymorphism in female preference for a polymorphic male trait in the swordtail fish *Xiphophorus cortezi*. *Anim Behav* **65**:45–52.

Müller K, Schlemper J, Kuhnert L, Kuhnert KD, 2014. Calibration and 3D ground truth data generation with orthogonal camera-setup and refraction compensation for aquaria in real-time. Proceedings of VISAPP 2014—International Conference of Computer Vision Theory and Applications. Lisbon, Portugal: IEEE Computer Society Press, 626–634.

Müller K, Gierszewski S, Witte K, Kuhnert KD, Forthcoming 2016. Where is my mate?—Real-time 3-D fish tracking for interactive mate-choice experiments. Accepted for ICPR 2016—23rd International Conference on Pattern Recognition; VAIB 2016, Proceedings.

Müller K, Smielik I, Hütwohl JM, Gierszewski S, Witte K et al., 2017. The virtual lover: variable and easily-guided 3D-fish animations as an innovative tool in mate-choice experiments with sailfin mollies. I. Design and implementation. *Curr Zool* **63**:55–64.

Nakayasu T, Watanabe E, 2014. Biological motion stimuli are attractive to medaka fish. *Anim Cogn* **17**:559–575.

Nelson XJ, Garnett DT, Evans CS, 2010. Receiver psychology and the design of the deceptive caudal luring signal of the death adder. *Anim Behav* **79**:555–561.

New STD, Peters RA, 2010. A framework for quantifying properties of 3-dimensional movement-based signals. *Curr Zool* **56**:327–336.

Nieder A, 2002. Seeing more than meets the eye: processing of illusory contours in animals. *J Comp Physiol A* **188**:249–260.

Novales Flamarique I, Hárosi FI, 2002. Visual pigments and dichroism of anchovy cones: a model system for polarization detection. *Vis Neurosci* **19**:467–473.

Novales Flamarique I, Hawryshyn CW, 1998. Photoreceptor types and their relation to the spectral and polarization sensitivities of clupeid fishes. *J Comp Physiol A* **182**:793–803.

Oliveira RF, Rosenthal GG, Schlupp I, Mcgregor PK, Cuthill IC et al., 2000. Considerations on the use of video playbacks as visual stimuli: the Lisbon workshop consensus. *Acta Ethol* **3**:61–65.

Ord TJ, Evans CS, 2002. Interactive video playback and opponent assessment in lizards. *Behav Process* **59**:55–65.

Parr LA, Waller BM, Heintz M, 2008. Facial expression categorization by chimpanzees using standardized stimuli. *Emotion* **8**:216–231.

Partan S, Yelda S, Price V, Shimizu T, 2005. Female pigeons *Columba livia*, respond to multisensory audio/video playbacks of male courtship behaviour. *Anim Behav* **70**:957–966.

Pather S, Gerlai R, 2009. Shuttle box learning in zebrafish *Danio rerio*. *Behav Brain Res* **196**:323–327.

Peckmezian T, Taylor PW, 2015. A virtual reality paradigm for the study of visually mediated behaviour and cognition in spiders and cognition in spiders. *Anim Behav* **107**:87–95.

Peters RA, Evans CS, 2007. Active space of a movement-based signal: response to the Jacky dragon *Amphibolurus muricatus* display is sensitive to distance, but independent of orientation. *J Exp Biol* **210**:395–402.

Powell DL, Rosenthal GG, 2017. What artifice can and cannot tell us about animal behavior. *Curr Zool* **63**:21–26.

Qin M, Wong A, Seguin D, Gerlai R, 2014. Induction of social behavior in zebrafish: live versus computer animated fish as stimuli. *Zebrafish* **11**:185–197.

Railton R, Foster T, Temple W, 2010. Transfer of stimulus control from a TFT to CRT screen. *Behav Process* **85**:111–115.

Robinson DM, Morris MR, 2010. Unraveling the complexities of variation in female mate preference for vertical bars in the swordtail *Xiphophorus cortezi*. *Behav Ecol Sociobiol* **64**:1537–1545.

Rosenthal GG, Evans CS, 1998. Female preference for swords in *Xiphophorus helleri* reflects a bias for large apparent size. *Proc Natl Acad Sci USA* **95**:4431–4436.

Rosenthal GG, 2000. Design considerations and techniques for constructing video stimuli. *Acta Ethol* **3**:49–54.

Rosenthal GG, Ryan MJ, 2005. Assortative preferences for stripes in danios. *Anim Behav* **70**:1063–1066.

Rowland WJ, 1999. Studying visual cues in fish behavior: a review of ethological techniques. *Environ Biol Fish* **56**:285–305.

Russel WMS, Burch RL, 1959. *The Principles of Humane Experimental Technique*. London: Methuen.

Schlupp I, Waschulewski M, Ryan MJ, 1999. Female preferences for naturally-occurring novel male traits. *Behaviour* **136**:519–527.

Seyama J, Nagayama RS, 2007. The Uncanny Valley: effect on realism on the impression of artificial human faces. *Presence* **16**:337–351.

Shashar N, Rutledge PS, Cronin TW, 1996. Polarization vision in cuttlefish: a concealed communication channel?. *J Exp Biol* **199**:2077–2084.

Siebeck UE, 2014. Communication in the ultraviolet: unravelling the secret language of fish. In: Witzany G, editor. *Biocommunication of Animals*. Netherlands: Springer, 299–320.

Smielik I, Müller K, Kuhnert KD, 2015. Fish motion simulation. In: Al-Akaidi M, Ayesh A, editors. *ESM—European Simulation and Modelling Conference 2015*. Leicester: EUROSIS, 392–396.

Steckenfinger SA, Ghazanfar AA, 2009. Monkey visual behavior falls into the uncanny valley. *Proc Natl Acad Sci USA* **106**:18362–18366.

Stewart FJ, Kinoshita M, Arikawa K, 2015. The butterfly *Papilio xuthus* detects visual motion using chromatic contrast. *Biol Lett* **11**:20150687.

Stowers JR, Fuhrmann A, Hofbauer M, Streinzer M, Schmid A et al., 2014. Reverse engineering animal vision with virtual reality and genetics. *Computer* **47**:38–45.

Straw AD, Branson K, Neumann TR, Dickinson MH, 2010. Multi-camera real-time three-dimensional tracking of multiple flying animals. *J R Soc Interface* **8**:395–409.

Swindells C, Dill JC, Booth KS, 2000. System lag tests for augmented and virtual environments. Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology (UIST '00). New York (NY): ACM, 161–170.

Tedore C, Johnsen S, 2013. Pheromones exert top-down effects on visual recognition in the jumping spider *Lyssomanes viridis*. *J Exp Biol* **216**:1744–1756.

Tedore C, Johnsen S, 2015. Visual mutual assessment of size in male *Lyssomanes viridis* jumping spider contests. *Behav Ecol* **26**:510–518.

Tedore C, Johnsen S, 2017. Using RGB displays to portray color realistic imagery to animal eyes. *Curr Zool* **63**:27–34.

Thoen HH, How MJ, Chiou T-H, Marshall J, 2014. A different form of color vision in mantis shrimp. *Science* **343**:411–413.

Thurley K, Henke J, Hermann J, Ludwig B, Tatarau C et al., 2014. Mongolian gerbils learn to navigate in complex virtual spaces. *Behav Brain Res* **266**:161–168.

Thurley K, Ayaz A, 2017. Virtual reality systems for rodents. *Curr Zool* **63**:109–119.

Thünken T, Bakker TCM, Baldauf SA, 2014. "Armpit effect" in an African cichlid fish: self-referent kin recognition in mating decisions of male *Pelvicachromis taeniatus*. *Behav Ecol Sociobiol* **8**:99–104.

Tinbergen N, 1948. Social releasers and the experimental method required for their study. *Wilson Bull* **60**:6–51.

Uetz GW, Clark DL, Stoffer B, Kozak E, Lallo CM et al., 2015. Multimodal communication in wolf spiders: playback studies with visual and vibratory signals. *J Acoust Soc Am* **137**:2395.

Van Dyk DA, Evans CS, 2008. Opponent assessment in lizards: examining the effect of aggressive and submissive signals. *Behav Ecol* **19**:895–901.

Veen T, Ingley SJ, Cui R, Simpson J, Asl MR et al., 2013. anyFish: an open-source software to generate animated fish models for behavioural studies. *Evol Ecol Res* **15**:361–375.

Ware E, Saunders DR, Troje NF, 2015. The influence of motion quality on responses towards video playback stimuli. *Biol Open* **4**:803–811.

Ware ELR, Saunders DR, Troje NF, 2017. Social interactivity in pigeon courtship behavior. *Curr Zool* **63**:85–95.

Watanabe S, Troje NF, 2006. Towards a "virtual pigeon": a new technique for investigating avian social perception. *Anim Cogn* **9**:271–279.

Watson AB, Ahumada Jr AJ, Farrell JE, 1986. Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. *J Opt Soc Am A* **3**:300–307.

Wehner R, 2001. Polarization vision: a uniform sensory capacity?. *J Exp Biol* **204**:2589–2596.

Witte K, Klink KB, 2005. No pre-existing bias in sailfin molly females *Poecilia latipinna* for a sword in males. *Behavior* **142**:283–303.

Witte K, Ueding K, 2003. Sailfin molly females copy the rejection of a male. *Behav Ecol* **14**:389–395.

Wong BBM, Rosenthal GG, 2006. Female disdain for swords in a swordtail fish. *Am Nat* **167**:136–140.

Woo KL, 2007. Computer-generated animal model stimuli. *J Vis Exp* **6**:2.

Woo KL, Hunt M, Harper D, Nelson NJ, Daugherty C et al., 2009. Discrimination of flicker frequency rates in the reptile tuatara (Sphenodon). *Naturwissenschaften* **96**:415–419.

Woo KL, Rieucau G, 2008. Considerations in video playback design: using optic flow analysis to examine motion characteristics of live and computer-generated animation sequences. *Behav Process* **78**:455–463.

Woo KL, Rieucau G, 2012. Aggressive signal design in the jacky dragon *Amphibolurus muricatus*: display duration affects efficiency. *Ethology* **118**:157–168.

Woo KL, Rieucau G, 2015. The importance of syntax in a dynamic visual signal: recognition of jacky dragon displays depends upon sequence. *Acta Ethol* **18**:255–263.

Zbinden M, Largiadèr CR, Bakker TCM, 2004. Body size of virtual rivals affects ejaculate size in sticklebacks. *Behav Ecol* **15**:137–140.

Zeil J, 2000. Depth cues, behavioural context, and natural illumination: some potential limitations of video playback techniques. *Acta Ethol* **3**:39–48.