



# The ‘Real-World Approach’ and Its Problems: A Critique of the Term Ecological Validity

Gijs A. Holleman<sup>1,2\*</sup>, Ignace T. C. Hooge<sup>1</sup>, Chantal Kemner<sup>1,2,3</sup> and Roy S. Hessels<sup>1,2</sup>

<sup>1</sup> Department of Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, Netherlands, <sup>2</sup> Department of Developmental Psychology, Utrecht University, Utrecht, Netherlands, <sup>3</sup> Brain Center, University Medical Center Utrecht, Utrecht, Netherlands

## OPEN ACCESS

### Edited by:

Matthias Gamer,  
Julius Maximilian University  
of Würzburg, Germany

### Reviewed by:

Yoni Pertzov,  
The Hebrew University of Jerusalem,  
Israel  
Nicola Jean Gregory,  
Bournemouth University,  
United Kingdom

### \*Correspondence:

Gijs A. Holleman  
g.a.holleman@uu.nl

### Specialty section:

This article was submitted to  
Cognitive Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 24 January 2020

**Accepted:** 25 March 2020

**Published:** 30 April 2020

### Citation:

Holleman GA, Hooge ITC,  
Kemner C and Hessels RS (2020) The  
‘Real-World Approach’ and Its  
Problems: A Critique of the Term  
Ecological Validity.  
Front. Psychol. 11:721.  
doi: 10.3389/fpsyg.2020.00721

A popular goal in psychological science is to understand human cognition and behavior in the ‘real-world.’ In contrast, researchers have typically conducted their research in experimental research settings, a.k.a. the ‘psychologist’s laboratory.’ Critics have often questioned whether psychology’s laboratory experiments permit generalizable results. This is known as the ‘real-world or the lab’-dilemma. To bridge the gap between lab and life, many researchers have called for experiments with more ‘ecological validity’ to ensure that experiments more closely resemble and generalize to the ‘real-world.’ However, researchers seldom explain what they mean with this term, nor how more ecological validity should be achieved. In our opinion, the popular concept of ecological validity is ill-formed, lacks specificity, and falls short of addressing the problem of generalizability. To move beyond the ‘real-world or the lab’-dilemma, we believe that researchers in psychological science should always specify the particular context of cognitive and behavioral functioning in which they are interested, instead of advocating that experiments should be more ‘ecologically valid’ in order to generalize to the ‘real-world.’ We believe this will be a more constructive way to uncover the context-specific and context-generic principles of cognition and behavior.

**Keywords:** ecological validity, experiments, real-world approach, generalizability, definitions

## INTRODUCTION

A popular goal in psychological science is to understand human cognition and behavior in the ‘real-world.’ In contrast, researchers have traditionally conducted experiments in specialized research settings, a.k.a. the ‘psychologist’s laboratory’ (Danziger, 1994; Hatfield, 2002). Over the course of psychology’s history, critics have often questioned whether psychology’s lab-based experiments permit the generalization of results beyond the laboratory settings within which these results are typically obtained. In response, many researchers have advocated for more ‘ecologically valid’ experiments, as opposed to the so-called ‘conventional’ laboratory methods (Neisser, 1976; Aanstoos, 1991; Kingstone et al., 2008; Shamay-Tsoory and Mendelsohn, 2019; Osborne-Crowley, 2020). In recent years, several technological advances (e.g., virtual reality, wearable eye trackers, mobile EEG devices, fNIRS, biosensors, etc.) have further galvanized researchers to emphasize the importance of studying human cognition and behavior in the ‘real-world,’ as new technologies will aid researchers in overcoming some of the inherent limitations of laboratory experiments (Schilbach, 2015; Shamay-Tsoory and Mendelsohn, 2019; Sonkusare et al., 2019).

In this article, we will argue that the general aspiration of researchers to understand human cognition and behavior in the ‘real-world’ by conducting experiments that are more ‘ecologically valid’ (henceforth referred to as the ‘real-world approach’) is not without its problems. Most notably, we will argue that the popular term ‘ecological validity,’ which is widely used nowadays by researchers to discuss whether experimental research resembles and generalizes to the ‘real-world,’ is shrouded in both conceptual and methodological confusion. As we ourselves are interested in cognitive and behavioral functioning in the context of people’s everyday experience, and conduct experiments across various ‘laboratory’ and ‘real-world’ environments, we have seen how the uncritical use of the term ‘ecological validity’ can lead to rather misleading and counterproductive discussions. This not only holds for how this concept is used in many scholarly articles and textbooks, but also in presentations and discussions of experimental research at conferences, during the review process, and when talking with students about experimental design and the analysis of evidence.

Although the usage of the term ecological validity has previously been criticized by several scholars (Hammond, 1998; Schmuckler, 2001; cf. Araujo et al., 2007; Dunlosky et al., 2009), we think that these critiques have largely been overlooked. Therefore, it will be necessary to cover some of the same ground. The contribution of this article is threefold. First, we extend the critique of the term ecological validity and apply it to the field of social attention. Second, we scrutinize some of the assumptions that guide the contemporary framework of ecological validity, specifically those regarding artificiality–naturalness and simplicity–complexity. Finally, our article is meant to educate a new generation of students and researchers on the historical roots and conceptual issues of the term ecological validity. This article consists of four parts. First, we will provide a brief history of the so-called ‘real-world or the lab’-dilemma and discuss several definitions and interpretations of the term ecological validity. Second, we will go into the historical roots of the concept of ecological validity and describe how the original meaning of this concept has transformed significantly. Third, we will scrutinize the prevailing assumptions that seem to guide how researchers are currently using the term ecological validity. Finally, we will apply our conceptual analysis to a specific field of study, namely the field of social attention. In recent years, this field has been particularly concerned with issues of ecological validity and generalizability. Therefore, the field of social attention offers an exemplary case to explain how the uncritical use of the terms ‘ecological validity’ and the ‘real-world’ may lead to misleading and counterproductive conclusions.

## A BRIEF HISTORY OF THE ‘REAL-WORLD OR THE LAB’-DILEMMA

The popular story of psychology (or the broader ‘cognitive sciences’) has it that “*psychology became a science by rising from the ‘armchair’ of speculation and uncontrolled observation, and entering the laboratory to undertake controlled observation and measurement*” (Hatfield, 2002, p. 208). The ‘psychologist’s laboratory,’ a special room furnished with all kinds of

lab paraphernalia and sophisticated equipment, has been regarded as the celebrated vehicle of psychology’s journey into sciencehood (Danziger, 1994; Goodwin, 2015). However, despite psychologists’ long tradition of laboratory experimentation (for a history and discussion, see Gillis and Schneider, 1966), there also have been many critical voices saying that psychology’s laboratory experiments are too limited in scope to study how people function in daily life. For example, Brunswik (1943, p. 262) once wrote that experimental psychology was limited to “*narrow-spanning problems of artificially isolated proximal or peripheral technicalities of mediation which are not representative of the larger patterns of life*”. Barker (1968, p. 3) wrote that “*it is impossible to create in the laboratory the frequency, duration, scope and magnitude of some important human conditions.*” Neisser (1976, p. 34) wrote that “*contemporary studies of cognitive processes usually use stimulus material that is abstract, discontinuous, and only marginally real.*” Bronfenbrenner (1977, p. 513) wrote that “*many of these experiments involve situations that are unfamiliar, artificial, and short-lived and that call for unusual behaviors that are difficult to generalize to other settings.*” Kingstone et al. (2008, p. 355) declared that “*the research performed in labs, and the findings they generate, are in principle and in practice unlikely to be of relevance to the more complex situations that people experience in everyday life,*” and Shamay-Tsoory and Mendelsohn (2019, p. 1) stated that “*conventional experimental psychological approaches have mainly focused on investigating behavior of individuals as isolated agents situated in artificial, sensory, and socially deprived environments, limiting our understanding of naturalistic cognitive, emotional, and social phenomena.*”

According to these scholars, psychological science is faced with a gloomy predicament: findings and results based on highly controlled and systematically designed laboratory experiments may not be a great discovery but only a “*mere laboratory curiosity*” (Gibson, 1970, pp. 426–427). As Anderson et al. (1999, p. 3) put it: “*A common truism has been that... laboratory studies are good at telling whether or not some manipulation of an independent variable causes changes in the dependent variable, but many scholars assume that these results do not generalize to the ‘real-world.’*” The general concern is that, due to the ‘artificiality’ and ‘simplicity’ of the laboratory, some (if not many) lab-based experiments do not adequately represent the ‘naturalness’ and ‘complexity’ of psychological phenomena in everyday life (see **Figure 1**). This problem has become familiar to psychologists as the ‘real-world or the lab’-dilemma (Hammond and Stewart, 2001). At the heart of psychology’s ‘real-world or the lab’-dilemma lies a pernicious methodological choice: “*Should it [psychological science] pursue the goal of generality by demanding that research be generalizable to “real life” (aka the “real-world”), or should it pursue generalizability by holding onto its traditional laboratory research paradigm?*” (Hammond and Stewart, 2001, p. 7).

Although psychological science is comprised of many specialized research areas, the goal to understand human cognition and behavior in the ‘real-world’ has become a critically acclaimed goal for psychologists and cognitive scientists of all stripes. Indeed, examples of the ‘real-world or the lab’-dilemma can be found not only in various ‘applied’ fields of psychology,



**FIGURE 1** | Examples of historical and contemporary laboratory rooms and field experiments. **(A)** A laboratory room from the early 20th century. A participant is seated in front of a ‘disc tachistoscope,’ an apparatus to display visual images (adapted from Hilton, 1920). **(B)** A picture of a field experiment by J. J. Gibson. Observers had to judge the size of an object in the distance (adapted from Gibson, 1950). **(C)** A 21st century eye tracking laboratory. A participant is seated in front of a *SMI Hi-Speed* tower-mounted eye tracker (based on Valtakari et al., 2019). **(D)** A wearable eye-tracker (barely visible) is used to measure gaze behavior while participants walked through corridors with human crowds (Hessels et al., 2020). Copyright statement – Panels **(A,B)**. All photographs are used under the provision of the “fair use” U.S. Copyright Act 107 and Dutch Copyright Law Article 15a for non-profit purposes of research, education and scholarly comment. The photograph from W. Hilton’s book: *Applied Psychology: Driving Power of Thought* (Original date of publication, 1920). Retrieved April 1, 2020, from <http://www.gutenberg.org/files/33076/33076-h/33076-h.htm>. The photograph from J. J. Gibson’s book: *The Perception of the Visual World* (Original date of publication, 1950, Figure 74, p. 184) was retrieved from a copy of the Utrecht University library. **(C,D)** Photographs are owned by the authors and the people depicted in the images gave consent for publication.

such as ergonomics (Hoc, 2001), clinical (neuro)psychology (Wilson, 1993; Parsons, 2015), educational psychology (Dunlosky et al., 2009), sport psychology (Davids, 1988), marketing and consumer psychology (Smith et al., 1998), and the psychology of driving (Rogers et al., 2005), but also in the so-called ‘basic’ fields of psychological science, such as the study of perception (Brunswik, 1956; Gibson, 1979/2014), attention (Simons and Levin, 1998; Peelen and Kastner, 2014), memory (Banaji and Crowder, 1989; Neisser, 1991; Cohen and Conway, 2007), social cognition (Schilbach et al., 2013; Schilbach, 2015; Shamay-Tsoory and Mendelsohn, 2019; Osborne-Crowley, 2020), judgment-and-decision making (Koehler, 1996), and child development (Lewkowicz, 2001; Schmuckler, 2001; Adolph, 2019).

## The ‘Real-World Approach’: A Call for Ecological Validity

In the past decades, researchers have often discussed how they may overcome some of the limitations of laboratory-based experiments. Perhaps the largest common denominator of

what we call the ‘real-world approach’ is a strong emphasis on ‘ecological validity.’ Over the past decades, the term ecological validity has made its appearance whenever researchers became concerned with the potential limitations of laboratory experiments (see e.g., Jenkins, 1974; Neisser, 1976; Banaji and Crowder, 1989; Aanstoos, 1991; Koehler, 1996; Smilek et al., 2006; Risko et al., 2012; Schilbach, 2015; Caruana et al., 2017; Shamay-Tsoory and Mendelsohn, 2019; Osborne-Crowley, 2020). As Neisser (1976, p. 33) famously put it:

*“The concept of ecological validity has become familiar to psychologists. It reminds them that the artificial situation created for an experiment may differ from the everyday world in crucial ways. When this is so, the results may be irrelevant to the phenomena that one would really like to explain.”*

The main problem, according to Neisser and many others, is that experiments in psychological science are generally “lacking in ecological validity” (Neisser, 1976, p. 7; Smilek et al., 2006; Shamay-Tsoory and Mendelsohn, 2019; Sonkusare et al., 2019). Aanstoos (1991, p. 77) even referred to this problem as the

“*ecological validity crisis*.” To counter this problem, many researchers have called for studies with ‘more’ or ‘greater’ ecological validity. For example, Koehler (1996, p. 1) advocated for a “*more ecologically valid research program*,” Schilbach (2015, p. 130) argued for “*the inclusion of more ecologically valid conditions*,” and Smilek et al. (2006, p. 104) suggested that “*in order for results to generalize to real-world scenarios we need to use tasks with greater ecological validity*.” Clearly, ecological validity is regarded as an important feature of experimental research by researchers who pursue the ‘real-world approach.’ However, in our opinion, and we are not alone in this regard (see also Hammond, 1998; Araujo et al., 2007; Dunlosky et al., 2009), this notion of ecological validity has caused considerable confusion. To foreshadow some of our criticism of ecological validity, we will show that this concept has largely been detached from its original parentage (cf. Brunswik, 1949), and is now host to different interpretations guided by questionable assumptions (for a history, see Hammond, 1998). Worst of all, the concept is often wielded as a blunt weapon to criticize and dismiss experiments, even though researchers seldom make explicit what definition of ecological validity they use or by which set of criteria they have evaluated a study’s ecological validity (as previously pointed out by Hammond, 1998; Schmuckler, 2001; Dunlosky et al., 2009).

## The Big Umbrella of Ecological Validity

In past decades, the concept of ecological validity has been related to various facets of psychological research, for example, the ecological validity of stimuli (Neisser, 1976; Risko et al., 2012; Jack and Schyns, 2017), the ecological validity of tasks (Smilek et al., 2006; Krakauer et al., 2017), the ecological validity of conditions (Schilbach, 2015; Blanco-Elorrieta and Pyllkänen, 2018), the ecological validity of research settings (Bronfenbrenner, 1977; Schmuckler, 2001), the ecological validity of results (Eaton and Clore, 1975; Greenwald, 1976; Silverstein and Stang, 1976), the ecological validity of theories (Neisser, 1976), the ecological validity of research designs (Rogers et al., 2005), the ecological validity of methods (Banaji and Crowder, 1989), the ecological validity of phenomena (Johnston et al., 2014), the ecological validity of data (Aspland and Gardner, 2003), and the ecological validity of paradigms (Macdonald and Tatler, 2013; Schilbach et al., 2013). However, despite the popular usage of this term, specific definitions and requirements of ecological validity are not always clear.

A closer look at the literature suggests that different definitions and interpretations are used by researchers. Let’s consider some examples of the literature where researchers have been more explicit in their definitions of ecological validity. For example, Ashcraft and Radvansky (2009, p. 511) defined ecological validity as: “*The hotly debated principle that research must resemble the situations and task demands that are characteristic of the real-world rather than rely on artificial laboratory settings and tasks so that results will generalize to the real-world, that is, will have ecological validity*.” Another influential definition of ecological validity was given by Bronfenbrenner (1977), who defined ecological validity as “*the extent to which the environment experienced by the subjects in a scientific investigation has the properties it is supposed or assumed to have by the*

*investigator*” (p. 516). In Bronfenbrenner’s view, a study’s ecological validity should not be predicated on the extent to which the research context resembles or is carried out in a ‘real-life’ environment. Instead, theoretical considerations should guide one’s methodological decisions on what type of research context is most appropriate given one’s focus of inquiry. For example, if one is interested in the behavioral responses of children when they are placed in a ‘strange situation’ then a laboratory room may be adequately suited for that particular research goal. However, if one is interested in how children behave within their home environment, then a laboratory room may not be the most suitable research context. As Bronfenbrenner (1977, p. 516) remarked: “*Specifically, so far as young children are concerned, the results indicate that the strangeness of the laboratory situation tends to increase anxiety and other negative feeling states and to decrease manifestations of social competence*.”

Ecological validity has also been used interchangeably with (or regarded as a necessary component of) ‘external validity’ (Berkowitz and Donnerstein, 1982; Mook, 1983; Hoc, 2001). The concept of external validity typically refers to whether a given study result or conclusion, usually obtained under one set of conditions and with one group of participants, can also be generalized to other people, tasks, and situations (Campbell, 1957). For example, in the literature on neuropsychological assessment and rehabilitation, ecological validity has primarily been conceptualized as “. . . *the degree to which clinical tests of cognitive functioning predict functional impairment*” (Higginson et al., 2000, p. 185). In this field, there has been much discussion about whether the neuropsychological tests used by clinicians accurately predict cognitive and behavioral impairments in everyday life (Heinrichs, 1990; Wilson, 1993). One major concern is that the test materials are either too abstract or too general to adequately represent the kind of problems that people with cognitive and neurological impairments encounter in their daily routines, for example, while cooking or buying food at the supermarket. In response, various efforts have been made to increase the ecological validity of neuropsychological tests, for example, by developing performance measures with relevance for everyday tasks and activities (Shallice and Burgess, 1991; Alderman et al., 2003), by combining and correlating tests results with behavioral observations and self-reports (Wilson, 1993; Higginson et al., 2000), and by using *Virtual Reality* (VR) applications to create test situations in which a patient’s cognitive and functional impairments are likely to be expressed (Parsons, 2015; Parsons et al., 2017).

## THE HISTORICAL ROOTS OF ECOLOGICAL VALIDITY

As we have seen, definitions and interpretations of ecological validity may not only differ among researchers, but also across various subfields within psychology. As such, it is not always clear how the concept should be interpreted. Interestingly, the term ecological validity used to have a very precise meaning when it was first introduced to psychological science by Brunswik (1949,

1952, 1955, 1956). Brunswik coined the term ‘ecological validity’ to describe the correlation between a proximal sensory cue (e.g., retinal stimulation) and a distal object-variable (e.g., object in the environment). In Brunswik’s terminology, ecological validity refers to a measure (a correlation coefficient) that describes a probabilistic relationship between the distal and proximal layers of an organism-environment system. According to Brunswik (1955): “A correlation between ecological variables, one which is capable of standing in this manner as a probability cue for the other, may thus be labeled “ecological validity”” (p. 199). Brunswik (1952) believed psychology to primarily be a science of organism-environment relations in which the “organism has to cope with an environment full of uncertainties” (p. 22). In Brunswik’s ‘lens model’ (Brunswik, 1952), the ecological validities of perceptual cues indicate the potential utility of these cues for the organism to achieve its behavioral goals. Note that Brunswik’s concept of ecological validity is very different from how the term is generally used nowadays, namely to discuss and evaluate whether some laboratory-based experiments resemble and generalize to the ‘real-world’ (cf. Neisser, 1976; Smilek et al., 2006; Ashcraft and Radvansky, 2009; Shamay-Tsoory and Mendelsohn, 2019).

The erosion and distortion of Brunswik’s definition of ecological validity has been documented by several scholars (e.g., Hammond, 1998; Araujo et al., 2007; Holleman et al., in press). As explained by Hammond (1998), the original definition of ecological validity, as Brunswik (1949, 1952) introduced it, has been conflated with Brunswik’s ‘representative design’ of experiments (Brunswik, 1955, 1956). Representative design was Brunswik’s methodological program for psychological science to achieve generalizability of results. To achieve this, researchers should not only conduct proper sampling on the side of the subjects, by sampling subjects who are representative of a specific ‘target population’ (e.g., children, patients), but researchers should also sample stimuli, tasks, and situations which are representative of a specific ‘target ecology.’ As such, an experiment may be treated as a sample of this ‘target ecology.’ By virtue of sampling theory, researchers may then determine whether results can be generalized to the intended conditions. In short, representative design requires researchers to first specify the conditions toward which they intend to generalize their findings, and then specify how those conditions are represented in the experimental arrangement (Brunswik, 1956). For more in-depth discussions on representative design, see Hammond and Stewart (2001); Dhami et al. (2004), and Hogarth (2005).

## A SYSTEMATIC APPROACH TO ECOLOGICAL VALIDITY?

The current lack of terminological precision surrounding ecological validity is, to say the least, problematic. There seems to be no agreed upon definition in the literature, nor any means of classification to determine or evaluate a study’s ecological validity. This seems to be at odds with the relative ease by which researchers routinely invoke this concept to discuss the limitations and shortcomings of laboratory experiments. All the while, researchers seldom make clear how they have

determined a study’s ecological (in)validity. As Schmuckler (2001, p. 419) pointed out: “One consequence of this problem is that concerns with ecological validity can be raised in most experimental situations.” To overcome these problems, several scholars have emphasized the need for a more systematic approach to ecological validity (Lewkowicz, 2001; Schmuckler, 2001; Kingstone et al., 2008; Risko et al., 2012). For example, Lewkowicz (2001, p. 443) wrote that: “What is missing is an independent, objective, and operational definition of the concept of ecological validity that makes it possible to quantify a stimulus or event as more or less ecologically valid.” According to Schmuckler (2001), ecological validity can be evaluated on at least three dimensions: (1) *the nature of the stimuli*; (2) *the nature of task, behavior, or response*; (3) *the nature of the research context*. Researchers have primarily discussed these dimensions in terms of their artificiality–naturalness (e.g., Hoc, 2001; Schmuckler, 2001; Risko et al., 2012; Shamay-Tsoory and Mendelsohn, 2019; Sonkusare et al., 2019), and their simplicity–complexity (e.g., Kingstone et al., 2008; Peelen and Kastner, 2014; Lappi, 2015). As such, a general framework can be construed where stimuli, tasks, behaviors, and research contexts can be evaluated on a continuum of artificiality–naturalness and simplicity–complexity (see also Risko et al., 2012; Lappi, 2015; Shamay-Tsoory and Mendelsohn, 2019; Osborne-Crowley, 2020). At one extreme is the laboratory, characterized by its artificiality and simplicity. At the other extreme is the ‘real-world,’ characterized by its naturalness and complexity. According to this multidimensional framework, researchers may determine a study’s overall ecological validity by combining (e.g., averaging or summing) the main components of ecological validity (i.e., stimuli, tasks/behaviors, research context) in terms of their relative artificiality–naturalness and simplicity–complexity. However, while many researchers have conceptualized ecological validity alongside these dimensions, we think there are several problems to consider. Since the dimensions of this framework are supposedly important to determine the ecological validity of experimental research, this then raises the question of how researchers can judge the artificiality–naturalness and simplicity–complexity of particular experiments. This question will be explored in the following sections.

## Artificiality – Naturalness

The contrast between ‘artificiality’ and ‘naturalness’ is a particularly prominent point of discussion in the ‘real-world or the lab’-dilemma and when researchers talk about the ecological validity of experimental research practices (Hoc, 2001; Kingstone et al., 2008; Shamay-Tsoory and Mendelsohn, 2019). According to Hoc (2001, pp. 282–283), ‘artificial’ situations are “those that are specifically designed for research” and ‘natural’ situations are “the target situations to be understood by research”. Importantly, Hoc (2001) notes that this distinction is made from the perspective of the researcher. However, this artificiality–naturalness distinction should also be considered from the subject’s point of view. For example, according to Sonkusare et al. (2019): “naturalistic paradigms can be heuristically defined as those that employ the rich, multimodal dynamic stimuli that represent our daily lived experience, such as film clips, TV advertisements, news items,

and spoken narratives, or that embody relatively unconstrained interactions with other agents, gaming environments, or virtual realities” (p. 700). Furthermore, researchers have long recognized that artificiality arises when the experimental methods employed by researchers interfere with the naturalness of the psychological phenomena one aims to study. Consequently, there is always an inherent trade-off between the degree of artificiality imposed by the experimental conditions and the naturalness of the phenomena under scientific investigation (Brunswik, 1956; Barker, 1968; Banaji and Crowder, 1989; Kingstone et al., 2008; Risko et al., 2012; Caruana et al., 2017). However, as Winograd (1988) has previously remarked, it remains difficult to “draw a line where artificiality ends and ecological validity... for real events begins” (p. 18).

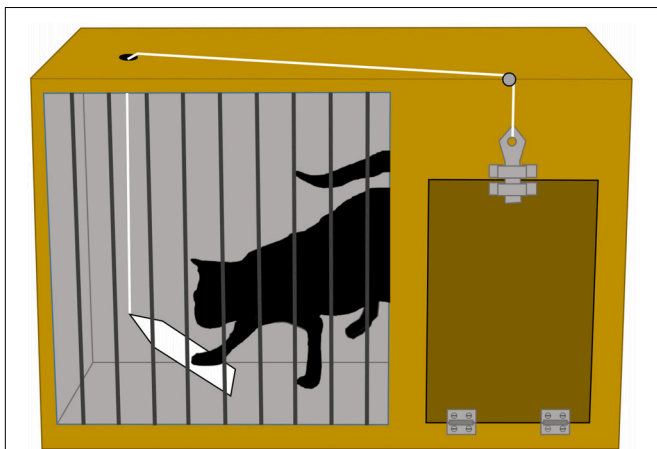
Interestingly, discussions on the naturalness–artificiality of experimental methods have a long pedigree in psychological science. By the end of the 19th century, Thorndike (1899) and Mills (1899) already argued fiercely about what methodology should be favored to study the behavior of cats. Mills dismissed Thorndike’s work because of the artificiality of the experimental methods employed by Thorndike (see **Figure 2**), whereas Thorndike regarded the ethological approach favored by Mills as a collection of uncritical observations and anecdotes. Mills (1899, p. 264) wrote that: “Dr. Thorndike... has given the impression that I have not made experiments, or ‘crucial experiments’... I may remark that a laboratory as ordinarily understood is not well suited for making psychological experiments on animals”. Mills’ point was that: “cats placed in small enclosures... cannot be expected to act naturally. Thus, nothing from about their normal behavior can be determined from their behavior in highly artificial,

abnormal surroundings” (Goodwin, 2015, p. 200). In response to Mills, Thorndike (1899, p. 414) replied: “Professor Mills does not argue in concrete terms, does not criticize concrete unfitness in the situations I devised for the animals. He simply names them unnatural.” Thorndike clearly did not accept Mills’ charge on the artificiality of his experimental arrangements to study the behavior of cats because Mills did not define what should be considered natural behavior in the first place.

We think that this historical discussion between Thorndike and Mills is illuminating, because it characterizes the heart of the discussion on ecological validity nowadays. Namely, what exactly did Mills consider to be ‘natural’ or ‘normal’ behavior? And how did Mills determine that Thorndike’s experiments failed to capture the ‘natural’ behavior of cats? Following Thorndike’s point on the matter, we think that researchers cannot readily determine the naturalness–artificiality of any given experimental arrangement, at least not without specifying what is entailed by these ascriptions. As Dunlosky et al. (2009, p. 431) previously remarked: “A naturalistic setting guarantees nothing, especially given that “naturalistic” is never unpacked – what does it mean?”. Indeed, our survey of the literature also shows that the historical discussion between Thorndike and Mills is by no means a discussion of the past. In fact, we regularly encounter discussions on the ‘artificiality’ and ‘naturalness’ of experimental setups, the presentation of stimuli, the behavior of participants, or the specific tasks and procedures used in experiments – not only in the literature, but also among our colleagues and reviewers. We must often ask for the specifics, because such remarks typically remain undefined by those who toss them around.

## Simplicity – Complexity

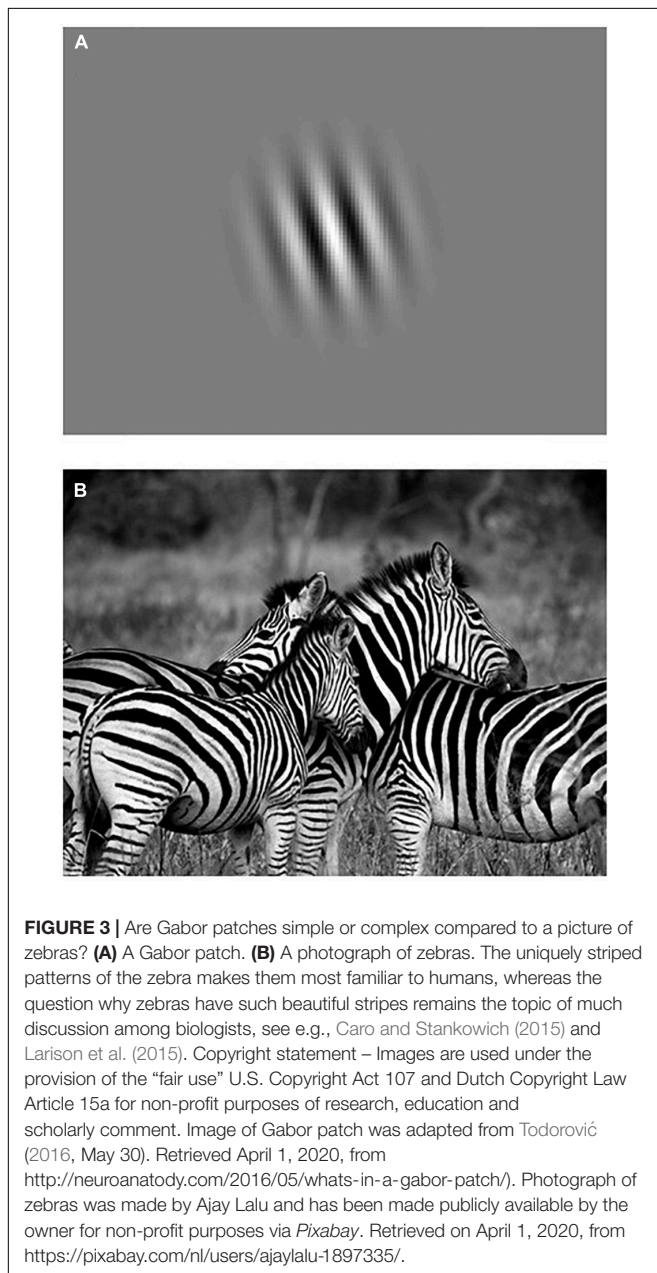
The contemporary framework of ecological validity also posits that the laboratory and the ‘real-world’ are inversely proportional in terms of their simplicity–complexity. Many researchers have lamented that laboratory experiments have a ‘reductionistic’ tendency to simplify the complexity of the psychological phenomena under study (e.g., Neisser, 1976; Kingstone et al., 2008; Shamay-Tsoory and Mendelsohn, 2019; Sonkusare et al., 2019). For example, Sonkusare et al. (2019, p. 699) stated that “the ecological validity of these abstract, laboratory-style experiments is debatable, as in many ways they do not resemble the complexity and dynamics of stimuli and behaviors in real-life.” But what exactly is meant by complexity? Let’s consider some examples from the literature. In the field of social attention, researchers have often used schematic images, photographs and videos of people and social scenes as stimuli to study the cognitive, behavioral, and physiological processes of face perception, gaze following and joint attention (Langton et al., 2000; Frischen et al., 2007; Puce and Bertenthal, 2015). However, in recent years, there has been considerable debate that such stimuli are not ‘ecologically valid’ because they do not “capture the complexity of real social situations” (Birmingham et al., 2012, p. 30). While we agree that looking at a photographic image of a person’s face is different from looking at a living and breathing person, in what ways do these situations differ in complexity? Do these scholars mean that looking at a ‘live’ person is more complex than looking at a picture of that person? Or



**FIGURE 2** | A ‘puzzle box’ devised by Thorndike (1899, 2017) to study learning behavior of cats. A hungry cat is placed in a box which can be opened if the cat pushes a latch. A food reward (‘positive reinforcer’) will be obtained by the cat if it figures out how to escape from the box. Thorndike discovered that after several trials, the time it takes the cat to escape from the box decreases. Experiments with puzzle boxes remain popular today to study the cognitive capacities of animals, for example, see Richter et al. (2016) for a study with octopuses. Copyright statement – Image created and owned by author IH and is based on E. L. Thorndike’s book: *Animal Intelligence* (Original date of publication, 1911, Figure 1, p. 30).

do they mean that the former is more complex than the latter from the perspective of the researcher who wants to understand the cognitive, behavioral, and physiological processes of face perception and social attention?

To take another example, Gabor patches are often used as stimuli by experimental psychologists to study 'low-level visual processing' (see **Figure 3**). Experimental psychologists use Gabor patches as visual stimuli because they offer a high degree of experimental control over various stimulus parameters (e.g., spatial frequency bandwidths, orientation, contrast, size, location). Gabor patches can be described with mathematical precision (i.e., "*Gaussian-windowed sinusoidal gratings*," Fredericksen et al., 1997, p. 1), and their spatial



properties are considered to be a good representation of the receptive field profiles in the primary visual cortex. While Gabor patches may be considered 'simple' to researchers who study the relation between low-level visual processing and neural activity in terms of orientation-tuning and hemodynamic response functions, they also point to the yet to be explained 'complexity' of the many possible relations between other cognitive processes and patterns of neural activity in the brain. On the other hand, a naïve participant (who likely has no clue about what researchers have discovered about low-level visual processing) may describe these Gabor patches as blurry, kind of stripy, zebra-like circles, and think that they are incredibly boring to look at for many trials while lying quietly in a MRI scanner.

Our point here is that simplicity–complexity is in the eye of the beholder. Who is to say what is more simple or complex? Physicists, computer scientists, information theorists, and evolutionary biologists have developed various definitions and measures of complexity (e.g., physical complexity, computational complexity, effective complexity, algorithmic complexity, statistical complexity, structural complexity, functional complexity, etc.), typically expressed in strictly mathematical terms (Edmonds, 1995; Gell-Mann, 1995; Adami, 2002). But what definitions and measures of complexity are used by psychologists and cognitive scientists? Researchers in psychological science seem to have more loosely used the term complexity, for example, to describe a wide range of biological, behavioral, cognitive, social, and cultural phenomena, which typically contain lots of *many's* (i.e., many parts, many variables, many degrees of freedom). Researchers may refer to various phenomena as 'complex' because they are simply not (yet) understood, as in "*the brain is too complex for us to understand*" (Edmonds, 1995, p. 4). Yet, such intuitive notions of complexity, whether they are caused by ignorance or whether they are used to describe something's size, number, or variety (Edmonds, 1995), are not very helpful to evaluate the simplicity–complexity of stimuli, tasks, and situations, nor do such notions provide any formula by which these components can be summed to determine the total ecological validity of a given study. According to Gell-Mann (1995, p. 16):

*"As measures of something like complexity for an entity in the real-world, all such quantities are to some extent context-dependent or even subjective. They depend on the coarse graining (level of detail) of the description of the entity, on the previous knowledge and understanding of the world that is assumed, on the language employed, on the coding method used for conversion from that language into a string of bits, and on the particular idealized computer chosen as a standard."*

## The 'Real World' or the 'Laboratory': Psychology's False Dilemma?

We have discussed several problems with how researchers have used the term 'ecological validity'. In short, the concept of ecological validity has transformed significantly over the past several decades since it was introduced by Brunswik (1949). It has lost most of its former theoretical and methodological cohesion (for a history, see Hammond, 1998), and the definitions and

requirements of ecological validity used by researchers nowadays are seldom made explicit. As such, some experiments may be regarded as ‘ecologically valid’ by one researcher while they can be casually dismissed as ‘ecologically invalid’ by others. A closer look at the literature suggests that many researchers seem to assume that everyone understands what is meant by this term, while in fact the concept of ecological validity is seldom defined. As such, the concept of ecological validity is primarily used nowadays to make hand-waving statements about whether some (lab-based) experiments resemble ‘real life,’ or whether some results obtained in the laboratory may or may not generalize to the ‘real-world.’

In our opinion, the contemporary framework of ecological validity eventually falls short of providing researchers with a tractable research program. Researchers seem to primarily base their judgments of ecological validity upon their own particular theoretical assumptions and considerations about the so-called artificiality–naturalness and simplicity–complexity of experimental situations, typically in the absence of a more formal set of criteria. As such, while we certainly sympathize with the ‘call for ecological validity,’ insofar it has motivated researchers to be critical about the limitations of experimental methods, we also think that the uncritical use of the term ecological validity has caused a lot of confusion, and in some cases has even been counterproductive. Perhaps the most problematic consequence of using the term ecological validity as an easy substitute for the ‘real-world’ was previously pointed out by Hammond (1998). He commented that:

*“There is, of course, no such thing as a “real-world.” It has been assigned no properties, and no definition; it is used simply because of the absence of a theory of tasks or other environments, and thus does not responsibly offer a frame of reference for the generalization”.*

In Hammond’s view, the aim to understand cognitive and behavioral functioning in the ‘real-world’ is basically pointless if one does not first define this notion of the ‘real-world.’ As such, researchers have locked themselves “onto the horns of a false dilemma” (Hammond and Stewart, 2001, p. 7). Thus, in order to talk sensibly about whether some results can also be generalized to particular situations beyond the experimental conditions in which those results were obtained, researchers first need to specify the range and distributions of the variables and conditions to which their results are supposed to be applicable. Since the notion of the ‘real-world’ patently lacks specificity, this phrase inevitably hampers researchers to specify the range and boundary conditions of cognitive and behavioral functioning in any given research context, and thus precludes one from getting at the context-specific and context-generic principles of cognition and behavior (see also Kruglanski, 1975; Simons et al., 2017).

## The Nature of the Environment?

Instead of trying to understand cognitive and behavioral functioning in the ‘real-world,’ we completely agree with Hammond (1998) that the charge of researchers is to always specify and describe the particular context of behavior in which one is interested. Ultimately, the real challenge for researchers is to develop a theory of how specific environmental contexts are related to various forms of cognitive and behavioral functioning.

But what constitutes a psychologist’s theory of the environment? Researchers in psychological science are typically concerned with the nature of the organism, yet, the nature of the environment and its relation to cognitive and behavioral functioning has received considerably less attention from a theoretical point of view (Barker, 1966; Heft, 2013). Interestingly, there have been several scholars who have dedicated themselves to precisely this question, and whose theories of cognition and behavior included a clear perspective on the nature of the environment.

According to Tolman and Brunswik (1935), the nature of the environment, as it appears to the organism, is full of uncertainties. The organism perceives the environment as an array of proximal ‘cues’ and ‘signs’ (i.e., information sources), which are the ‘local representatives’ of various distal objects and events in the organism’s environment. To function more or less efficiently, the organism needs to accumulate, combine, and substitute the information it derives from the available ‘cues’ and ‘signs,’ so that it can adequately adjust its means to achieve its behavioral goals (e.g., finding food or shelter). However, since the environment is inherently probabilistic and only partly predictable, the organism continually needs to adjust its assumptions about the state of the environment based on the available information sources. Another example is given by Barker (1968), whose concept of ‘behavior settings’ (see also Heft, 2001) is key in describing how the environment shapes the frequency and occurrence of human cognition and behavior. Important to behavior settings is that they are the product of the collective actions of a group of individuals. Their geographical location can be specified (e.g., the supermarket, the cinema, etc.), and they have clear temporal and physical boundaries (e.g., opening hours, a door to enter and exit the building). Behavior settings are ‘independent’ of an individual’s subjective experience, yet what goes on inside any behavior setting is characterized by a high degree of interdependency and equivalence of actions between individuals (e.g., most people who are inside a supermarket are shopping for groceries and people in cinemas are watching movies). Another ‘classic’ example of a theory of the environment can be found in J. J. Gibson’s book *The Ecological Approach to Visual Perception* (1979/2014). According to Gibson, there exists a strong mutuality and reciprocity between the organism and its environment. He introduced the concept of ‘affordances’ to explain how the inherent ‘meaning’ of things (i.e., functional significance to the individual) can be directly perceived by an individual perceiver and how this ‘information’ shapes the possibilities for potential actions and experiences. For example, a sufficiently firm and smooth surface may be walk-on-able, run-on-able, or dance-on-able, whereas a rough surface cluttered with obstacles does not afford such actions (Heft, 2001). In short, affordances are properties of an organism-environment system. They are perceiver-relative functional qualities of an object, event or place in the environment and they are dependent on the particular features of the environment and their relationships with the functional capabilities of a particular individual (for more in-depth discussions, see e.g., Heft, 2001; Stoffregen, 2003).

In order to describe and specify the environment and its relation to cognitive and behavioral functioning, we may draw on these scholars to guide us in a more specific direction. While



we do not specifically recommend any of these perspectives, we think they are illuminating because these scholars motivate us to ask questions such as: What is the specific functional context of the cognitive and behavioral processes one is interested in? What are the relevant variables and conditions in this context given one's focus of inquiry and level of analysis? What do we know or assume to know about the range and distribution of these variables and conditions? And how can these variables and conditions be represented in experimental designs to study specific patterns of cognitive and behavioral functioning? In order to answer some these questions, several researchers have emphasized the importance of first observing how people behave in everyday situations prior to experimentation. For example, Kingstone et al. (2008) advocated for an approach called *Cognitive Ethology*, which proposes that researchers should first observe how people behave in everyday situations before moving into the laboratory. In a similar vein, Adolph (2019) proposes that researchers should start with a rich description of the behaviors they are interested in order to first identify the “essential invariants” of these behaviors (p. 187).

## THE FIELD OF SOCIAL ATTENTION: AWAY FROM THE REAL-WORLD AND TOWARD SPECIFICITY ABOUT CONTEXT

To exemplify how some of the ideas outlined above may be useful to researchers, we will apply these ideas to a research topic of our interest: social attention. The field of social attention, as briefly discussed previously, is primarily focused on how attention is influenced by socially relevant objects, events, and situations, most notably, interactions with other social agents. In recent decades, it has been argued extensively that the experimental arrangements used by researchers in this field need more ‘ecological validity’ in order to adequately study the relevant characteristics of social attention in the ‘real-world’ (Risko et al., 2012, 2016; Schilbach et al., 2013; Caruana et al., 2017; Macdonald and Tatler, 2018; Shamay-Tsoory and Mendelsohn, 2019). In the light of these concerns, several researchers have advocated to study “*real-world social attention*” (Risko et al., 2016, p. 1) and “*real-world social interaction*” (Macdonald and Tatler, 2018, p. 1; see also Shamay-Tsoory and Mendelsohn, 2019). One example of this is given by Macdonald and Tatler (2018). In this study, Macdonald and Tatler (2018) investigated how social roles given to participants influenced their social gaze behavior during a collaborative task: baking a cake together. Participants were either not given explicit social roles, or they were given a ‘Chef’ or ‘Gatherer’ role. Macdonald and Tatler (2018) showed that, regardless of whether social roles were assigned or not, participants did not gaze at their cake-baking partners very often while carrying out the task. After comparing their results with other so-called ‘real-world interaction studies’ (e.g., Laidlaw et al., 2011; Wu et al., 2013), the authors stated that: “*we are not able to generalize*

*about the specific amount of partner gaze during any given real-world interaction*” (Macdonald and Tatler, 2018, p. 2171). We think that this statement clearly illustrates how the use of ‘real-world’ and ‘real life’ labels may lead to misleading and potentially counterproductive conclusions, as it seems to imply that ‘real-world interactions’ encompass a clearly defined category of behaviors. However, as argued previously, these so-called ‘real-world interactions’ are *not* a clearly defined category of behaviors. Instead, statements about generalizability need to be considered within a more constrained and carefully defined context (cf. Brunswik, 1956; Simons et al., 2017). This would make it more clear what researchers are talking about instead of subsuming studies under the big umbrella of the ‘real-world.’ For example, if the goal is to study how the cognitive and behavioral processes of social attention are influenced by different contexts and situations, researchers need to specify social gaze behavior as a function of these different contexts and situations.

Thus, instead of studying ‘real-world’ social attention in the context of ‘real-world’ social interactions, researchers should first try to describe and understand cake-baking attention (Macdonald and Tatler, 2018), sharing-a-meal attention (Wu et al., 2013), waiting-room attention (Laidlaw et al., 2011), walking-on-campus attention (Foulsham et al., 2011), Lego-block-building attention (Macdonald and Tatler, 2013), playing-word-games attention (Ho et al., 2015), interviewee-attention (Freeth et al., 2013), and garage-sale attention (Rubo and Gamer, 2018). By doing so, we may begin to understand the context-generic and context-specific aspects of attentional processes, allowing for a more sophisticated theory of social attention. These examples not only show the wide variety of behavioral tasks and contexts that are possible to study in relation to social attention, they also show that uncritical references to ‘ecological validity’ a.k.a. ‘real-worldliness’ are not very helpful to specify the relevant characteristics of particular behavioral contexts.

There are also good examples where researchers have been more explicit about the specific characteristics of social situations that they are interested in. Researchers in the field of social attention have, for example, tried to unravel the different functions of gaze behavior. One important function of gaze behavior is to acquire visual information from the world, however, within a social context, gaze may also signal important information to others which may be used to initiate and facilitate social interaction (see e.g., Gobel et al., 2015; Risko et al., 2016). In a series of experiments, researchers have systematically varied whether, and the degree to which social interaction between two people was possible, and measured how gaze was modulated as a function of the social context (Laidlaw et al., 2011; Gobel et al., 2015; Gregory and Antolin, 2019; Holleman et al., 2020). In other studies, researchers have been explicit about the task-demands and social contexts that elicit specific patterns of gaze behavior, for example, in the context of face-to-face interactions and conversational exchanges (Ho et al., 2015; Hessels et al., 2019). We think that, if researchers would try to be more explicit in their descriptions of task-demands and social contexts in relation to gaze, this may prove to be a solid basis for a more

sophisticated theory of social attention, yet such work remains challenging (for a recent review, see Hessels, in press).

## CONCLUSION

We have argued that the ‘real-world approach’ and its call for ecological validity has several problems. The concept of ecological validity itself is seldom defined and interpretations differ among researchers. We believe that references to ecological validity and the ‘real-world’ can become superfluous if researchers would clearly specify and describe the particular contexts of behavior in which they are interested. This will be a more constructive way to uncover the context-specific and context-generic principles of cognition and behavior. As a final note, we hope that editors and reviewers will safeguard journals from publishing papers where terms

such as ‘ecological validity’ and the ‘real-world’ are used without specification.

## AUTHOR CONTRIBUTIONS

GH and RH drafted the manuscript. RH, IH, and CK edited and revised the manuscript.

## FUNDING

GH and RH were supported by the Consortium on Individual Development (CID). CID is funded through the Gravitation programme of the Dutch Ministry of Education, Culture, and Science and the NWO (grant no. 024.001.003 awarded to author CK).

## REFERENCES

- Aanstoos, C. M. (1991). Experimental psychology and the challenge of real life. *Am. Psychol.* 1:77. doi: 10.1037/0003-066x.46.1.77
- Adami, C. (2002). What is complexity? *Bioessays* 24, 1085–1094.
- Adolph, K. E. (2019). “Ecological validity: mistaking the lab for real life,” in *My Biggest Research Mistake: Adventures and Misadventures in Psychological Research*, Ed. R. Sternberg (New York, NY: Sage), 187–190. doi: 10.4135/9781071802601.n58
- Alderman, N., Burgess, P. W., Knight, C., and Henman, C. (2003). Ecological validity of a simplified version of the multiple errands shopping test. *J. Int. Neuropsychol. Soc.* 9, 31–44. doi: 10.1017/s1355617703910046
- Anderson, C. A., Lindsay, J. J., and Bushman, B. J. (1999). Research in the psychological laboratory: truth or triviality? *Curr. Direct. Psychol. Sci.* 8, 3–9. doi: 10.1111/1467-8721.00002
- Araujo, D., Davids, K., and Passos, P. (2007). Ecological validity, representative design, and correspondence between experimental task constraints and behavioral setting: comment on Rogers, Kadar, and Costall (2005). *Ecol. Psychol.* 19, 69–78. doi: 10.1080/10407410709336951
- Ashcraft, M., and Radvansky, G. (2009). *Cognition*, 5th Edn. Upper Saddle River, NJ: Pearson Education, Inc.
- Aspland, H., and Gardner, F. (2003). Observational measures of parent-child interaction: an introductory review. *Child Adolesc. Mental Health* 8, 136–143. doi: 10.1111/1475-3588.00061
- Banaji, M. R., and Crowder, R. G. (1989). The bankruptcy of everyday memory. *Am. Psychol.* 44:1185. doi: 10.1037/0003-066x.44.9.1185
- Barker, R. G. (1966). “On the nature of the environment,” in *The Psychology of Egon Brunswik*, ed. K. R. Hammond (New York: Holt, Rinehart and Winston).
- Barker, R. G. (1968). *Ecological Psychology: Concepts and Methods for Studying the Environment of Human Behavior*. Stanford, CA: Stanford University Press.
- Berkowitz, L., and Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *Am. Psychol.* 37:245. doi: 10.1037/0003-066x.37.3.245
- Birmingham, E., Ristic, J., and Kingstone, A. (2012). “Investigating social attention: a case for increasing stimulus complexity in the laboratory,” in *Cognitive Neuroscience, Development, and Psychopathology: Typical and Atypical Developmental Trajectories of Attention*, eds J. A. Burack, J. T. Enns, and N. A. Fox (Oxford University Press), 251–276. doi: 10.1093/acprof:oso/9780195315455.003.0010
- Blanco-Elorrieta, E., and Pykkänen, L. (2018). Ecological validity in bilingualism research and the bilingual advantage. *Trends Cogn. Sci.* 22, 1117–1126. doi: 10.1016/j.tics.2018.10.001
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *Am. Psychol.* 32:513. doi: 10.1037/0003-066x.32.7.513
- Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychol. Rev.* 50:255. doi: 10.1037/h0060889
- Brunswik, E. (1949). Remarks on functionalism in perception. *J. Pers.* 18, 56–65. doi: 10.1111/j.1467-6494.1949.tb01233.x
- Brunswik, E. (1952). *The Conceptual Framework of Psychology*. Chicago: University of Chicago Press.
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychol. Rev.* 62:193. doi: 10.1037/h0047470
- Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. Berkeley: University of California Press.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychol. Bull.* 54:297. doi: 10.1037/h0040950
- Caro, T., and Stankowich, T. (2015). Concordance on zebra stripes: a comment on Larison et al. (2015). *R. Soc. Open Sci.* 2:150323. doi: 10.1098/rsos.150323
- Caruana, N., McArthur, G., Woolgar, A., and Brock, J. (2017). Simulating social interactions for the experimental investigation of joint attention. *Neurosci. Biobehav. Rev.* 74, 115–125. doi: 10.1016/j.neubiorev.2016.12.022
- Cohen, G., and Conway, M. A. (2007). *Memory in the Real World*. Abingdon: Psychology Press.
- Danziger, K. (1994). *Constructing the Subject: Historical Origins of Psychological Research*. Cambridge: Cambridge University Press.
- Davids, K. (1988). Ecological validity in understanding sport performance: some problems of definition. *Quest* 40, 126–136. doi: 10.1080/00336297.1988.10483894
- Dhmi, M. K., Hertwig, R., and Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychol. Bull.* 130:959. doi: 10.1037/0033-2909.130.6.959
- Dunlosky, J., Bottiroli, S., and Hartwig, M. (2009). “Sins committed in the name of ecological validity: A call for representative design in education science,” in *Handbook of Metacognition in Education*, eds D. J. Hacker, J. Dunlosky, and A. C. Graesser (Abingdon: Routledge), 442–452.
- Eaton, W. O., and Clore, G. L. (1975). Interracial imitation at a summer camp. *J. Pers. Soc. Psychol.* 32:1099. doi: 10.1037/0022-3514.32.6.1099
- Edmonds, B. (1995). “What is complexity?—the philosophy of complexity per se with application to some examples in evolution,” in *The Evolution of Complexity*, Ed. J. T. Bonner (Dordrecht: Kluwer).
- Foulsham, T., Walker, E., and Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vis. Res.* 51, 1920–1931. doi: 10.1016/j.visres.2011.07.002
- Fredericksen, R., Bex, P. J., and Verstraten, F. A. (1997). How big is a Gabor patch, and why should we care? *JOSA A* 14, 1–12.
- Freeth, M., Foulsham, T., and Kingstone, A. (2013). What affects social attention? Social presence, eye contact and autistic traits. *PLoS One* 8:e53286. doi: 10.1371/journal.pone.0053286
- Frischen, A., Bayliss, A. P., and Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol. Bull.* 133:694. doi: 10.1037/0033-2909.133.4.694

- Gell-Mann, M. (1995). What is complexity? Remarks on simplicity and complexity by the Nobel Prize-winning author of *The Quark and the Jaguar*. *Complexity* 1, 16–19. doi: 10.1002/cplx.6130010105
- Gibson, J. J. (1950). *The Perception of the Visual World*. Cambridge: Houghton Mifflin Company.
- Gibson, J. J. (1970). On the relation between hallucination and perception. *Leonardo* 3, 425–427.
- Gibson, J. J. (2014). *The Ecological Approach to Visual Perception: Classic Edition*. New York, NY: Psychology Press. (Original date of publication 1979).
- Gillis, J., and Schneider, C. (1966). “The historical preconditions of representative design,” in *The Psychology of Egon Brunswik*, ed. K. R. Hammond (New York, NY: Holt, Rinehart & Winston, Inc), 204–236.
- Gobel, M. S., Kim, H. S., and Richardson, D. C. (2015). The dual function of social gaze. *Cognition* 136, 359–364. doi: 10.1016/j.cognition.2014.11.040
- Goodwin, C. J. (2015). *A History of Modern Psychology*, 5 Edn. Hoboken, NJ: John Wiley & Sons.
- Greenwald, A. G. (1976). Within-subjects designs: to use or not to use? *Psychol. Bull.* 83:314. doi: 10.1037/0033-2909.83.2.314
- Gregory, N. J., and Antolin, J. V. (2019). Does social presence or the potential for interaction reduce social gaze in online social scenarios? Introducing the “live lab” paradigm. *Q. J. Exp. Psychol.* 72, 779–791. doi: 10.1177/1747021818772812
- Hammond, K. R. (1998). *Ecological Validity: Then and Now*. Available online at: <http://www.brunswik.org/notes/essay2.html> (accessed April 1, 2020).
- Hammond, K. R., and Stewart, T. R. (2001). *The Essential Brunswik: Beginnings, Explications, Applications*. New York, NY: Oxford University Press.
- Hatfield, G. (2002). Psychology, philosophy, and cognitive science: reflections on the history and philosophy of experimental psychology. *Mind Lang.* 17, 207–232. doi: 10.1111/1468-0017.00196
- Heft, H. (2001). *Ecological Psychology in Context: James Gibson, Roger Barker, and the Legacy of William James’s Radical Empiricism*. Hove: Psychology Press.
- Heft, H. (2013). An ecological approach to psychology. *Rev. Gen. Psychol.* 17, 162–167. doi: 10.1037/a0032928
- Heinrichs, R. W. (1990). Current and emergent applications of neuropsychological assessment: problems of validity and utility. *Prof. Psychol.* 21:171. doi: 10.1037/0735-7028.21.3.171
- Hessels, R. S. (in press). How does gaze to faces support face-to-face interaction? A review and perspective. *Psychonom. Bull. Rev.* doi: 10.31219/osf.io/8zta5
- Hessels, R. S., Holleman, G. A., Kingstone, A., Hooge, I. T. C., and Kemner, C. (2019). Gaze allocation in face-to-face communication is affected primarily by task structure and social context, not stimulus-driven factors. *Cognition* 184, 28–43. doi: 10.1016/j.cognition.2018.12.005
- Hessels, R. S., van Doorn, A. J., Benjamins, J. S., Holleman, G. A., and Hooge, I. T. C. (2020). Task-related gaze control in human crowd navigation. *Atten. Percept. Psychophys.* doi: 10.3758/s13414-019-01952-9 [Online ahead of print]
- Higginson, C. I., Arnett, P. A., and Voss, W. D. (2000). The ecological validity of clinical tests of memory and attention in multiple sclerosis. *Arch. Clin. Neuropsychol.* 15, 185–204. doi: 10.1016/s0887-6177(99)00004-9
- Hilton, W. (1920). *Applied Psychology: Driving Power of Thought*. The Society of Applied Psychology. Available online at: <http://www.gutenberg.org/files/33076/33076-h/33076-h.htm> (accessed April 1, 2020).
- Ho, S., Foulsham, T., and Kingstone, A. (2015). Speaking and listening with the eyes: gaze signaling during dyadic interactions. *PLoS One* 10:e0136905. doi: 10.1371/journal.pone.0136905
- Hoc, J.-M. (2001). Towards ecological validity of research in cognitive ergonomics. *Theor. Issues Ergon. Sci.* 2, 278–288. doi: 10.1371/journal.pone.0184488
- Hogarth, R. M. (2005). The challenge of representative design in psychology and economics. *J. Econ. Methodol.* 12, 253–263. doi: 10.1177/0269216311399663
- Holleman, G. A., Hessels, R. S., Kemner, C., and Hooge, I. T. C. (2020). Implying social interaction and its influence on gaze behavior to the eyes. *PLoS One* 15:e0229203. doi: 10.1371/journal.pone.0229203
- Holleman, G. A., Hooge, I. T. C., Kemner, C., and Hessels, R. S. (in press). The reality of ‘real-life’ neuroscience: a commentary on Shamay-Tsoory & Mendelsohn. *Perspect. Psychol. Sci.*
- Jack, R. E., and Schyns, P. G. (2017). Toward a social psychophysics of face communication. *Annu. Rev. Psychol.* 68, 269–297. doi: 10.1146/annurev-psych-010416-044242
- Jenkins, J. J. (1974). Remember that old theory of memory? Well, forget it. *Am. Psychol.* 29:785. doi: 10.1037/h0037399
- Johnston, P., Molyneux, R., and Young, A. W. (2014). The N170 observed ‘in the wild’: robust event-related potentials to faces in cluttered dynamic visual scenes. *Soc. Cogn. Affect. Neurosci.* 10, 938–944. doi: 10.1093/scan/nsu136
- Kingstone, A., Smilek, D., and Eastwood, J. D. (2008). Cognitive ethology: a new approach for studying human cognition. *Br. J. Psychol.* 99, 317–340. doi: 10.1348/000712607x251243
- Koehler, J. J. (1996). The base rate fallacy reconsidered: descriptive, normative, and methodological challenges. *Behav. Brain Sci.* 19, 1–17. doi: 10.1017/s0140525x00041157
- Krakauer, J. W., Ghazanfar, A. A., Gomez-Marín, A., MacIver, M. A., and Poeppel, D. (2017). Neuroscience needs behavior: correcting a reductionist bias. *Neuron* 93, 480–490. doi: 10.1016/j.neuron.2016.12.041
- Kruglanski, A. W. (1975). The two meanings of external invalidity. *Hum. Relat.* 28, 653–659. doi: 10.1177/001872677502800704
- Laidlaw, K. E., Foulsham, T., Kuhn, G., and Kingstone, A. (2011). Potential social interactions are important to social attention. *Proc. Natl. Acad. Sci. U.S.A.* 108, 5548–5553. doi: 10.1073/pnas.1017022108
- Langton, S. R., Watt, R. J., and Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends Cogn. Sci.* 4, 50–59. doi: 10.1016/s1364-6613(99)01436-9
- Lappi, O. (2015). Eye tracking in the wild: the good, the bad and the ugly. *J. Eye Mov. Res.* 8:1. doi: 10.1016/j.dcn.2019.100710
- Larison, B., Harrigan, R. J., Thomassen, H. A., Rubenstein, D. I., Chan-Golston, A. M., Li, E., et al. (2015). How the zebra got its stripes: a problem with too many solutions. *R. Soc. Open Science* 2:140452. doi: 10.1098/rsos.140452
- Lewkowicz, D. J. (2001). The concept of ecological validity: what are its limitations and is it bad to be invalid? *Infancy* 2, 437–450. doi: 10.1207/s15327078in0204\_03
- Macdonald, R. G., and Tatler, B. W. (2013). Do as eye say: gaze cueing and language in a real-world social interaction. *J. Vis.* 13, 1–12. doi: 10.1167/13.4.6
- Macdonald, R. G., and Tatler, B. W. (2018). Gaze in a real-world social interaction: a dual eye-tracking study. *Q. J. Exp. Psychol.* 71, 2162–2173. doi: 10.1177/1747021817739221
- Mills, W. (1899). The nature of animal intelligence and the methods of investigating it. *Psychol. Rev.* 6:262. doi: 10.1037/h0074808
- Mook, D. G. (1983). In defense of external invalidity. *Am. Psychol.* 38:379. doi: 10.1037/0003-066x.38.4.379
- Neisser, U. (1976). *Cognition and Reality: Principles and Implications Of Cognitive Psychology*. San Francisco, CA: W. H. Freeman and Company.
- Neisser, U. (1991). A case of misplaced nostalgia. *Am. Psychol.* 46:34–36. doi: 10.1037/0003-066x.46.1.34
- Osborne-Crowley, K. (2020). Social cognition in the real world: reconnecting the study of social cognition with social reality. *Rev. Gen. Psychol.* 1–15. doi: 10.4324/9781315648156-1
- Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Front. Hum. Neurosci.* 9:660. doi: 10.3389/fnhum.2015.00660
- Parsons, T. D., Carlew, A. R., Magtoto, J., and Stonecipher, K. (2017). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychol. Rehabil.* 27, 777–807. doi: 10.1080/09602011.2015.1109524
- Peelen, M. V., and Kastner, S. (2014). Attention in the real world: toward understanding its neural basis. *Trends Cogn. Sci.* 18, 242–250. doi: 10.1016/j.tics.2014.02.004
- Puce, A., and Bertenthal, B. I. (2015). *The Many Faces of Social Attention: Behavioral and Neural Measures*, eds A. Puce and B. I. Bertenthal (Switzerland: Springer).
- Richter, J. N., Hochner, B., and Kuba, M. J. (2016). Pull or push? Octopuses solve a puzzle problem. *PLoS One* 11:e0152048. doi: 10.1371/journal.pone.0152048
- Risko, E. F., Laidlaw, K., Freeth, M., Foulsham, T., and Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Front. Hum. Neurosci.* 6:143. doi: 10.3389/fnhum.2012.00143
- Risko, E. F., Richardson, D. C., and Kingstone, A. (2016). Breaking the fourth wall of cognitive science: real-world social attention and the dual function of gaze. *Curr. Direct. Psychol. Sci.* 25, 70–74. doi: 10.1177/0963721415617806

- Rogers, S. D., Kadar, E. E., and Costall, A. (2005). Gaze patterns in the visual control of straight-road driving and braking as a function of speed and expertise. *Ecol. Psychol.* 17, 19–38. doi: 10.1207/s15326969eco1701\_2
- Rubo, M., and Gamer, M. (2018). “Virtual reality as a proxy for real-life social attention?” Paper presented at the Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications. New York, NY.
- Schilbach, L. (2015). Eye to eye, face to face and brain to brain: novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. *Curr. Opin. Behav. Sci.* 3, 130–135. doi: 10.1016/j.cobeha.2015.03.006
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2013). Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. doi: 10.1017/s0140525x12000660
- Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy* 2, 419–436. doi: 10.1207/s15327078in0204\_02
- Shallice, T., and Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain* 114, 727–741. doi: 10.1093/brain/114.2.727
- Shamay-Tsoory, S. G., and Mendelsohn, A. (2019). Real-life neuroscience: an ecological approach to brain and behavior research. *Perspect. Psychol. Sci.* 14, 841–859. doi: 10.1177/1745691619856350
- Silverstein, C. H., and Stang, D. J. (1976). Seating position and interaction in triads: a field study. *Sociometry* 39, 166–170.
- Simons, D. J., and Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonom. Bull. Rev.* 5, 644–649. doi: 10.3758/bf03208840
- Simons, D. J., Shoda, Y., and Lindsay, D. S. (2017). Constraints on generality (COG): a proposed addition to all empirical papers. *Perspect. Psychol. Sci.* 12, 1123–1128. doi: 10.1177/1745691617708630
- Smilek, D., Birmingham, E., Cameron, D., Bischof, W., and Kingstone, A. (2006). Cognitive ethology and exploring attention in real-world scenes. *Brain Res.* 1080, 101–119. doi: 10.1016/j.brainres.2005.12.090
- Smith, P. W., Feinberg, R. A., and Burns, D. J. (1998). An examination of classical conditioning principles in an ecologically valid advertising context. *J. Market. Theor. Pract.* 6, 63–72. doi: 10.1080/10696679.1998.11501789
- Sonkusare, S., Breakspear, M., and Guo, C. (2019). Naturalistic stimuli in neuroscience: critically acclaimed. *Trends Cogn. Sci.* 23, 699–714. doi: 10.1016/j.tics.2019.05.004
- Stoffregen, T. A. (2003). Affordances as properties of the animal-environment system. *Ecol. Psychol.* 15, 115–134. doi: 10.1016/j.humov.2019.01.002
- Thorndike, E. (1899). A reply to “The nature of animal intelligence and the methods of investigating it”. *Psychol. Rev.* 6, 412–420. doi: 10.1037/h0073289
- Thorndike, E. (2017). *Animal Intelligence: Experimental Studies*. Abingdon: Routledge.
- Todorović, A. (2016) *What’s in a Gabor Patch?*, Vol. 2016. Available online at: <http://neuroanatomy.com/2016/05/whats-in-a-gabor-patch/> (accessed April 1, 2020).
- Tolman, E. C., and Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychol. Rev.* 42:43. doi: 10.1037/h0062156
- Valtakari, N. V., Hooge, I. T. C., Benjamins, J. S., and Keizer, A. (2019). An eye-tracking approach to Autonomous sensory meridian response (ASMR): the physiology and nature of tingles in relation to the pupil. *PLoS One* 14:e226692. doi: 10.1371/journal.pone.0226692
- Wilson, B. A. (1993). Ecological validity of neuropsychological assessment: do neuropsychological indexes predict performance in everyday activities? *Appl. Prevent. Psychol.* 2, 209–215. doi: 10.1016/s0962-1849(05)80091-5
- Winograd, E. (1988). “Continuities between ecological and laboratory approaches to memory,” in *Emory Symposia in Cognition, 2. Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory* eds U. Neisser and E. Winograd (Cambridge: Cambridge University Press), 11–20. doi: 10.1017/cbo9780511664014.003
- Wu, D. W.-L., Bischof, W. F., and Kingstone, A. (2013). Looking while eating: the importance of social context to social attention. *Sci. Rep.* 3:2356. doi: 10.1038/srep02356

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Holleman, Hooge, Kemmer and Hessels. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.