The Treachery of Images: How Realism Influences Brain and Behavior

Jacqueline C. Snow[1], **Jody C. Culham**[2,3,*]

[1]Department of Psychology, University of Nevada Reno, Reno, Nevada, USA

[2]Department of Psychology, University of Western Ontario, London, Ontario, Canada

[3]Brain and Mind Institute, Western Interdisciplinary Research Building, University of Western Ontario, London, Ontario, N6A 3K7, Canada

Abstract

Although the cognitive sciences aim to ultimately understand behavior and brain function in the real world, for historical and practical reasons, the field has relied heavily on artificial stimuli, typically pictures. We review a growing body of evidence that both behavior and brain function differ between image proxies and real, tangible objects. We also propose a new framework for "immersive neuroscience" to combine two approaches: (1) the traditional "build-up" approach of gradually combining simplified stimuli, tasks, and processes; and (2) a newer "tear-down" approach that begins with reality and compelling simulations such as virtual reality to determine which elements critically affect behavior and brain processing.

Keywords

real-world behavior; immersive neuroscience; real objects; images; actions; virtual reality

The Dependence on Proxies for Realism in Cognitive Sciences

In a famous painting, *The Treachery of Images*, artist René Magritte painted a pipe above the words "Ceci n'est pas une pipe" ("This is not a pipe"). When asked about the painting, he replied, "Could you stuff my pipe? No, i s just a representation, is it not? So if I had written on my picture 'This is a pipe', I'd have been lying!" [1]. Magritte had an insight that applies even to psychology and neuroimaging: a picture of an object is only a limited representation of the real thing.

Despite such intuitions regarding the importance of realism, researchers in the cognitive sciences often employ experimental **proxies** (see Glossary) for reality. One ubiquitous proxy is the use of artificial two-dimensional (2-D) images of stimuli that are not actually

*Correspondence: jculham@uwo.ca (J. C. Culham).

present, evoking indirect perceptions of depicted objects or scenes [2]. Images predominate over real objects in research because they are easy to create, easy to present rapidly with accurate timing on computer displays, and easy to control (e.g., for low-level attributes like brightness). Researchers will often employ the phrase "real-world" to imply that some feature of the image matches some aspect of reality (such as category, image statistics, familiar properties like size, or implied depth), even though the stimuli are not real objects, which we define as physical, tangible 3-D solids (see Box 1).

Here we address whether behavior and brain responses to real objects may differ from images. Though the potential theoretical importance of realism has been recognized for decades [2], there is an emerging body of empirical evidence for such differences. With the advent of human neuroimaging, our understanding has expanded from the earliest stages of visual processing (e.g., V1) to higher-level brain regions involved in recognition and actions in the real world. Correspondingly, our understanding of vision and cognition may need to progress beyond light patterns and images to encompass the tangible objects and scenes that determine everyday behavior.

Why Might Proxies Differ from the Real World?

Before reviewing empirical evidence for differences between proxies and the real world, we note that philosophically, there are reasons to posit such differences.

Realism vs. symbolism

As Magritte highlighted best, while tangible objects satisfy needs, images are representational and often symbolic (e.g., cartoons). Images have a unique 'duality' because they are both physical objects comprised of markings on a surface, and at the same time, they are representations of something else [2–5]. Humans can easily recognize images such as line drawings; however, such representational images are a relatively new invention of humankind, with cave drawings appearing only within the last 45,000 years [6] of the 4–5 million years of hominin evolution. Moreover, pictures were highly schematic (such as outlines with no portrayal of depth) until artists learned how to render cues like perspective during the Renaissance and to use photography and film in the 19th century [7]. With increasingly compelling simulations, there is increased awareness of the potential importance of **presence**, the sense of realism and immersion [8,9].

Actability

Animals evolved to survive in the natural world using brain mechanisms that perform actions guided by sophisticated perceptual and cognitive systems. Modern cognitive neuroscience relies heavily on reductionism, using impoverished stimuli and tasks that neglect the importance of actions as the outcomes upon which survival has depended for millions of years [10]. Despite the importance of action outcomes for evolutionary selection, historically, psychology has neglected motor control [11], even though all cognitive processes, even those that are not explicitly motor (e.g., attention and memory), evolved in ultimate service of affecting motor behavior [12]. Increasingly, researchers are realizing that theories developed in motor control, such as the importance of forward models and

feedback loops [13], may explain many other cognitive functions from perception through social interactions [14]. As such, one must question why so many cognitive studies use stimuli – particularly static images -- that do not afford motor behavior or use tasks no more sophisticated than pressing a button or uttering a verbal response. Images do not afford actual actions. Images may evoke notions of **affordances** [15] and **action associations** [16] but they lack **actability**, the potential to interact with the represented object meaningfully. Put simply, one cannot pound a nail with a photo of a hammer.

In addition, images are purely visual and thus lack genuine cross-modal processing. For example, when determining the ripeness of a peach, color provides some cues to ripeness but ideally this is confirmed by assessing how well it yields to a squeeze, smells fragrant, and tastes sweet and juicy. Multisensory processing not only allows more accurate interpretation of objects in the world, it enables hypotheses generated from one sense to be tested using another. Yet pictures are rarely touched (by adults) and haptic exploration would merely confirm the flatness conveyed by vision.

Motion and depth

Even in terms of vision alone, images are impoverished. Most notably, images never change, although videos can be used to convey animation. In addition, images lack a veridical combination of depth cues. Whereas pictures can accurately convey relative depth relationships (e.g., the ball is in front of the block), they have limited utility in computing absolute distances (e.g., the ball is 40 cm away, within reach, and the block is 90 cm away, beyond reach) and thus inferring real-world size and actability. With images, depth cues may be in conflict: monocular cues (e.g., shading, occlusion) convey three-dimensionality while **stereopsis** and oculomotor cues convey flatness and implausible size-distance relationships. For example, in typical fMRI studies, stimuli may include large landmarks (e.g., the Eiffel tower) and small objects (e.g., coins) that subtend the same retinal angle at the same viewing distance even though their typical sizes differ by orders of magnitude in the real world. Compared to static images, dynamic movies can evoke stronger activation in lateral (but not ventral) occipitotemporal cortex [17–19]. The addition of 3-D cues from stereopsis evokes stronger brain responses and stronger inter-regional connectivity, compared to flat 2D movies [8,20]. 3-D effects are stronger for stimuli that appear near (vs. far) and the effects increase through the visual hierarchy [21]. Motion and depth are also likely to be important for the perception of real objects, although none of these studies actually employed real objects, where depth may be especially important for both action and perception.

Development

Not only are these factors important for object processing in adults, they are fundamental during early development. Infants learn to make sense of their environment largely through the interactions and multisensory integration provided in the tangible world. Young infants fail to understand that images are symbolic and may try to engage with them [22]. Bodily movement and interaction with objects are crucial to child development, as shown by the quantum leaps in cognitive abilities that occur once infants become mobile [23–25]. Indeed, as classic research showed, humans [26] and other animals [27] require self-initiated active exploration of the visual environment for normal development and coordination. Evidence

is mixed as to whether other species comprehend images as representations of reality [28], even though images are common stimuli for neurophysiological research. Realism can also have striking effects on object processing in human infants and young children, as well as in animals (see Box 2).

In sum, because humans have evolved and developed in the real world, our behaviors and brain processes are likely to reflect the important features of tangible objects and environments – tangibility, actability, multisensory processing, and motion and depth.

Empirical Evidence for The Importance of Real Objects

Despite theoretical arguments for the importance of real objects for understanding behavior and brain function, surprisingly few empirical studies have used real-world stimuli or directly compared responses to stimuli in different formats. However, with the development of novel methods for presenting real objects under controlled conditions (see Figure 1), there is growing evidence that real objects elicit different responses compared to images.

Object Recognition

Recognition performance may be better for real objects compared to pictures, an effect known as a **real-object advantage**, in both neuropsychological patients and typical individuals. Although patients with visual agnosia are typically unable to recognize 2-D pictures of everyday objects, they can often recognize objects presented as real-world solids [29–36]. Interestingly, the real-object advantage in agnosia patients appears to be related to expectations about the typical real-world size of tangible objects – information that is not available in 2-D images. Specifically, patients with visual agnosia performed better at object recognition for tangible objects than pictures but only when the physical size of the real objects was consistent with the typical real-world size [37].

A driving factor in the real-object advantage seems to be actability. An electroencephalography (EEG) study that found that, compared to matched images, real tools invoked a stronger and more sustained neural signature of motor preparation contralateral to the dominant hand of participants [38]. Moreover a neuroimaging study found different neural representations for real, tangible objects vs. similar images during hand actions, particularly when 3-D cues conveyed important information for grasping [39]. Notably, not all phenomena show a real-object advantage, and this can provide clues as to the nature of processing. For example, realism does not influence tool priming, suggesting that this particular phenomenon relies on a semantic process unaffected by actability [40,41].

Though we have focussed on visual objects, realism may also be important for sensory processing and recognition in domains other than vision. For example, audition researchers are coming to realize that a large body of research using simple tones may not generalize to natural sounds [42–45].

Memory

Real objects are more memorable than 2-D photographs of the same objects. When asked to remember everyday objects, free recall and recognition were both superior for participants

who viewed real objects compared to those who viewed colored photographs or line drawings [46]. Realism may be particularly important for the study of episodic memory, which is heavily framed by the context [47].

In an electroencephalography study, a neural correlate of familiarity for previously seen stimuli (the 'old-new' effect) was stronger when objects were first seen in real format then picture format compared to the converse [38]. These results suggest that stimuli that were first encountered as real-world objects were more memorable than their pictures [46].

Attention and gaze preferences

Real objects also capture attention more so than 2-D or 3-D images. When participants were asked to discriminate the orientation of tools (spoons), their reaction times were more affected by the orientation of irrelevant distractor objects when the stimuli were real objects compared to pictures or stereoscopic 3-D images [48]. Critically, the stronger attentional capture for real stimuli depended on actability. When the stimuli were positioned out of reach of the observer, or behind a transparent barrier that prevented in-the-moment interaction with the stimuli, the real objects elicited comparable interference effects as did the 2-D images. The use of a transparent barrier is a particularly elegant manipulation because it removes immediate actability while keeping the visual stimulus (including 3-D visual cues) nearly identical. Studies of eye gaze also suggest that real objects capture attention more effectively than pictures, an effect we term the **real-object preference**. For example, when infants as young as 7 months see a real object beside a matched photo, they spend more time gazing at the real object, even if they had already habituated to it [49]. Moreover, the preference for real objects is correlated with the frequency with which individual infants use touch to explore objects, suggesting that actability and multisensory interactions are key factors [50]. Macaque monkeys also spontaneously look at real objects longer than pictorial stimuli [51].

Neural evidence suggests that real objects are processed differently than matched pictures, consistent with behavioral preferences and mnemonic advantages for real objects. While repeated presentations of object images leads to reduced fMRI activation levels, a phenomenon called repetition suppression [52], surprisingly, such repetition effects were weak, if not absent, when real objects were repeated [53].

Neural evidence from humans and n0n-human primates suggests that brain responses in action-selective regions are driven by actability, with stronger responses in action-selective brain regions to real objects that are within reach [54], especially of the dominant hand [55]. Indeed the responses in some neurons is reduced or eliminated when a real object is placed beyond reach or blocked by a transparent barrier [56]. Similarly, in studies of human decision-making, real food is considered more valuable [57,58] and more irresistible [59] than images of food [60], particularly when it is seen within reach [58,61].

Attention also appears to be modulated by the cognitive knowledge that an object is real, including the recognition that real interactions would have real consequences. That is, actability may include not just the ability to grasp and manipulate stimuli but for those stimuli to have real ecological consequences. As one example, participants who believed

a real tarantula [62] or snake [63] was approaching showed robust activation in emotion-related regions, apparently stronger than for images [64].

Quantitative vs. Qualitative Differences Between Real Objects and Images

Accumulating evidence certainly suggests that, compared to artificial stimuli, realistic stimuli lead to *quantitative* changes in responses, including improvements in memory [46,47] object recognition [37], attention and gaze capture [48,49,51], or valuation [57,58]. These findings signal that realistic stimuli can amplify or strengthen behavioral and brain responses that might otherwise be difficult to observe when relying on proxies.

Though it may be tempting to dismiss quantitative changes as trivial, large changes can nevertheless be meaningful. When research statistics have limited power, especially in neuroimaging studies where costs limit sample sizes, boosts to effect sizes can affect the detectability of meaningful effects. However, given that artificial stimuli are easier to generate and present in the laboratory than real ones, an alternative approach to tackling quantitative differences would be to rely on proxies but compensate for attenuated effects, for example by designing experiments to maximize power (for example, with larger sample sizes).

Although quantitative differences are of theoretical and practical interest, a vital question for ongoing and future research is whether stimulus realism leads to *qualitative* differences. Qualitative differences would be reflected, for example, by different patterns of behavior, or by activation in different brain areas or networks, for real objects versus artificial stimuli. This question is vital because qualitative differences would suggest that the conclusions about behavior and brain processing generated from studies using images may not generalize to real stimuli. Finding qualitative differences associated with realism could enrich our understanding of naturalistic behavior and brain function, and critically inform theoretical frameworks of vision and action. While quantitative differences in responses to real objects and pictures that stem from lower-level attributes (such as an absence of binocular disparity cues) may seem uninteresting, these cues provide a gateway to higher-level stimulus characteristics, such as egocentric distance and real-world size, which are important for actability.

New methodological approaches that address the similarity of behavioral [65] or neural [66] "representations" (based on the similarity of ratings, or patterns of fMRI activation within a brain region, respectively) can assess qualitative differences in the way real versus artificial stimuli are processed, over and above differences in response magnitude. Using these approaches, recent studies have begun to reveal qualitative differences arising from realism.

Real and simulated objects appear to be "represented" differently in typical adult participants. When observers were tasked with manually arranging a set of graspable objects by their similarities, their groupings differed between pictures, real objects, and virtual 3-D objects presented using augmented reality (AR) [67]. Using a computer mouse, observers grouped pictures of objects by a conceptual factor, the objects' typical location. In contrast,

using the hands to move virtual 3-D projections of objects using AR, observers arranged the items not by their typical location but rather according to their physical characteristics of real-world size, elongation, and weight (rather surprisingly as virtual objects have no actual mass). Observers who lifted and moved real objects incorporated both conceptual and physical object properties into their arrangements. Thus, changes in the format of an object can lead to striking shifts in responses. Objects that can be manipulated using the hands, either directly in reality, or indirectly via AR, evoke richer processing of physical properties, such as real-world size and weight. Such results suggest that experimental approaches that rely solely on images may overestimate the role of cognitive factors and underestimate physical ones.

We propose a testable hypothesis: the factors that contribute to behavioral and neural differences (both qualitative and quantitative) will depend upon the processes being studied and the brain regions and networks that subserve them. For example, images may be a perfectly good, arguably preferable [68] stimulus for studying low-level visual processes at the earliest stages of the visual system (e.g., primary visual cortex). Even at higher levels of visual processing concerned with perception and semantics (the ventral visual stream), images may effectively evoke concepts to a comparable degree as real objects [41,69]. However, higher levels of visual processing for representing space and actions (the dorsal visual stream) [70,71] are likely to be strongly affected by depth [72–75], and actability [54].

From Stimulus Realism to Task Realism

In addition to enhancing the realness of stimuli, cognitive neuroscience will also benefit from approaches that allow more natural unfolding of cognitive *processes*. Growing evidence suggests that real actions affect behavior and brain processes differently than simulations. For example, real actions differ from pantomimed or delayed actions, both behaviorally and neurally [76–79].

Cognitive neuroscience uses a reductionist taxonomy of siloed cognitive functions – perception, action, attention, memory, emotion – studied largely in isolation [80] and by structured experimental paradigms with experimenter-defined trials and blocks [81]. Yet in everyday life, a multitude of cognitive functions and the brain networks that subserve them are seamlessly and dynamically integrated. New naturalistic approaches examine correlated brain activation across participants watching the same naturalistic movie segments [82,83]; the results corroborate findings from conventional approaches [84,85], providing a balance between rigour and realism [86]. Nevertheless, movie viewing is a passive act and may neglect active, top-down cognitive or motor processes.

Many experimental approaches in cognitive neuroscience map stimuli to responses without "closing the feedback loop" such that the consequences of one response become the stimulus for the next response [13]. This issue has long been recognized [87] but has had limited impact on mainstream experimental approaches. Alternative approaches advocate for the study of the "freely behaving brain" [88]. For example, studies of taxi drivers' neuroanatomy [89] and spontaneous cognitions during virtual driving tasks [90,91] have

unveiled the neural basis of dynamic real-world navigation. Emerging technologies may enable recording of human neural activity in real-world scenarios [92–96]. Emerging data-driven analysis strategies enable the study of brain processes during broad-ranging stimuli or behaviors [95,97,98], even under natural situations. Rather than trying to isolate stimulus or task features, users of these approaches realize that features that co-occur in the real world are likely jointly represented in brain organizational principles [99,100].

Immersive Neuroscience

Based on the theoretical and empirical differences between reality and proxies, we propose a potential new direction in the cognitive sciences, an approach we call **immersive neuroscience**. The goal of immersive neuroscience is to push the field of experimental cognitive sciences closer to the real world using a combination of realism and, where realism is challenging or impossible, compelling simulations such as VR/AR. The approach resonates with historical advocacy for an ecological approach ([15,101] and reviewed in depth elsewhere [47,102]). However, the development of new technologies for studying brain and behavior in realistic contexts has dramatically improved in recent decades (see Box 3).

We conceptualize the move toward realism as a continuum, with highly reduced stimuli (and tasks) on one end, fully real stimuli (and tasks) on the other, and gradations in between (see Figure 2). A typical assumption is that reduced stimuli provide high experimental control and convenience while sacrificing ecological validity; whereas, real-world stimuli provide the converse (but note that concepts like "ecological validity" and "real world" have been criticized as ill-defined and context dependent [103]). However, there are ways of optimizing both control/convenience and ecological validity through well-designed apparatus and protocols [104]. Although we have depicted ecological validity (and control/convenience) as a continuum, effects of stimulus richness may be monotonic but not necessarily gradual. That is, although gradual increases in stimulus realism could lead to gradual changes in behavior and brain responses, it is also plausible that there could be abrupt qualitative changes, for example between tangible solids and all simulations. Thus, although we are advocating for increasing realism, including better simulations like AR/VR, we view full reality as the empirical gold standard against which simulations should be assessed. Moreover, although we have depicted a continuum of realness according to visual richness, the relevant stimulus dimensions may turn out to be highly multidimensional and the dimensions may not be straightforward to define [100].

The immersive neuroscience approach is not intended to suggest that centuries of research using reductionist approaches are invalid, nor that all research necessitates realism. Reductionism is one essential approach in science and "starting simple" is especially important in the early years of new fields (such as neuroscience, little over a century old, and cognitive neuroscience, mere decades old). Moreover, research in the full-blown real world involves many technical challenges that can limit rapid progress [105]. That said, we promote an alternative to the reductionist or "build-up" approach of only using minimalist stimuli that become gradually combined to add complexity. The alternative is a "tear-down" approach in which we start with reality and remove components to see which matter

(Figure 2). Importantly, the "tear-down" approach is a complement, not a replacement, to the "build-up" approach. A third alternative is to relinquish the notion of experimentally dissecting behavior and brain processing at all, instead embracing the complexity of the natural environment in its own right [101,102]. The challenge remains that many current neuroscience techniques are not entirely "reality compatible". For example, "bringing the real world into the MRI scanner" [53] introduces elaborate technical challenges. In such cases, the tear-down approach can determine whether heroic efforts to enhance realism are worthwhile or whether easier proxies (e.g., VR/AR) may suffice.

A comprehensive understanding of the cognitive sciences will likely benefit from the combination of approaches: building up, tearing down, and studying fully natural situations and environments. While controlled experiments are useful for testing hypotheses about the contributions of components (e.g., whether a component is necessary or sufficient), ecological experiments are useful for testing whether those hypotheses generalize to natural settings, and for generating new hypotheses that consider the complexities of the organism in its environment.

The use of reality and validated proxies also informs cutting-edge computational techniques, such as artificial neural networks [106] which can handle the messy complexities of the natural world, as does the brain itself [107]. As some have recently argued [102], a key factor in the success of modern neural networks is the utilization of large data sets sampled from "the real world", including its inherent nonlinearities and interactions. This approach has proven far more successful than earlier approaches in artificial intelligence that sought sterilized and comprehendible algorithms based on limited input data. Nevertheless, artificial neural networks are often trained on data sets that lack potentially crucial aspects of realism. In vision science, for example, giant databases of static images [108] are used to train artificial neural networks and compare the "representations" to those in the brain [109–111]. While this approach has been enlightening, especially for ventral-stream perceptual processes, next-generation endeavors could benefit from incorporating 3-D depth structure [112], self-generated motion [113], embodiment [107], and active manipulation [114,115]. Such approaches could lead to artificial intelligence that learns in a manner akin to how human infants learn to comprehend and act within the real world through a series of transitions in their active experiences [25,116].

Concluding Remarks

We reviewed a growing body of literature that emphasizes (1) the theoretical arguments for why stimulus realism might matter for perception, cognition and action; (2) the feasibility of developing approaches, both real and virtual, for enhancing naturalism in research paradigms; (3) the evidence that realism affects multiple domains of cognition; (4) the factors that influence the real-object advantage, with actability seeming the most prevalent; and (5) a proposal for how an immersive neuroscience approach could operate and could benefit the field (see Outstanding Questions). In experimental psychology and neuroscience, restricting the stimuli, the tasks and the processes that we investigate may limit our understanding of the integrated workings of sensory-cognitive-motor loops and our theoretical framework of human and animal cognition. Such limitations may hamper

the development of real-world applications such as robotics [117] and brain-computer interfaces [118]. We argue that the importance of various components of realism is an empirical question, not merely a philosophical one. Reductionist and virtual proxies may prove appropriate for investigating cognition, but greater validation of these approaches in natural contexts would serve the community well.

Acknowledgments

Glossary

Actability
Whether a person can perform a genuine action upon a stimulus. For example, a nearby object may be reachable, and if it has an appropriate size and surface characteristics, also graspable. For example, a tennis ball within reach or being lobbed towards one is directly actable, a tennis ball lying on the other side of the court is not, and a cactus within reach may not be.

Action Associations
The semantic concepts evoked by a stimulus, even an image or a word, based on long-term experiences. For example, seeing an image of a knife may evoke the idea of cutting but not be actable.

Affordance
Defined by J. J. Gibson as "an action possibility formed between an agent and the environment." Despite its widespread use, there is little consensus on its meaning, which can include action possibilities in strict Gibsonian definition, semantic associations with objects (see Action Associations), or the potential for genuine interaction (see Actability). Depending on the definition, it is unclear whether images can evoke affordances; in our view, they may evoke associations and allow inferred affordances but not enable actability or true affordances.

Agency
The feeling of control over one's environment. For example, a computer user who sees a cursor move as he moves the mouse will experience a sense of agency.

Iconicity
the degree of visual similarity between a picture and its real-world referent.

Immersive neuroscience
A proposal for a new approach in cognitive neuroscience that places a stronger emphasis on real stimuli and tasks as well as compelling simulations such as virtual reality.

Motion parallax

Differences in the speed of retinal motion from objects at different depths for a moving observer, which can be used to extract 3-D relationships from retinal images, even monocularly.

Presence

The compelling feeling that a virtual stimulus or environment is real.

Proxy

A stimulus or task that is assumed to accurately represent a counterpart in the real world.

Real-Object Advantage

Improvements in behavioral performance, such as improved memory or recognition, for real objects compared to images.

Real-Object Preference

A preference to look at real objects more than images.

Stereopsis

Differences in the relative positions of a visual stimulus on the two retinas (also called binocular disparity) based on differences in depth, which can be used to extract 3-D relationships from the retinal images.

Vergence-accommodation conflict

Two important cues to absolute distance are vergence, the degree to which the eyeballs rotate inwardly so that the fixated point lands on the foveas, and accommodation, the degree to which the lenses of the eyes flex to focus the fixated object on the retinas. On VR/AR displays, the lenses accommodate to the screen at a fixed distance, but the eyes converge at distances further away, often causing discomfort, fatigue and feelings of sickness.

References

1. Magritte R and Torczyner H (1994) Letters Between Friends, Harry N. Abrams Inc

2. Gibson JJ (1954) A theory of pictorial perception. Audiov. Commun. Rev 2, 3–23

3. Gregory RL (1970) The Intelligent Eye, McGraw-Hill.

4. Ittelson WH (1996) Visual perception of markings. Psychon. Bull. Rev 3, 171–187 [PubMed: 24213865]

5. Kennedy JM and Ostry DJ (1976) Approaches to picture perception: perceptual experience and ecological optics. Can. J. Psychol 30, 90–98 [PubMed: 974898]

6. Aubert M et al. (2019) Earliest hunting scene in prehistoric art. Nature 576, 442–445 [PubMed: 31827284]

7. Hockney D and Gayford M (2018) A History of Pictures for Children, Abrams.

8. Forlim CG et al. (2019) Stereoscopic rendering via goggles elicits higher functional connectivity during virtual reality gaming. Front. Hum. Neurosci 13,

9. Baumgartner T (2008) Feeling present in arousing virtual reality worlds: prefrontal brain regions differentially orchestrate presence experience in adults and children. Front. Hum. Neurosci 2,

10. Cisek P and Kalaska JF (2010) Neural mechanisms for interacting with a world full of action choices. Annu. Rev. Neurosci 33, 269–98 [PubMed: 20345247]

11. Rosenbaum DA (2005) The Cinderella of psychology: The neglect of motor control in the science of mental life and behavior. Am. Psychol 60, 308–317 [PubMed: 15943523]

12. Anderson M and Chemero A (2016) The brain evolved to guide action. In Wiley Handbook of Evolutionary Neuroscience (Shepherd, S., ed), Wiley

13. Wolpert DM and Flanagan JR (2001) Motor prediction. Curr. Biol 11, R729–32 [PubMed: 11566114]

14. Sokolov AA et al. The cerebellum: Adaptive prediction for movement and cognition., Trends in Cognitive Sciences, 21. (2017), Elsevier Ltd, 313–332 [PubMed: 28385461]

15. Gibson JJ The Ecological Approach to Visual Perception, Erlbaum

16. Tucker M and Ellis R (1998) On the relations between seen objects and components of potential actions. J. Exp. Psychol. Hum. Percept. Perform 24, 830–46 [PubMed: 9627419]

17. Pitcher D et al. (2019) A functional dissociation of face-, body- and scene-selective brain areas based on their response to moving and static stimuli. Sci. Rep 9, 1–9 [PubMed: 30626917]

18. Pitcher D et al. (2011) Differential selectivity for dynamic versus static information in face-selective cortical regions. Neuroimage 56, 2356–2363 [PubMed: 21473921]

19. Beauchamp MS et al. (2002) Parallel visual motion processing streams for manipulable objects and human movements. Neuron 34, 149–159 [PubMed: 11931749]

20. Gaebler M et al. (2014) Stereoscopic depth increases intersubject correlations of brain networks. Neuroimage 100, 427–434 [PubMed: 24945664]

21. Finlayson NJ et al. (2017) Differential patterns of 2D location versus depth decoding along the visual hierarchy. Neuroimage 147, 507–516 [PubMed: 28039760]

22. DeLoache JS et al. (1998) Grasping the nature of pictures. Psychol. Sci. 9, 205–210

23. Gibson EJ (1988) Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. Annu. Rev. Psychol 39, 1–42

24. Franchak JM (2020) The ecology of infants' perceptual-motor exploration. Curr. Opin. Psychol 32, 110–114 [PubMed: 31445428]

25. Adolph KE and Hoch JE (2019) Motor development: Embodied, embedded, enculturated, and enabling. Annu. Rev. Psychol 70, 141–164 [PubMed: 30256718]

26. Held R and Bossom J (1961) Neonatal deprivation and adult rearrangement: Complementary techniques for analyzing plastic sensory-motor coordinations. J. Comp. Physiol. Psychol 54, 33–37 [PubMed: 13713069]

27. Held R and Hein A (1963) Movement-produced stimulation in the development of visually guided behavior. J. Comp. Physiol. Psychol 56, 872–876 [PubMed: 14050177]

28. Bovet D and Vauclair J (2000) Picture recognition in animals and humans. Behav. Brain Res 109, 143–165 [PubMed: 10762685]

29. Chainay H and Humphreys GW (2001) The real-object advantage in agnosia: Evidence for a role of surface and depth information in object recognition. Cogn. Neuropsychol 18, 175–91 [PubMed: 20945210]

30. Hiraoka K et al. (2009) Visual agnosia for line drawings and silhouettes without apparent impairment of real-object recognition: a case report. Behav. Neurol 21, 187–92 [PubMed: 19996516]

31. Farah M (1990) Visual Agnosia, MIT Press.

32. Humphrey GK et al. (1994) The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. Perception 23, 1457–81 [PubMed: 7792135]

33. Riddoch MJ and Humphreys GW (1987) A case of integrative visual agnosia. Brain 110 (Pt 6, 1431–62 [PubMed: 3427396]

34. Servos P et al. (1993) The drawing of objects by a visual form agnosic: contribution of surface properties and memorial representations. Neuropsychologia 31, 251–9 [PubMed: 8492878]

35. Turnbull OH et al. (1998) 2D but not 3D: pictorial-depth deficits in a case of visual agnosia. Cortex 40, 723–38

36. Ratcliff G and Newcombe F (1982) Object recognition: some deductions from the clinical evidence. In The Handbook of Adult Language Disorders (Ellis A, ed), Psychology Press

37. Holler DE et al. (2019) Real-world size coding of solid objects, but not 2-D or 3-D images, in visual agnosia patients with bilateral ventral lesions. Cortex 119, 555–568 [PubMed: 30987739]

38. Marini F et al. (2019) Distinct visuo-motor brain dynamics for real-world objects versus planar images. Neuroimage 195, 232–242 [PubMed: 30776529]

39. Freud E et al. (2018) Getting a grip on reality: Grasping movements directed to real objects and images rely on dissociable neural representations. Cortex; 98, 34–48 [PubMed: 28431740]

40. Snow J et al. (2016) fMRI reveals different activation patterns for real objects vs. photographs of objects. J. Vis 16, 512

41. Kithu MC et al. (2021) A priming study on naming real versus pictures of tools. Exp. Brain Res 9, 15c

42. Schutz M and Gillard J (2020) On the generalization of tones: a detailed exploration of non-speech auditory perception stimuli. Sci. Rep DOI: 10.1038/s41598-020-63132-2

43. Hamilton LS and Huth AG (2020) The revolution will not be controlled: natural stimuli in speech neuroscience. Lang. Cogn. Neurosci 35, 573–582 [PubMed: 32656294]

44. Neuhoff JG (2004) Ecological Psychoacoustics, Elsevier.

45. Gaver WW (1993) What in the world do we hear?: An ecological approach to auditory event perception. Ecol. Psychol 5, 1–29

46. Snow JC et al. (2014) Real-world objects are more memorable than photographs of objects. Front. Hum. Neurosci 8, 837 [PubMed: 25368568]

47. Shamay-Tsoory SG and Mendelsohn A (2019) Real-life neuroscience: An ecological approach to brain and behavior research. Perspect. Psychol. Sci 14, 841–859 [PubMed: 31408614]

48. Gomez MA et al. (2018) Graspable objects grab attention more than images do. Psychol. Sci 29, 206–218 [PubMed: 29215960]

49. Gerhard TM et al. (2016) Distinct visual processing of real objects and pictures of those objects in 7- to 9-month-old infants. Front. Psychol 7, 1–9 [PubMed: 26858668]

50. Gerhard TM et al. (2021) Manual exploration of objects is related to 7-month-old infants' visual preference for real objects. Infant Behav. Dev 62, 101512 [PubMed: 33310403]

51. Mustafar F et al. (2015) Enhanced visual exploration for real objects compared to pictures during free viewing in the macaque monkey. Behav. Processes 118, 8–20 [PubMed: 26003135]

52. Grill-Spector K et al. (2006) Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci 10, 14–23 [PubMed: 16321563]

53. Snow JC et al. (2011) Bringing the real world into the fMRI scanner: Repetition effects for pictures versus real objects. Sci. Rep. 1,

54. Gallivan JP et al. (2009) Is that within reach? fMRI reveals that the human superior parieto-occipital cortex encodes objects reachable by the hand. J. Neurosci 29, 4381–91 [PubMed: 19357266]

55. Gallivan JP et al. (2011) Neuroimaging reveals enhanced activation in a reach-selective brain area for objects located within participants' typical hand workspaces. Neuropsychologia 49, 3710–3721 [PubMed: 21958649]

56. Bonini L et al. (2014) Space-Dependent Representation of Objects and Other's Action in Monkey Ventral Premotor Grasping Neurons. J. Neurosci 34, 4108–4119 [PubMed: 24623789]

57. Romero CA et al. (2018) The real deal: Willingness-to-pay and satiety expectations are greater for real foods versus their images. Cortex 107, 78–91 [PubMed: 29233524]

58. Bushong B et al. (2010) Pavlovian processes in consumer choice: The physical presence of a good increases willingness-to-pay. Am. Econ. Rev 100, 1556–1571

59. Medic N et al. (2016) The presence of real food usurps hypothetical health value judgment in overweight people. eNeuro 3, 412–427

60. Mischel W and Moore B (1973) Effects of attention to symbolically presented rewards on self-control. J. Pers. Soc. Psychol 28, 172–179 [PubMed: 4747220]

61. Mischel W et al. (1972) Cognitive and attentional mechanisms in delay of gratification. J. Pers. Soc. Psychol 21, 204–18 [PubMed: 5010404]

62. Mobbs D et al. (2010) Neural activity associated with monitoring the oscillating threat value of a tarantula. Proc. Natl. Acad. Sci. U. S. A 107, 20582–20586 [PubMed: 21059963]

63. Nili U et al. (2010) Fear Thou Not: Activity of Frontal and Temporal Circuits in Moments of Real-Life Courage. Neuron 66, 949–962 [PubMed: 20620879]

64. Camerer C and Mobbs D (2017) Differences in behavior and brain activity during hypothetical and real choices. Trends Cogn. Sci 21, 46–56 [PubMed: 27979604]

65. Kriegeskorte N and Mur M (2012) Inverse MDS: Inferring dissimilarity structure from multiple item arrangements. Front. Psychol 3, 1–13 [PubMed: 22279440]

66. Kriegeskorte N et al. (2008) Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci 2, 4 [PubMed: 19104670]

67. Holler DE et al. (2020) Object responses are highly malleable, rather than invariant, with changes in object appearance. Sci. Rep 10, 1–14 [PubMed: 31913322]

68. Rust NC and Movshon JA (2005) In praise of artifice. Nat. Neurosci 8, 1647–1650 [PubMed: 16306892]

69. Squires SD et al. (2016) Priming tool actions: Are real objects more effective primes than pictures? Exp. Brain Res 234, 963–976 [PubMed: 26686378]

70. Ungerleider LG and Mishkin M (1982) Two cortical visual systems. In Analysis of Visual Behavior (Ingle DJ et al., eds), pp. 549–586, MIT Press

71. Goodale MA and Milner AD (1992) Separate visual pathways for perception and action. Trends Neurosci 15, 20–5 [PubMed: 1374953]

72. Welchman AE (2016) The human brain in depth: How we see in 3D. Annu. Rev. Vis. Sci 2, 345–376 [PubMed: 28532360]

73. Minini L et al. (2010) Neural modulation by binocular disparity greatest in human dorsal visual stream. J. Neurophysiol 104, 169–78 [PubMed: 20445027]

74. Sakata H et al. (1998) Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. Philos. Trans. R. Soc. London B 353, 1363–73 [PubMed: 9770229]

75. Theys T et al. (2015) Shape representations in the primate dorsal visual stream. Front. Comput. Neurosci 9,

76. Goodale M et al. (1994) Differences in the visual control of pantomimed and natural grasping movements. Neuropsychologia 32, 1159–78 [PubMed: 7845558]

77. Króliczak G et al. (2007) What does the brain do when you fake it? An FMRI study of pantomimed and real grasping. J. Neurophysiol 97, 2410–22 [PubMed: 17229828]

78. Singhal A et al. (2013) Human fMRI reveals that delayed action re-recruits visual perception. PLoS One 8, e73629 [PubMed: 24040007]

79. Cohen NR et al. (2009) Ventral and dorsal stream contributions to the online control of immediate and delayed grasping: a TMS approach. Neuropsychologia 47, 1553–62 [PubMed: 19168086]

80. Hastings J et al. (2014) Interdisciplinary perspectives on the development, integration, and application of cognitive ontologies. Front. Neuroinform 8, 1–7 [PubMed: 24501593]

81. Huk A et al. (2018) Beyond trial-based paradigms: Continuous behavior, ongoing neural activity, and natural stimuli. J. Neurosci 38, 7551–7558 [PubMed: 30037835]

82. Hasson U et al. (2004) Intersubject synchronization of cortical activity during natural vision. Science 303, 1634–40 [PubMed: 15016991]

83. Hasson U and Honey CJ (2012) Future trends in neuroimaging: Neural processes as expressed within real-life contexts. Neuroimage 62, 1272–8 [PubMed: 22348879]

84. Cantlon JF and Li R (2013) Neural activity during natural viewing of Sesame Street statistically predicts test scores in early childhood. PLoS Biol 11,

85. Richardson H et al. (2018) Development of the social brain from age three to twelve years. Nat. Commun 9, 1027 [PubMed: 29531321]

86. Cantlon JF (2020) The balance of rigor and reality in developmental neuroscience. Neuroimage DOI: 10.1016/j.neuroimage.2019.116464

87. Dewey J (1896) The reflex arc concept in psychology. Psychol. Rev 3, 357–370

88. Maguire EA (2012) Studying the freely-behaving brain with fMRI. Neuroimage 62, 1170–1176 [PubMed: 22245643]

89. Maguire EA et al. (2000) Navigation-related structural change in the hippocampi of taxi drivers. Proc. Natl. Acad. Sci 97, 4398–4403 [PubMed: 10716738]

90. Spiers HJ and Maguire EA (2007) Decoding human brain activity during real-world experiences. Trends Cogn. Sci 11, 356–365 [PubMed: 17618161]

91. Spiers HJ and Maguire EA (2008) The dynamic nature of cognition during wayfinding. J. Environ. Psychol 28, 232–249 [PubMed: 19325934]

92. Boto E et al. (2018) Moving magnetoencephalography towards real-world applications with a wearable system. Nature 555, 657–661 [PubMed: 29562238]

93. Melroy S et al. (2017) Development and design of next-generation head-mounted Ambulatory Microdose Positron-Emission Tomography (AM-PET) System. Sensors (Basel) 17,

94. Pinti P et al. A review on the use of wearable functional near-infrared spectroscopy in naturalistic environments., Japanese Psychological Research, 60. (2018), 347–373 [PubMed: 30643322]

95. Gabriel PG et al. (2019) Neural correlates of unstructured motor behaviors. J. Neural Eng 16, 066026 [PubMed: 31342926]

96. Fishell AK et al. (2019) Mapping brain function during naturalistic viewing using high-density diffuse optical tomography. Sci. Rep 9, 11115 [PubMed: 31366956]

97. Pinti P et al. (2017) A novel GLM-based method for the Automatic IDentification of functional Events (AIDE) in fNIRS data recorded in naturalistic environments. Neuroimage 155, 291–304 [PubMed: 28476662]

98. Baldassano C et al. (2017) Discovering Event Structure in Continuous Narrative Perception and Memory. Neuron 95, 709–721.e5 [PubMed: 28772125]

99. Graziano MSA and Aflalo TN (2007) Rethinking cortical organization: Moving away from discrete areas arranged in hierarchies. Neuroscientist 13, 138–147 [PubMed: 17404374]

100. Bao P et al. (2020) A map of object space in primate inferotemporal cortex. Nature 583, 103–108 [PubMed: 32494012]

101. Brunswik E (1947) Perception and the Representive Design of Psychological Experiments, University of California Press.

102. Nastase SA et al. (2020) Keep it real: rethinking the primacy of experimental control in cognitive neuroscience. Neuroimage 222, 117254 [PubMed: 32800992]

103. Holleman GA et al. (2020) The 'real-world approach' and its problems: A critique of the term ecological validity. Front. Psychol 11, 1–12 [PubMed: 32038435]

104. Romero CA and Snow JC (2019) Methods for presenting real-world objects under controlled laboratory conditions. J. Vis. Exp DOI: 10.3791/59762

105. Hessels RS et al. (2020) Wearable technology for "real-world research": Realistic or not? Perception 49, 611–615 [PubMed: 32552490]

106. Lecun Y et al. (2015) Deep learning. Nature 521, 436–444 [PubMed: 26017442]

107. Hasson U et al. (2020) Direct fit to nature: An evolutionary perspective on biological and artificial neural networks. Neuron 105, 416–434 [PubMed: 32027833]

108. Deng J et al. (2009), ImageNet: A large-scale hierarchical image database., in 2009 Conference on Computer Vision and Pattern Recognition

109. Kriegeskorte N (2015) Deep neural networks: A new framework for modeling biological vision and brain information processing. Annu. Rev. Vis. Sci 1, 417–446 [PubMed: 28532370]

110. Rajalingham R et al. (2018) Large-scale, high-resolution comparison of the core visual object recognition behavior of humans, monkeys, and state-of-the-art deep artificial neural networks. J. Neurosci 38, 7255–7269 [PubMed: 30006365]

111. Chang N et al. (2019) BOLD5000, a public fMRI dataset while viewing 5000 visual images. Sci. Data 6, 1–18 [PubMed: 30647409]

112. Rahman MM et al. (2019) Recent advances in 3D object detection in the era of deep neural networks: A survey. IEEE Trans. Image Process DOI: 10.1109/TIP.2019.2955239

113. Durdevic P and Ortiz-Arroyo D (2020) A deep neural network sensor for visual servoing in 3D spaces. Sensors 20,

114. Harman KL et al. (1999) Active manual control of object views facilitates visual recognition. Curr. Biol 9, 1315–1318 [PubMed: 10574764]

115. Arapi V et al. (2018) DeepDynamicHand: A deep neural architecture for labeling hand manipulation strategies in video sources exploiting temporal information. Front. Neurorobot 12, 86 [PubMed: 30618707]

116. Smith LB and Slone LK (2017) A developmental approach to machine learning? Front. Psychol 8, 1–10 [PubMed: 28197108]

117. Pfeifer R et al. (2007) Self-organization, embodiment, and biologically inspired robotics. Science 318, 1088–93 [PubMed: 18006736]

118. Hochberg LR et al. (2012) Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. Nature 485, 372–5 [PubMed: 22596161]

119. Fagan JF (1977) Infant recognition memory: studies in forgetting. Child Dev 48, 68–78 [PubMed: 844363]

120. Slater A et al. (1983) Perception of shape by the new-born baby. Br. J. Dev. Psychol 1, 135–142

121. Slater A et al. (1984) New-born infants' perception of similarities and differences between two-and three-dimensional stimuli. Br. J. Dev. Psychol 2, 287–294

122. DeLoache JS et al. (1979) Picture perception in infancy. Infant Behav. Dev 2, 77–89

123. Dirks J and Gibson E (1977) Infants' perception of similarity between live people and their photographs. Child Dev 48, 124–30 [PubMed: 844350]

124. Rose SA (1977) Infants ' transfer of response between two-dimensional and three-dimensional stimuli. Child Dev 48, 1086–1091

125. Shinskey JL and Jachens LJ (2014) Picturing objects in infancy. Child Dev 85, 1813–1820 [PubMed: 24779447]

126. Ganea PA et al. (2009) Toddlers' referential understanding of pictures. Nat. Rev. Neurosci 19, 283–95

127. Pierroutsakos SL and DeLoache JS (2003) Infants' manual exploration of pictorial objects varying in realism. Infancy 4, 141–156

128. Troje NF (2019) Reality check. Perception 48, 1033–1038 [PubMed: 31570079]

129. Harris DJ et al. (2019) Virtually the same? How impaired sensory information in virtual reality may disrupt vision for action. Exp. Brain Res DOI: 10.1007/s00221-019-05642-8

130. Culham JC et al. (2003) Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. Exp. Brain Res 153, 180–9 [PubMed: 12961051]

131. Brandi ML et al. (2014) The neural correlates of planning and executing actual tool use. J. Neurosci 34, 13183–13194 [PubMed: 25253863]

132. Schaffelhofer S and Scherberger H (2016) Object vision to hand action in macaque parietal, premotor, and motor cortices. Elife 5, 1–24

133. Julesz B (1971) Foundations of Cyclopean Perception, University of Chicago Press.

134. Coutant BE and Westheimer G (1993) Population distribution of stereoscopic ability. Ophthalmic Physiol. Opt 13, 3–7 [PubMed: 8510945]

135. Hess RF et al. (2015) Stereo vision: The haves and have-nots. Iperception 6, 2041669515593028 [PubMed: 27433314]

Box 1: Terminology for Visual Stimuli

One of the challenges in understanding differences between stimuli is the conflicting nomenclature used by different researchers. Some of the confusion could be remedied by adopting of a consistent set of terms:

2-D Images

2-D Images are planar displays that lack consistent cues to depth. Most commonly, 2-D images are presented via a computer to a monitor or projection screen; however, they can also be printed. They may differ in **iconicity,** the degree to which the picture resembles the real object [126], ranging from line drawings to photographs. Typically, 2-D images provide only monocular cues to depth (e.g., shading, occlusion) but no stereoscopic or motion-parallax cues to the depth of components within the image. 2-D images often misrepresent the size of the depicted object, with unrealistic relationships between physical size, distance and familiar size.

3-D Images

3-D Images are generated by presenting two 2-D images, one to each eye, that provide stereoscopic cues to depth. They assume a fixed head position and do not provide motion parallax if the observer moves.

Real Objects

Real Objects are tangible solids that can be interacted with. We prefer not to use the term "3-D Objects," which implies that the key difference from 2-D images is the third dimension (vs. other potential differences such as tangibility or actability). Although many researchers use the terms "real" or "real-world" for images that depict real objects, confusion could be avoided by limiting the use of "real" to physical objects and using other descriptors like "realistic" for representations that capture reality incompletely.

Simulated Realities

Simulated Realities (SR) includes approaches that aim to induce a sense of immersion, presence and **agency**. In Virtual Reality (VR), a computer-generated simulated environment is rendered in a head-mounted display (HMD). The display provides stereoscopic 3-D depth cues and changes as the observer moves, giving the observer the illusions of presence and agency. VR may also enable interactions with objects, typically through hand-held controllers. In Augmented Reality (AR) computer-generated stimuli are superimposed on an observer's view of the real environment (through a transparent head-mounted display (HMD) or passed-through live video from cameras attached to the HMD). Some use the term Mixed Reality (XR) to encompass VR and AR; however, the term is used inconsistently (sometimes being treated as synonymous with AR) so we propose the term Simulated Realities. SR could include other approaches that are technically not VR or AR, such as game engines that present stimuli on monitors or 3-D projectors rather than HMDs.

Box 2: Realness influences behavior in human infants and animals

Similar to adults, children's object recognition is influenced by the realness of the stimuli and tasks. Object recognition in infants has traditionally been investigated using habituation tasks, which are based on infants' general preference for new stimuli over previously seen items, as reflected in looking and grasping behaviors [119]. Capitalizing on this novelty preference in habituation tasks, infants are initially exposed (or 'habituated') to a stimulus, such as a picture of a shape or toy, and are subsequently presented with a different stimulus, such as a real object, to see whether or not habituation transfers to the new item. Habituation studies have shown that human infants can perceive the information carried in pictures even with remarkably little pictorial experience. For example, newborn babies can discriminate between basic geometric figures that differ in shape alone [120] and between 2-D versus 3-D stimuli [121,122]. By 6–9 months of age, infants can distinguish pictures from real objects. Moreover, infants show generalization of habituation from real objects to the corresponding 2-D pictures [122–124], and from 2-D pictures to their real objects [124,125]. Although picture-to-real object transfer in recognition performance improves with age and iconicity [126], this does not mean that infants *understand* what pictures are. Elegant work by Judy DeLoache and colleagues, for example, has shown that when given colored photos of toys, young infants initially attempted to grasp the pictures off the page, as do children raised in environments without pictures (such as Beng infants from West Africa) [22]. As in habituation studies showing picture-to-object transfer, manual investigation of pictures of toys is influenced by how realistic the pictures look, with more realistic pictures triggering greater manual investigation [127]. In a comprehensive review of the literature on picture perception, Bovet & Vauclair [28] concluded that a diverse range of animals (from invertebrates through great apes) can show appropriate responses to depicted stimuli, particularly ecologically relevant stimuli such as predators or conspecifics. Similar to human infants, such effects in animals appear strongest when animation or three-dimensionality are present and with increased exposure to pictorial stimuli. Nevertheless, their review highlights many cases where animals, infants and individuals from cultures without pictures fail to recognize depicted objects.

Box 3: The Potential of Simulated Realities in Research

Thus far we have presented empirical evidence that the realness of stimuli and tasks affects behavior and brain activation; one key question for moving forward is to determine which aspects of realness matter. Studies that use fully real stimuli are providing important new insights into perception, cognition and action. Nevertheless, studies that use real stimuli and tasks also present practical challenges that are not typically encountered when using images, such as the need to control or factor out potential confounds. Yet there are powerful arguments for moving away from approaches in which cognitive and neural processes are studied as if they are discrete de-contextualized events, but rather, as a continuous stream of sensory-cognitive-motor loops, reminiscent of how they unfold in naturalistic environments [81].

One emerging tool for both enhancing realism and testing the importance of components of realism is virtual reality (VR) and augmented reality (AR). There is growing enthusiasm for simulated realities (SR) in cognitive sciences because the technology enables a much more compelling experience for a user or research participant than conventional (typically computer-screen based) stimuli. VR/AR has the potential to optimize the trade-off between realism and control, though how well they evoke natural behavior and brain responses remains an open question and one that must be actively addressed (see Figure I). That is, researchers should not take for granted that VR/AR is a perfect proxy for reality, particularly in light of current technical limitations. Some contend that, because VR and AR can compellingly place the observer egocentrically and actively in a scene, especially through self-generated **motion parallax**, they evoke a sense of presence and recruit the dorsal visual stream more than much more reduced simulations like 2-D images [128]. Others contend that present-day technology is limited in verisimilitude because it may not fully engage action systems without genuine interactions directly with the body (vs. handheld controllers without haptic feedback) and displays are limited (e.g., low spatiotemporal resolution, small field of view) [129]. Many of these limitations are under active development by technology firms and will likely improve dramatically in the near future. However, some features are hard to improve, particularly the **vergence-accommodation conflict** and the need for compelling, affordable haptics that do not require cumbersome cybergloves. Moreover, issues such as motion sickness may limit the utility of simulation and no matter how good the technology becomes, people may remain aware that even compelling environments are not real and thus lack complete presence.

Outstanding Questions

- How well do common stimuli (e.g., pictures) and tasks developed for cognitive neuroscience research evoke the same behaviors and neural processes as real-world situations? Are differences qualitative or merely quantitative?

- When differences between proxies and reality are found, what drives those differences (e.g., three-dimensionality, multisensory attributes/tangibility, actability, ability to fulfill goals, presence)?

- Given the long-history of the "build-up" approach (reductionism) in cognitive neuroscience, how can the "tear-down" approach be used to ask interesting new questions?

- How can simulated realities (VR/AR/game engines) be used to bring cognitive neuroscience closer to realism? How do these technologies need to improve to better simulate reality? How can we optimize the balance between experimental control and convenience vs. ecological validity in research? How can simulated realities be optimized for commercial, practical and clinical applications (e.g., image-guided surgery, VR training, rehabilitation)?

- How can cognitive neuroscience study not just natural stimuli and tasks but also the natural cognitive processes of the "freely behaving brain"?

- As the "real-world" comes to include more technology with simulated stimuli and interactions (e.g., smartphones, computers, VR), how does this affect behavior and brain processing?

- As artificial neural network approaches improve, how can they take advantage of the complexities of the natural world and the means by which organisms learn in the natural world?

- How can emerging technologies measure brain activation in humans with fewer constraints (e.g., functional near-infrared spectroscopy) and enable real-world neuroscience?

Highlights

Although commonly utilized in cognitive neuroscience, images evoke different behavior and brain processing compared to the real, tangible objects they aim to approximate. Differences have been found in perception, memory, and attention. One key factor in differences appears to be that only real objects can be acted upon.

Evolution and development are shaped by the real world. Shortfalls between proxies and reality are evident in other species and in young children, who have not learned to comprehend representations as human adults can.

New technologies and new experimental approaches provide the means to study cognitive neuroscience with a balance between experimental control and ecological validity.
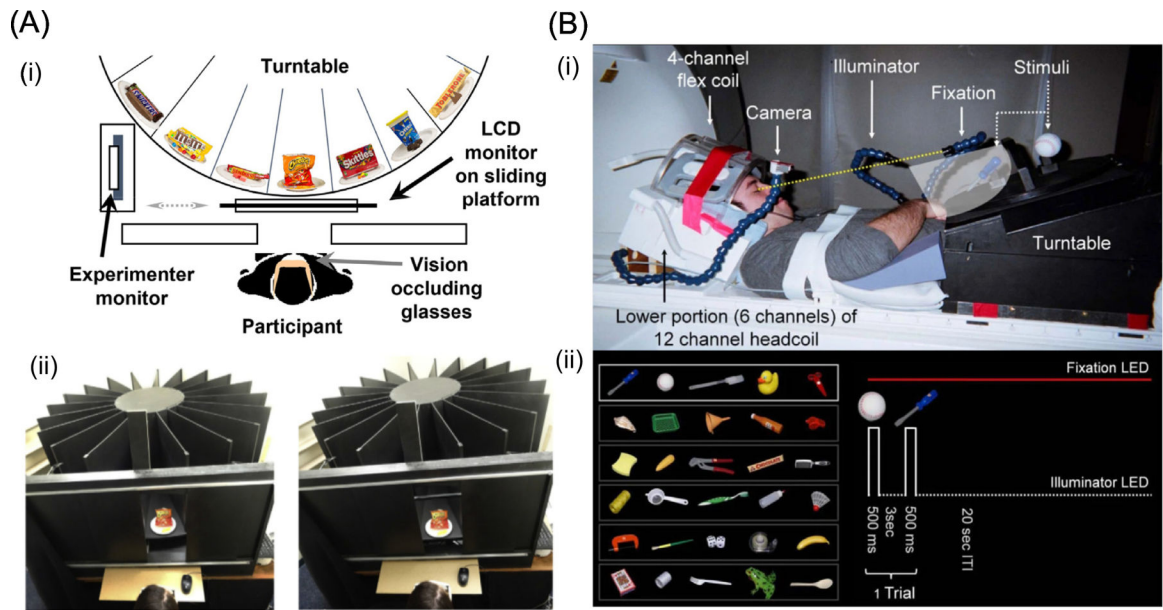
In addition to the standard approach of "building up" ecologically valid stimuli from simpler components, a complementary approach is to use reality as the gold standard and "tear down" various components to infer their contributions to behavior and brain processes.

Figure 1: Methods used to study behavior and brain responses to real objects

Innovative methods used to compare responses to real objects and representations. **(A)** Example from behavior. **(i)** In a recent study of decision-making, Romero et al. [57] used a custom-built turntable device to display a large set of real objects and closely-matched 2-D computerized images of everyday snack foods. Schematic shows the experimental setup from above. On real object trials the stimuli were visible on one sector of the turntable; on image trials the stimuli were displayed on a retractable monitor mounted on a sliding platform. Stimulus viewing on all trials was controlled using liquid-crystal glasses that alternated from transparent ('closed') to opaque ('open') states. **(ii)** Real object trial (left); image trial (right). Though stimuli are shown from above here, from the participants' viewpoint, displays appeared similar except for differences in stereopsis. **(B)** Presenting observers with real objects is especially challenging within fMRI environments. **(i)** Snow et al. [53] used fMRI to compare brain responses to everyday real-world objects versus photos. Using a repetition-suppression design, pairs of real/picture stimuli were presented from trial-to-trial on a turntable mounted over the participant's waist. Following from Culham et al., [130], the head coil was tilted forwards to enable participants to view the stimuli directly, without the use of mirrors. **(ii)** On each trial, two objects (lower left), each mounted on opposite sides of a central partition, were presented in rapid succession. Stimulus viewing on each trial was controlled using time-locked LED illumination; gaze was controlled using a red fixation light (lower right).
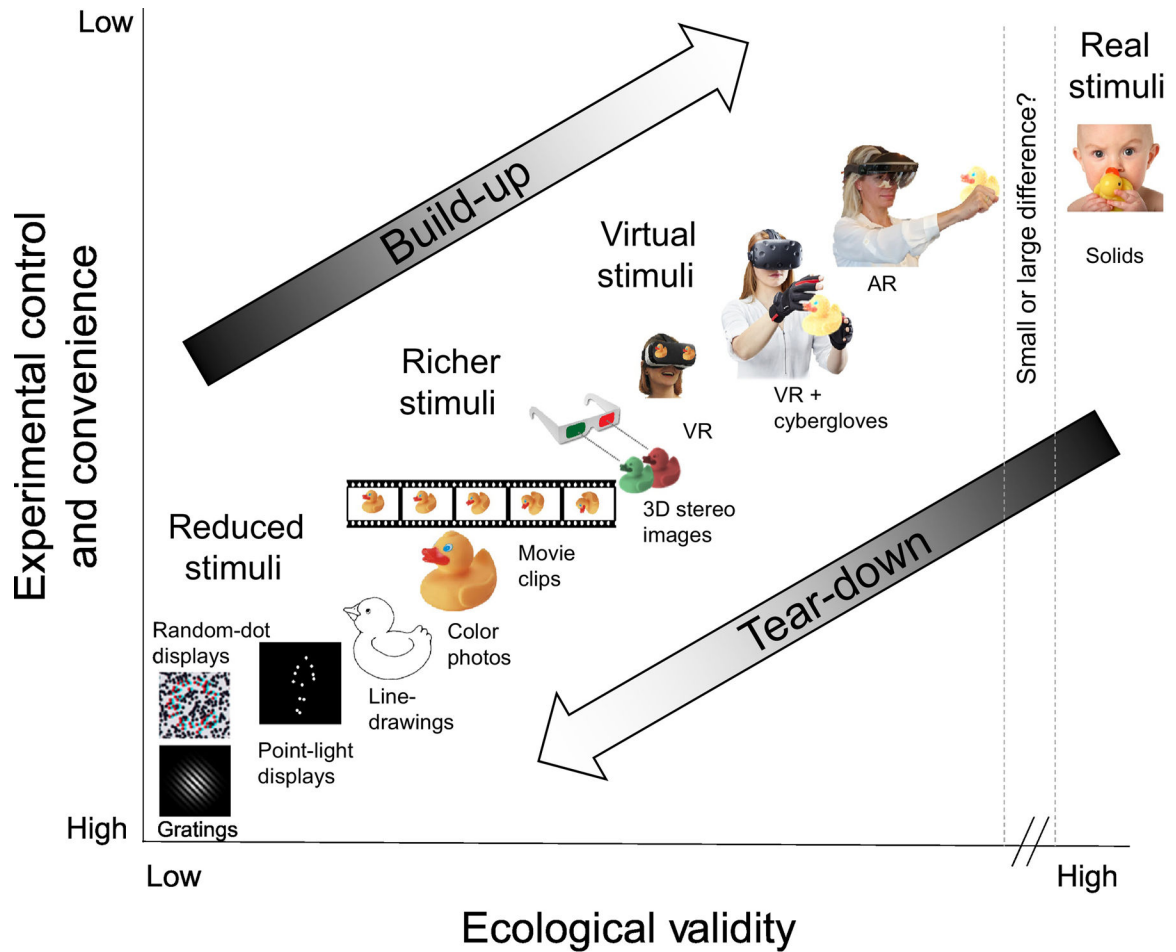
Trends in Cognitive Sciences

Figure 2: Tearing down vs. building up approaches
Different stimuli can be conceptualized as falling along a continuum of realness, from reduced or artificial (low in ecological validity, as shown in the *lower left*), to fully real (high in ecological validity, as shown in the *upper right*). Although ecological validity and control/convenience are thought to trade off, immersive neuroscience approaches can optimize both through well-designed apparatus and protocols. Answering questions about the importance of realness requires a fundamental shift away from the traditional "build-up" approach, in which cognition is studied by making reduced stimuli gradually more complex, to a "tear-down" approach, in which we start by studying responses to fully real stimuli and then gradually remove components. Although "tear-down" and "build-up" approaches may not always yield the same results, combining the two methods will permit a fuller understanding of the cognitive and brain mechanisms that support naturalistic vision and action. For example, the importance of stereopsis as a depth cue may differ between a build-up approach using random-dot stereograms [133] and a tear-down approach in which other depth cues (especially motion parallax) are available. A tear-down approach can reveal whether solids are processed qualitatively differently than artificial stimuli (represented by the distance between vertical dashed gray lines), in which case responses to real-world solids cannot be predicted by those to pictures or virtual stimuli. We postulate that the gap between
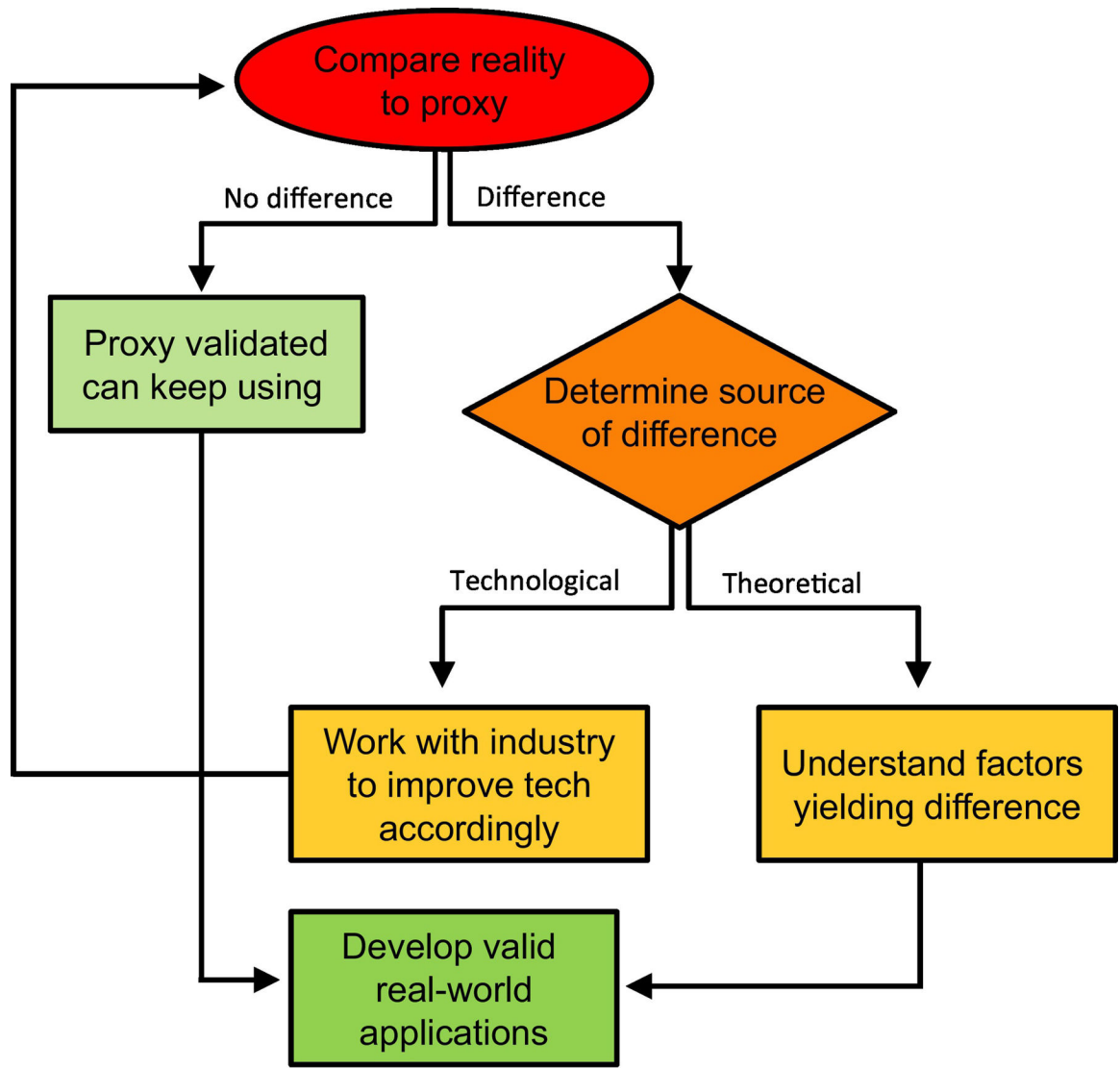
artificial stimuli vs. real objects may be more quantitative or qualitative depending on the participants' task and the brain area under study.

Figure I (Box 3). Which aspects of reality matter?
The flowchart illustrates a sequence to determine and optimize the validity of a proxy for reality.