Neuroscience and Biobehavioral Reviews 155 (2023) 105463

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Can cognitive neuroscience solve the lab-dilemma by going wild?

William Vallet[a,b,*,1], Virginie van Wassenhove[a]

[a] CEA DRF/Joliot, NeuroSpin, INSERM, Cognitive Neuroimaging Unit, Université Paris Saclay, 91191 Gif-sur-Yvette, France
[b] INSERM U1028, CNRS UMR 5292, PSYR2 Team, Centre de recherche en Neurosciences de Lyon (CRNL), Université Lyon 1, 69000 Lyon, France

ARTICLE INFO

Keywords:
Behavioral sciences
Ecological neurosciences
Experimental design
In situ design
methods

ABSTRACT

Reproducibility, measurability, and refutability are the foundation of the scientific method applied to empirical work. In the study of animal and human behavior, experimental protocols conducted in the lab are the most reliable means by which scientists can operationalize behaviors using controlled and parameterized setups. However, whether observations in the lab fully generalize in the real world remain legitimately disputed. The notion of "experimental design" was originally intended to ensure the generalizability of experimental findings to real-world situations. Experiments in the wild are more frequently explored and significant technological advances have been made allowing mobile neuroimaging. Yet some methodological limitations remain when testing scientific hypotheses in ecological conditions. Herein, we discuss the limitations of inferential processes derive from empirical observations in the wild. The multi-causal property of an ecological situation often lacks controls, and this major concern may prevent the replication and the reliability of behavioral observations. We discuss the epistemological and historical grounds of the induction process for behavioral and cognitive neurosciences and provide some possible heuristics for In situ experimental designs compatible with psychophysics in the wild.

1. Introduction

Over the years, the need to conduct experiments outside typical lab settings have become critical to make perceptual and cognitive neurosciences more ecologically valid. The impetus for ecological validity of experimental designs has enabled significant advances in human cognitive neuroscience. It has also facilitated the development of recording techniques and methods adapted to an experimental terrain in the wild. For instance, we can now use bio-sensing and record brain activity (Bateson et al., 2017; Boto et al., 2018; Siddharth et al., 2019) or even whole body imaging (Makeig et al., 2009) outside the lab. The ability to record temporally-sensitive signals with non-invasive brain imaging is made possible by mobile electroencephalography (mEEG), which can be recorded outside your typical Faraday cage and still provide reliable brain signals when participants are moving (Gramann et al., 2014; Makeig et al., 2009). Online experimentation has also blossomed and is reliably explored in a wide range of behavioral research areas (Anwyl-Irvine et al., 2021; Chaumon et al., 2022; Zhao et al., 2022).

These new experimental methods have expanded the possibilities of quantitative research fields and built the momentum for the emergence of innovative explorations of various cognitive processes in the wild (De Sanctis et al., 2021). The range of cognitive processes that can be, and has been, explored in real-life situations has grown quite impressively with recent studies addressing motor control (Mustile et al., 2021), cognitive-motor interference (Liebherr et al., 2021; Nenna et al., 2021; Reiser et al., 2021), performance monitoring (Lange and Osinsky, 2021), spatial navigation (Delaux et al., 2021; Mavros et al., 2016; Wunderlich and Gramann, 2021), attention (De Vos et al., 2014; Debener et al., 2012; Hölle et al., 2021; Kingstone et al., 2003; Ladouce et al., 2019; Piñeyro Salvidegoitia et al., 2019) but also inter-personal synchronization with important implications for social studies and neuro-education (Bevilacqua et al., 2019; Dikker et al., 2017; Poulsen et al., 2017).

Such a wide range of research questions attests that a new way of experimenting in the wild is growing and becoming ripe for ecological behavioral studies, providing an excellent potential for experimental psychology and cognitive neurosciences. However, the ability to test outside the lab also raises a canonical and essential question that is well-captured by the concept of the "lab-dilemma" (Hammond and Stewart, 2001; Holleman et al., 2020). The central challenge of the lab-dilemma

* Correspondence to: CEA, DRF/Joliot, Neurospin, Bâtiment 145 PC 156, 91191 Gif-sur-Yvette, France.
E-mail address: William.vallet@inserm.fr (W. Vallet).
[1] ORCID ID: https://orcid.org/0000-0002-3325-4360

https://doi.org/10.1016/j.neubiorev.2023.105463
Received 1 September 2023; Received in revised form 18 October 2023; Accepted 8 November 2023
Available online 13 November 2023
0149-7634/

is to ask how scientists can effectively combine ecological validity with well-controlled experimental designs when experimenting in the real-world.

The focus of the current paper is to discuss the possibility of incorporating ecological factors within the experimental design tested in the wild, allowing to enrich the experimental practice while ensuring that reliable inferences be made as to the causal factor on the ecological behavior. We will call this experimental approach the In situ design: the intent is to offer an innovative solution to enrich current experimental designs and exploit the testing taking place in the wild as opposed to simply translating the lab design into the real-world. Put simply, the idea is to adapt the experimental design to the real world situation and integrate an ecological variable of a situation to benefit the experimental inquiry. We will first introduce the history behind the lab-dilemma issue in psychology extended to cognitive neurosciences. We then illustrate this problem with a concrete example of setting up a study design in train stations and high-speed trains (Wiltdimes, 2018) and discuss how to fulfill the three requirements of scientific experimenting (reproducibility, measurability, refutability) when adapting a lab experiment to the wild using an In situ design.

2. What is the lab-dilemma?

The lab dilemma raises the problem of the generalizability of a given (animal or human) behavior in contemporary scientific approaches of behavioral sciences - encompassing cognitive neurosciences, experimental psychology, and perceptual psychophysics (Yarkoni, 2022). Since its inception, experimental psychology has demonstrated that experiments in the lab met the scientific criteria, with the application of sophisticated statistical methods on experimental data acquired in dedicated research settings and under highly controlled paradigms. The purpose of experimentation in psychology is to demonstrate the operationalization of human behavior to enable its measurements. However, while lab experimentation is considered a scientific standard for a large community of behavioral researchers, it also comes with significant limitations: the lab itself. The radical shift of experimental psychology from its original focus - explaining natural human behaviors - to behavior under highly controlled settings raised the question of generalizability of lab results. As a trade-off to its scientific rigor in the lab, experimental psychology has to demonstrate its capacity to generalize constrained lab results to human psychology in the real world (Hammond and Stewart, 2001). The questions raised by the lab-dilemma affect human (Kingstone et al., 2003, 2008) and animal (Gomez-Marin et al., 2014) research in similar ways, and behavioral (neuro)sciences as a whole.

2.1. Fisher's experimental design and the misuse of the induction process in experimental psychology

In the 1940 s, to adapt to the imperative request of generalization, experimental psychologists turned to Fisher's theory of factorial design for their experiments (Fisher, 1936; Fisher et al., 1926). The factorial design described in the first article, "*The arrangement of field experiment*", originally aimed at improving the agriculturist's control over farming parameters such as manured vs. unmanured, chloride vs. sulfate, or early vs. late manure applications. From a scientific perspective, an experimental design corresponds to the systematic preparation in assigning an experimental measurement to the levels of treatment. A design allows performing statistical analyses of the measurements from which the validity of a conclusion can be drawn (Kirk, 1995; Smith, 2000). Fisher's theory proposed a foundational framework from which modern experimental design in experimental psychology has emerged.

However, the limitations of inductive inference led to a shortcoming of the generalization process between the psychologist and the agriculturist (Brunswik, 1943, 1952). The agriculturist controls the parameters (e.g., the process of cultivation in a given field) to match his method with the circumstances of discovery. Indeed, inductive inference over uncontrollable parameters is not required in the agriculturist's research to achieve the generalization process. The uselessness of induction (of uncontrollable parameters such as the weather) can be seen in the fact that the testing situations fit with the real-world context that the agriculturist encounters (e.g., the cultivation of a particular fruit occurs at a specific time during the year and in a particular region).

To the contrary, the psychologist has very few options to choose from in designing an experiment. The ecological situations include an infinity of possible contexts that could virtually lead to the same behavior. To solve this difficulty, psychologists use inductive inference on a sample of participants assumed to be representative of a population to warrant the generalizability of their results (Brunswik, 1955). However, the induction process following experimentation does not generalize the results to the diversity of possible situations in the real world; instead, the induction process relies on inferences informed by statistics derived from the selected sample of participants. Thus, the induction is restricted to generalizing the results collected in lab settings to the general population. In short, making inductive inferences from the lab to the real world appears impossible, and the generalizability of observations is unwarranted.

Consequently, the limited validity of inductive inference towards real world situations in experimental psychology depends on the implicit quality of the controlled situations and how well it may (or not) represent the diversity of situations encountered in the natural habitat of human (or animal) lives. Furthermore, when studying a targeted behavior within an experimental design, it is crucial to recognize that the amount of variance captured by the experiment may not necessarily reflect or overlap the variance observed in real-world behavior.

Egon Brunswik formulated the fundamental limitation of experimental psychology as the double standard, which explained the logic of the induction process applied to the individual or participant but not to the environment (Brunswik, 1943). To solve this fundamental issue in experimental psychology, Brunswik introduced a novel approach and experimental design, improving the possibility of generalizing results obtained in the lab to the real world.

2.2. Brunswik's ecological validity is a logical turn

In the early 1950 s, Brunswik introduced the notion of ecological validity to the field of experimental psychology (Brunswik, 1952). In his theoretical framework, the central proposal was to dissociate the classic *experimental* design from the *representative* design. In the representative design, sampling stimuli from the environment, or artificial stimuli in which environmental properties were preserved, is a fundamental prerequisite for capturing psychological processes. Fundamental principle of the representative design is the rule that one may generalize the results from experiment only to those circumstances or objects that have been sampled in the design.

For Brunswick, goal-directed behaviors are, by default, adapted to the environment. Furthermore, the environment includes multicausality and probabilistic relations between variables of interest - i.e., a probabilistic functionalism. Therefore, he argued that the design of experimental tasks should emulate as close as possible the ecological settings towards which the generalization was initially intended (Araujo, Davids, 2009; Dhami et al., 2004; Hammond and Stewart, 2001). In this representative design, ecological validity can be assessed if psychologists take into account the correlations between the relation of the organism to proximal cues (e.g., the processing of acoustic stimulations) and the objects indicated by distal cues e.g., the visual stimuli appearing in the surrounding environment- (Koffka, 1936). He states, *"Any fairly consistent rapport, be it intuitively perceptual or explicitly rational, with distal layers of the environment presupposes the existence of proximal sensory cues of some degree of ecological validity to serve as mediators of the relationship"* (Brunswik, 1956, p.48).

For Brunswik, the observer has an uncertain access to the distal object in the world. Therefore, the observer's attending can only infer the existence of the distal object to the appropriate local sensory cue. This means that the observer should preferentially attend to the valid sources of information about their objects in the world or sensory cues that are "ecologically valid".

In the representative design, the statistical analysis of distal-proximal correlations, where one stimulus stands as a probability cue for the other stimulus, would be termed psychological ecology. Additionally, the strength of the correlation – i.e., the probability that one stimulus predicts the other in the environment – is coined ecological validity (Brunswik, 1952). Hence, in Brunswik's theoretical framework, the crucial take-home message of the representative design is that sampling situations and stimuli determine the target ecological situation in which generalization is intended to: a sample of the situations in an experiment thus becomes a unique sample of the target ecological situation (Holleman et al., 2020). Crucially, the representative design changes the logic of the inductive inference in experimental psychology and provides a generalization process. Indeed, the representative design confers the ecological value not to the representative sampling of participants as in classical experimental design but to a representative sampling of a situation in the inference process (Brunswik, 1943; Schmuckler, 2001).

With his work, Brunswik raised awareness in experimental psychology on the issue of the generalization process when using Fisher's experimental design. The representative design is highly appealing but in its application, extracting the formal properties of the ecological situation and operationalizing them with behavior is a complex challenge (Araújo et al., 2007). Furthermore, the inter-correlations among cues assuming ecological validity prevent the experimenters from independently manipulating the cues' co-occurrence in the real world (Massaro, 2014). Because of these concrete difficulties, very few researchers claim to be Brunswikian in their approach, apart from a few studies in learning and decision-making (Steiner and Frey, 2021). As for our current understanding of ecological validity, psychologists mostly elude representative designs since they do not integrate into their experimental tasks the sampling variables from the environment and a measure of natural statistics (i.e., correlation between cues (Hammond, 1996)). Therefore, ecological validity has been used to cover a variety of concept. For example, the notion of external validity (Campbell & Stanley, 2015) refers only to the generalization process from the experimental study to a larger population (Pinder et al., 2011), the capacity of some tasks in a study to reflect real-world situations (Ashcraft & Radvansky, 2009) and more largely to the generalization of experimental findings toward the real world. In the following lines, we use the term "ecological validity" in line with Brunswik's definition. More precisely, we state that information generated by sources from the real world (e.g., the optical flow) shares a high probabilistic value to predict another stimulus in the real-world (e.g., time to contact with an approaching object). Following this definition, we assume three roots of the terminology regrouping: The validity of the source, the probability that one stimulus predicts another and the capacity to generalize the results.

2.3. Where do we stand?

The central issue of generalization described in the lab dilemma is alive and well today in all fields using the empirical approach of experimental psychology, including modern cognitive neurosciences. As it stands, the experimental design does not fully ensure the replicability of real-world experiments: one main reason is that the ecological variable provided by the experimental context (aka the real world) does not integrate the experimental design as a factorized and controlled variable. In most experiments, the context is used as a group factor to test a laboratory effect in the real world but is not actually incorporated in the design.

Critics have often questioned whether experimental results observed in the lab truly allow for a generalizable comprehension of ecological behavior. As Neisser (1976) or Wong and Bronfenbrenner (1977) put it, the risk of assessing behavior under artificial lab conditions is that outcomes may radically differ from everyday life. This limitation of modern cognitive neuroscience is clearly demonstrated by studies on economic decision-making. To comprehend decision-making in real-life situations involving risk, it is essential for researchers to ascertain concern whether the decision-making processes observed in economic experiments truly mirror real-life risky decision-making behavior. Without addressing this, the findings of the study are irrelevant in terms of simulating the ecological context it aims to replicate. The Balloon Analogue Risk Task (BART (Lejuez et al., 2002)) is widely regarded as the gold standard for evaluating individual differences in real-life risky decision-making. However, its practical application is hindered by its subpar psychometric properties, including low convergent validity and test-retest reliability. As a result, its association with real-life risky decision is also compromised (Ju and Wallraven, 2023; Pleskac et al., 2008; Steiner and Frey, 2021). An additional illustration of how laboratory findings may not apply to real-world situations can be observed in electrophysiological data, specifically in terms of response reliability. It is widely acknowledged that natural stimulation elicits remarkably consistent and synchronized activity in regions that exhibit no modulation of response in controlled laboratory experiments (Hasson et al., 2011; Eisenberg et al., 2019; David et al., 2004).

Moreover, in humans and animals alike, the combination of task goals in lab settings may have such insignificant ecological validity that they may represent neural responses never used in natural behaviors (Gomez-Marin et al., 2014; Krakauer et al., 2017; Wong and Bronfenbrenner, 1977). In their view, the lab approach could only enhance our understanding of behavior under particular artificial circumstances.

The dichotomy between lab conditions (highly controlled but ecologically invalid) and ecological conditions (challenging to operationalize and control for but ecologically valid) could lead experimental research into a possible deadlock. To remedy to this difficulty, some research is heading towards new experimental approaches using virtual reality (e.g., Miller et al., 2019) naturalistic task settings (Redcay and Moraczewski, 2020) or testing behaviors in real world contexts through online tasks (Chaumon et al., 2022; Rogers, 2021). In parallel, new technical advances such as mEEG allow testing for the complexity of an ecological situation with a naturalistic behavior.

As underlined by Nastase (2021); Nastase et al. (2020), new ways of thinking have emerged in modern cognitive neurosciences to build relevant brain and behavior models for our understanding of complex behaviors in the real world. One possibility is to progress by incorporating and adapting Brunswik's theory and the original ecological validity into lab experimentation. Another possibility is to build new ways of experimenting with modern techniques and experimental data science (e.g., automatic method modeling or machine learning models). These approaches offer the potential to overcome some limitations of the traditional experimental approach, which relies on statistical inference from a small sample size. As previously mentioned, the concept of inductive inference is crucial in the experimental paradigm of behavioral science. However, it is worth noting that statistical inference based on data collected within a sample does not enable generalization or prediction in real-world scenarios. At most, we can only conclude that a sample can reproduce the same behavior in the same situation, but we cannot assume that the observed behavior will be replicated in the multitude of real-world situations. Therefore, classical inference, as understood within the experimental paradigm, does not guarantee accurate predictions outside of the sample (Yarkoni, 2022). Additionally, data modeling is typically limited to smaller datasets collected from a limited number of sources, while complex and large datasets require automated methods such as algorithmic modeling (Breiman, 2001). To overcome these inherent limitations of the experimental paradigm, the research community has made significant progress in understanding brain function using machine learning tools, which offer numerous advantages. The major improvement is the ability to surpass the constraints of statistical inference based on a sample and to eliminate the need for control inherent in experimental design. Instead, the machine learning approach strives to gather a significant amount of unrestricted training data that faithfully captures the complexities of the real world and allows prediction out-of-sample (Nastase et al., 2020). Based on this approach, machine learning could create models that possess a high degree of predictive power for real-world phenomena. However, the machine learning approach also has its limitations due to the methods used and the challenges associated with interpreting the results. Firstly, it is crucial that the source of information holds ecological value; otherwise, the predictive value becomes detached from the real world. This scenario is akin to applying machine learning in experimental situations, where the classifier decodes data based on the experiment's parameters. In such cases, the strength of machine learning lies in its ability to solely explain the targeted phenomena within the experimental context. Secondly, if searcher fails to identify the sources of information utilized by the classifier to decode the data, he will gain no knowledge about the real-world phenomena targeted. Consequently, the effectiveness of this data-driven approach depends solely on the researcher's ability to identify, a posteriori, the specific information and its combinations within the sensors that drive the classification (Carlson et al., 2018).

In complement to the representative design and the machine learning approach, a complementary and transitioning solution for real world experimentation is necessary, which can respond to two primary scientific criteria: the reproducibility and the reliability of behavioral and brain measurements. Improving real world experimentation through the adaptation of the experimental design may be the most practical way for cognitive neuroscience because of the possibility of understanding behavior and the brain in situ, i.e. by going directly beyond the reductionism bias and misusing the induction process in lab experiments. Furthermore, a significant distinction is arising between the machine learning approach and the application of experimental design in real-world scenarios. The former prioritizes correlation over causality and relies on inductive inference through data-driven methods. Conversely, the latter adheres to the traditional hypothetico-deductive approach, aiming to comprehend phenomena through causal relationships (Kitchin, 2014).

Despite such progress, the ecological situation of most studies taking place in the real world and published in the last decade tends to be set aside. The implication of multi-causal variables in the experimental design are set aside, reduced to a minimum and unaccounted for. Multi-causality herein refers to the fixed and identified number of variables in a given situation that have the potential to modulate a behavior. This definition applies well to the lab in which the situation can be fully controlled but does not fit the real world, in which since the entire set of possible variables affecting the targeted behavior cannot be listed. Alternatively, multi-causality can be seen as the features of a given variable acting on a targeted behavior. For example, in the real-world, the impact of an optical flow on behavior has multi-causal properties that are: speed, luminosity, and contrast, vestibular or haptic signals. In isolation, these properties can be independent variables. This conception of multi-causality fits with behavioral science's requisites and is consistent with the imperative of sampling situation proposed by Brunswik.

For instance, most studies incorporate the ecological situation into their design as context, amounting to contrasting the external environment with the lab environment (De Vos et al., 2014; Edwards and Trujillo, 2021; Scanlon et al., 2020; Zink et al., 2016) while a subset of studies do integrate the ecological situations (or the different external environments) as an ecological context variable (Aspinall et al., 2015; Piñeyro Salvidegoitia et al., 2019; Reiser et al., 2019; Scanlon et al., 2020; Shiffman et al., 2008). Scanlon and collaborators (2020) designed an experiment in two different outdoor environments: they asked participants to perform an oddball task while cycling outdoors in a quiet park or near a noisy road. In this way, the ecological situation integrates the experimental paradigm because it is factorized into quiet vs. noisy environment. A first benefit of this approach is to directly test the effects of different ecological situations (here, noisy vs. quiet environment) on participants' electrophysiological responses even if the ecological situation cannot be controlled. A second benefit is to demonstrate that mEEG can be reliably recorded in noisy environments.

Other works also explored cognitive mechanisms and brain dynamics in ecological situations through process-oriented instructions or passive psychological tasks. For example, studies exploring spatial cognition looked at free or guided exploration in different urban environments (Aspinall et al., 2015; Wenczel et al., 2017). Interestingly, a recent study exploring spatial navigation proposed quantifying the saccade-related potentials linked to information processing in the real world as the effect of specific instructions on brain activity (Wunderlich and Gramann, 2021).

Lab experiments have also provided some tasks for real world experiments. For example, an oddball auditory task presenting a series of frequent and infrequent stimuli have been used in different ecological situations (De Vos et al., 2014; Hölle et al., 2021; Ladouce et al., 2017; Scanlon et al., 2020; Zink et al., 2016) to explore attentional mechanisms in real world contexts. One limitation is that these kinds of tasks remain substantially separated from the ecological situation itself: the oddball paradigm calls cognitive mechanisms (attentional and predictive processes) without incorporating any of the given ecological situations (e.g., the fact that the person is currently moving in the real world).

Last, previous studies have explored a targeted behavior directly in concordance with a real-world situation. Lee and collaborators (Lee et al., 1984) have explored how children and adults perceived an approaching car during a road-crossing task. In their study, they demonstrated that an experimental paradigm can be adapted to an ecological situation. Their initial hypothesis was that crossing the road safely requires perceiving the affordance of gaps between vehicles and that this ability could only be learnt by acting in relation to the traffic. From this assumption, an ecological situation was the only possibility to test their working hypothesis properly, but it was obviously too risky to operationalize such an experimental situation. To overcome the risk, a fake street was created adjacent to the real street so that children could safely make decisions about crossing the street. The researchers instructed the children to cross the fake street "*as if they were crossing the first half of the adjacent real street to a traffic island. It was emphasized that they should watch the real traffic on their side of the road and only cross the fake road - without swerving - when they were sure they could safely cross the real road*". Using this ecological paradigm, the authors safely demonstrated children's tendency to accept gaps that are too short before crossing the road. This kind of adaptation of experimental paradigms to ecological experiments is surprisingly rare in behavioral neurosciences. Still, although this ecological paradigm approximated an ecological situation, the ecological variable was not controlled for or factorized in the design: e.g., the speed of the approaching car could be manipulated.

Because of the inherent difficulties of conducting experiments in the real world, most experimental work in the wild used passive tasks or used ecological variables as context effects. As a result, psychophysical tasks are not being exploited and ecological variables are not controlled for and factorized in the design. As such, they do not provide the needed quantitative approach to study brain processes and behavior in relation to the real world. The process of generalization in cognitive neuroscience largely relies on statistical methods applied to quantitative data obtained in an experimental design that factorizes the independent variables. In almost all ecological studies, the absence of controlled and factorized ecological variable prevents exploring the causal relationship between an independent variable and the quantitative measures of behavior and the neural signals, besides the situational aspect. This difficulty in implementing experiments in ecological situations has arisen for one reason: ecological situations imply variable with multi-causal properties, leading an infinite number of working hypotheses about the cause of a behavioral effect. Consequently, as for an experimental paradigm in laboratory, if ecological variable is not controlled many ecological experiments cannot assume the standard scientific criteria of reproducibility, measurability, and refutability.

2.4. Technical difficulties aside, what challenges do we face in the wild?

Let us take an example of the complexity of exploring cognitive processes without control over an ecological situation and contextual stimuli. In a study, exploring the relation between urban context and brain activity, a sample of participants were recorded with mEEG while they followed a path through three distinct zones in the city center of Edinburgh (Aspinall et al., 2015). The authors expected that the three urban zones would modulate participants' brain activity in relation to their emotional state. To do this, they used an algorithm to filter and translate combinations of mEEG signals into four variables indicating the emotional states of the participants as they followed the different paths through the city. The authors reported that when participants walked in a social interaction zone, compared to a busy street, the mEEG activity was more likely to be categorized as an "excited state". However, one severe limitation of the study was the impossibility to fully control events in each zone (context) that was visited by the participants. This prevented homogenizing the effects across participants. For example, a new object (a bike locked to the streetlamp) or an event (a cat crossing the street or the siren of an ambulance passing by) could radically change the experimental context, arousal, and emotional state of the participants. Eventually, each of these natural occurrences in the urban environment provides a plausible confounding factor for the question of interest.

Alternative methodological solutions could be entertained for testing in the wild. A first solution is to develop new data processing to investigate natural cognition in the real world context. In their important study, Wunderlich and Gramann (2021) underlined the difficulties of exploring cognitive processes without control over the contextual situation and stimuli. To address this issue, the authors proposed a blink-related brain potential analysis during real world navigation. Their analysis aimed to link eye movement-related brain potentials during stimuli perception related to navigation in the real world. Under this approach, the working hypothesis is that blink-related brain potentials could specify the involvement of higher cognitive processes in the perception of stimuli. This design offers a tangible possibility to improve real world experiments. However, it can also run the risk of reverse inference so that the putative engagement of cognitive processes is inferred from correlated neuronal activities or activations of particular brain regions (Poldrack, 2006).

Hence, to rigorously address the lab-dilemma issues, we wish to reach an experimental compromise between lab and ecological settings. For this, we introduce a set of eight guidelines and discuss a novel In situ design.

3. Challenges of an approach neither fully wild, nor fully lab

The key aspect of the In situ design is to adapt a lab experiment and its parameters to an ecological setting while minimizing unregulated variables from the real world. Below, we provide a list of guiding principles and present a case study demonstrating the potential of the In situ design to perform psychophysical tasks in a real world setting by implementing an ecological parameter as an integrated variable in the experimental design. Descriptive statistics of two experiments illustrate the feasibility of such approach.

3.1. Guiding principles of an In situ experimentation

The ability to operationalize behaviors and ecological situations in a scientific framework are critical: experimentation in the real world must not only allow for the possibility of fitting the research questions with the ecological situation, it must also integrate elements of the context and situation as controlled parameters in the experiment itself. Conceiving experimentation in that manner partially solves the double standard stated by Brunswik (1943) with an inference process that is possible towards the diversity of real world situations. We describe below eight issues that raised by the lab-dilemma and provide sample guidelines intended to control for multi-causal properties in ecological experimentation.

1) *How to match controlled experimental parameters with the main ecological parameter*? The chosen ecological situation must contain an overall variable effect, operationalized as a controlled parameter in the experimental paradigm (i.e., an independent variable). This way of thinking about an experimental design contributes to the independent variable's measurement quality.

2) *How to limit the effect of other multi-causal variables in the experiment*? To control for other sources of effect, it is necessary to identify them and adapt the experimental protocol around them so that their statistical expectation can be close to zero. The real-world situation must be sufficiently flexible and controllable to reduce the effect variables by applying rigorous control over the situation. Since the experiment could take place in the real world, experimenters must adapt the controls according to their ecological situation. It is important to note that the In situ design excludes the other sources of effect of the real world but not the multi-causality properties of the ecological variable, including the design.

3) *How to reproduce the experiment across testing sessions under the same ecological settings*? This aspect is central and most challenging to implement, given the time and the necessary material. For a given participant, behavioral testing combined with mEEG necessitates between 30 and 60 min of installation time followed by 60–120 min of maximum recording time while the participant performs a task. Therefore, material limitations can be a limiting factor for the inclusion of many participants in a single testing session. Consequently, the real world situation should allow replicating the same ecological situation over time to correspond as closely as possible to the design of the experimentation.

4) *How do we abide by the scientific replicability criterion in ecological settings*? Any research team should be able a priori to replicate a study under In situ design given the implementation of control parameters over the ecological situation. This is particularly difficult, as previously noticed (Aspinall et al., 2015), and factors or events without the possibility to be parametrized may be considered random enough not to affect the observations' statistical reliability.

5) *How to ensure the safety of people and equipment*? Using mobile equipment during testing sessions in the wild must not affect the safety of the participants or other people in the vicinity of the experimental setup.

6) *To stabilize the experiment in time and space*: The necessity to carry out psychophysical experiments outside the lab (possibly with the same apparatus as in the lab) can be straining for the equipment. Consequently, the chosen environment and devices should still allow keeping the experimental setup steady without much manipulation.

7) *To limit the artifact generation for psychophysical tasks*: Since one of the goals of In situ experimentation is to collect behavioral data from complex psychophysical tasks similar to lab conditions, the recording area needs to be sufficiently quiet to minimize the perturbations that may interfere with the behavioral engagement in the task.

8) *To limit the artifact generation for mobile neuroimaging*: One of the main issues using neuroimagery methods such as mEEG remains the electrical artifacts in brain signals. Therefore, it is necessary to partly control the electrical sources around the setup that may affect the equipment. A primary recommendation is to use a power supply (without electrical outlets) for all equipment completed with a check of the recording area for foreign equipment to the setup.

4. In situ experimenting in trains and train stations

As part of a research project studying spatial and temporal cognition in the real world (Wildtimes ANR 2019), we encountered the limitations of experimental protocols when transferring lab experiments to real world experiments. To overcome these limitations, we developed several experiments that incorporated, ad hoc, ecological variables from real-world situations. Hence, ecological variables were factorized and controlled for. Below, we explain the In situ approach generally describing why the train context provided a relevant ecological situation for our questions. We then illustrate the feasibility of the approach for two psychophysical studies, one of which was ran in combination with mEEG.

In high-speed trains like the French Train à Grande Vitesse (TGV), the motion generated by the displacement at high velocity provides an interesting solution to test the idea of In situ design. First, compared to other modes of transportation, the TGV provides a reliable place to do experiments by enabling the replication of identical ecological situation across several testing sessions, within and across individuals. In a study, experiments can be scheduled daily, at precisely the same time of day, and on the same journey. Departure and arrival procedures for a train journey are indeed standardized across days.

The suitability of testing in trains deals with the various logistical advantages of the environment. With the help of SNCF (Société Nationale SNCF; primary operator for the French TGV), one can access TGV during their typical journeys and perform experimental work during the days that are carefully selected (Fig. 1). Participants can be seated at tables, providing a situation comparable to the lab, which is a very decisive advantage for the tests. In our ceas studies, SNCF provided the needed logistical support and the infrastructure, such as rooms in train stations for setting up participants with mEEG without inconveniencing other passengers. Importantly also, conducting experiments in trains does not cause a security risk to the other passengers, unlike experimenting in cars on the road, for instance – and is perhaps less logistically demanding than road closure (Protzak and Gramann, 2018).

The stability of TGV is also a substantial advantage for mEEG recordings and the prevention of possible artifacts. The electrical sources around the setup can be controlled for, which allowed recording mEEG signals without major electrical artifacts. In fact, TGV tends to act a bit like a Faraday cage, protecting against the magnetic field originating from catenaries.

Importantly, in the context of In situ design, the speed of the TGV, its acceleration and its deceleration, can be factorized in an experimental design. Similarly, the orientation of participants in the train (facing forward or backward with the direction of the train) can be factorized and incorporated as experimental variables of interest. In one of our experiments, half of the participants were seated according to the randomization process of the factor level: facing in the same or the opposite direction of the train motion. Whether a participant is seated or moving in the train as well as the salience of the optical flow can be manipulated, along with the distance to a window and whether blinds are pulled or not. The double inverted row seating arrangements in the TGV is beneficial in that it allows exploiting the optical flow and factorizing ecological parameters. The TGV is also notoriously quiet and devoid of auditory cues typically heard in other regular trains on the train tracks.

4.1. Experimental case 1: spontaneous tapping during train journey

Spontaneous motor tapping is a sensorimotor task in which participants pace a movement at their preferred tempo. It requires no special training and can be measured using finger-tapping task in which participants tap their index finger on a keyboard at a comfortable self-paced manner. Studies have shown that self-paced movements - from finger tapping to whole-body movements such as walking – spontaneously fall in the range of about 1–3 Hz or a mean time interval of about 500 ms (Collyer et al., 1994; MacDougall and Moore, 2005; Styns et al., 2007) with some natural inter-individual differences (Hammerschmidt et al., 2021). Spontaneous tapping may reflect the speed of the internal clock and can provide insights on an individual's feeling of time passing. Exploiting the In situ from the lab to the wild, our question was whether spontaneous tempo could capture the anecdotal report that waiting for the train seems to drag as compared to traveling in the train. We used this task during various episodes of a train journey to assess an individual's variation of the speed of the internal clock. The test was performed during five episodes (our independent variable): while waiting for the train at the station, at the beginning, middle and end of the train travel and at the arrival train station. The inter-tap-interval

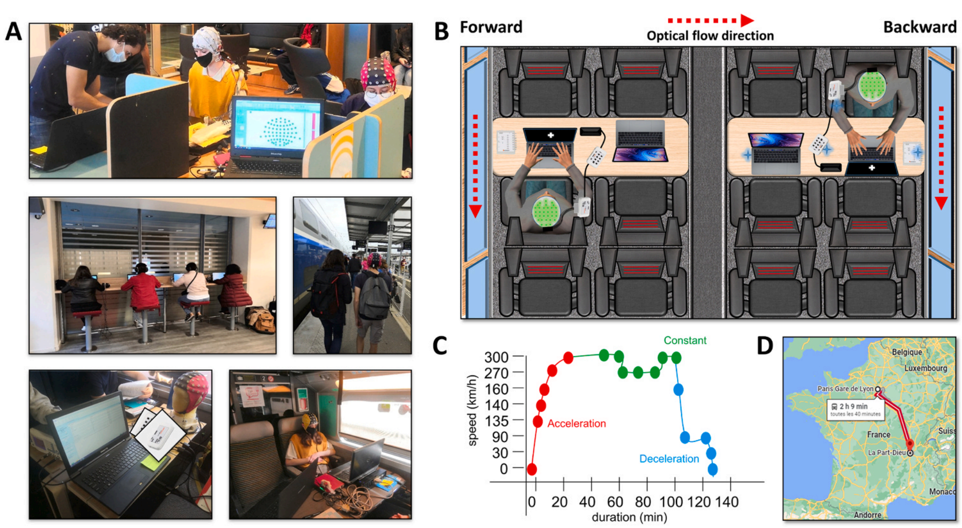

Fig. 1. Example of an In Situ design in train stations and trains. A: The photos show the sequence of an experimental session in the ecological Top: mEEG setup. Middle left: behavioral session at the train station testing four participants at the same time. Middle right: Boarding the TGV with a participant equipped with mEEG. Bottom left and right: Participants setup in the train. **B**: The ecological factor was the seating position of the participant so that the optical flow was forward (left panel) or backward (right panel). When the participant is sitting facing the direction of displacement (forward), the direction of the optic flow varies toward the egocentric reference frame. Conversely, when the participant is seated back to the direction of displacement (backward), the optic flow varies in the opposite direction to the egocentric frame of reference. The direction of the optical flow is given by the dotted red arrow. **C**: Speed profile of the TGV during the selected journey. **D**: Geographical path of the TGV between Paris and Lyon (629 kilometers of railway (Copyrights @ Google Maps)).

(ITI) was our dependent variable (Fig. 2).

4.1.1. Participants

22 participants (12 women, 19–48 y.o) took part in the experiment. All were right-handed, non-smokers, with normal or corrected-to-normal vision and audition, and with no known neurological or psychiatric antecedents. The participants were non-expert travelers on the that journey, they were not regularly practicing music nor singing and they were daily laptop users. The experiment was conducted in accordance to the ethics guidelines and the study was approved by the Comité Ethique de l'Université Paris-Saclay (CER-2018–034-UPSAY). Each participant signed a consent form prior to the study. This study was not preregistered.

4.1.2. Procedure

Participants traveled from Paris to Bordeaux using the TGV. The outbound TGV (Paris-Bordeaux) traveled from 10 a.m. to 12 a.m. and the inbound TGV (Bordeaux-Paris) from 2 p.m. to 4 p.m. The selected journey satisfied all experimental criteria, namely: the journey was not too long, the trip had a reliable and systematic duration, the line was recent which prevented massive occupation, yet the destination is popular, easing the recruitment of participants. The participants were always tested on the same portion of the journey following the exact same time schedule (TGV inOui 8573 and 8508). The systematic scheduled provided a good control for the speed of the train, the vibrations, the auditory cues (absent in the Paris-Bordeaux) and the landscape. In the train, upstairs and downstairs seats, facing forward and backward, seats near windows and corridors were all counter-balanced across participants. They were controlled for and not factorized in this instance; the second study instead factorized this aspect. A light laptop (HP EliteBook 850 G3) was used to collect the data via Psychtoolbox (Matlab). Participants wear headphones (DT 770 PRO Beyedynamic 250 Ohms chosen to help with passive noise reduction). They were asked to produce a self-paced rhythm and to keep it as precisely as possible for one minute. This task was realized five times during the journey: while waiting for the train, during the journey (at departure, during constant speed, and at arrival) and at the arrival station.

4.1.3. Statistical analysis

All statistical analyses were carried out in the R programming language (R Core Team, 2017) and RStudio environment (v.2023.06.0; RStudioTeam 2015) and emmeans (Lenth,2017) software packages. The trial exclusion rule was based on the interquartile range and used within participant to remove outliers: 12703 trials out of a total of 13210 trials were excluded. We used a general linear modelling approach with inter-individual variability treated as a random effect (Knoblauch K., Maloney L.T., 2012) and Tukey method for p adjustment. Pairwise comparisons were running post-hoc. Effect sizes were corrected with a sigma of.27

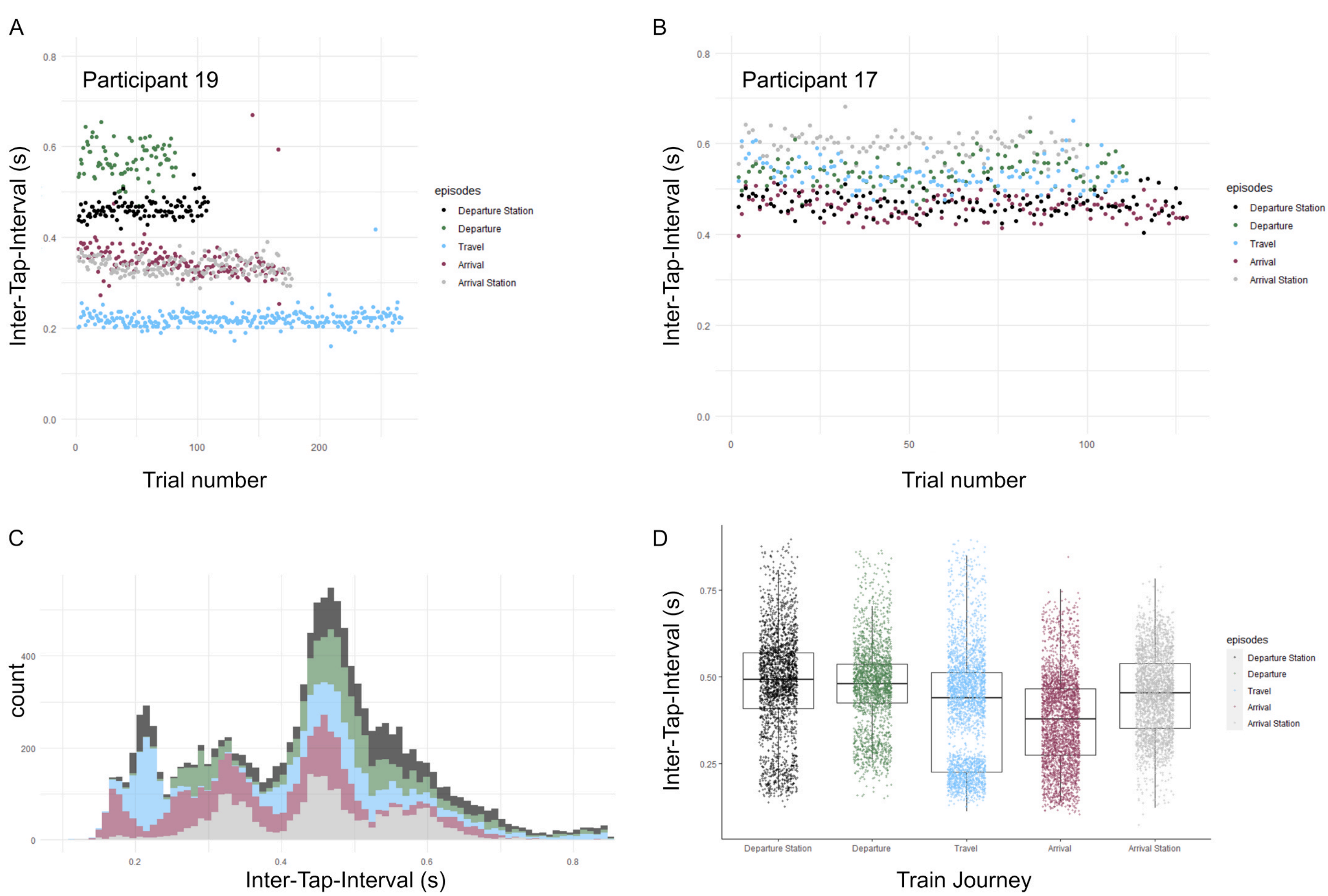

Fig. 2. : Finger-tapping task during episodes of a train journey. A-B: Examples of inter-tap-intervals (ITIs, in seconds (s)) for two participants collected at each episode of a train journey for one minute. Departure station (black) and arrival station (gray) were collected at the train station. Departure (green), travel (blue) and arrival (purple) were collected in the train. Each dot is a sample ITI for the participant. The faster the ITIs, the more sample in the minute. **A:** Participant 19 shows distinct tapping rates during the journey, with the fastest one during train travel. **B:** Participant 17 shows a more homogenous pattern of tapping across the episodes. **C:**Descriptive statistics. Distribution of ITIs per train episode. **D:** Mean ITIs as a function of the episodes. All episodes significantly differed in terms of tapping rate from each other (Table 1). The shorter the ITIs, the faster the tapping. At the train stations before departure and after arrival (black), participants tapped significantly more slowly than during travel episodes. One dot is an individual trial datapoint.

and a confidence level of 0.95. The statistical analysis provided here is not intended to provide a full interpretation of the study.

4.1.4. Observations

In all episodes of the journey, the data collected in this task showed an expected amount of inter-individual variability in tapping rate that is within the range of previous studies. The tapping rate is quantified as inter-tap-intervals (ITIs, in seconds). Figs. 2A and 2B illustrate the behavioral profile for two participants. The distribution of ITIs for the tested population is provided in Fig. 2C. In Fig. 2D, we report the box plots for the study.

Overall, the descriptive statistics (Fig. 2) show a comparable dispersion of ITIs at the train station (at departure in black or at arrival in gray) and during train travels (departure, travel, arrival). As time went by during the train travel, the ITIs appears to shorten (i.e.tapping rates became faster).

4.1.5. Interpretations and limitations

With this experiment, we demonstrate that a finger-tapping task can be easily exported to the real-world and can provide insights on how traveling may impact an individual's internal clock. A major limitation of these observations is the within-individual design, preventing to firmly conclude whether a given train episode, or the chronology of the testing is the most important factor explaining changes in the rate of finger-tapping. We notably seen during the train journey that finger tapping fastens compared to the train stations. These pilot observations illustrate the feasibility of the In situ design and a simple case of how to incorporate an ecological experimental factor in the design.

4.2. Experimental case 2: time-to-contact during train travel

In a second example of an experimental In situ design during train travels, we used an auditory Time-To-Contact (TTC) task with mEEG to explore the impact of optical flow on the behavioral estimation of a sound trajectory and its associated brain activity. Predicting the time course of an approaching object enables anticipatory movements for interceptive or avoidance action. From an evolutionary perspective, this naturalistic behavior is relevant because it determines the direction of potential predators and prey (Cade et al., 2020; Hall and Moore, 2003; Neuhoff, 2001). In the lab, artificially-induced TTC can generate temporal expectations (Chang and Jazayeri, 2018) and engage brain's areas supporting temporal attention and orienting processes (Coull et al., 2008). One objective was to explore the impact of the congruence of optical flow with auditory TTC production: the most salient effect in the ecological situation of the train was thus exploited as an experimental parameter. We used the optical flow generated by the passive linear displacement of the train as an independent variable. We also included an additional factor at two levels - facing forward or facing backward with the direction of train motion. We expected that participants would provide faster responses and produce shorter TTC (pTTC) when facing forward (in the direction of the train travel).

4.2.1. Participants

A total of 71 participants were recruited for the study (35 in the lab experiment and 36 in the TGV experiment; the two samples were distinct). All participants provided a written informed consent. The experiments were approved by the independent ethics committee Comité d'Ethique pour la Recherche de l'Université Paris-Saclay (WildTimes, CER-2018–034 UPSAY). No participants reported known neurological or psychiatric disorders. All were free of medication and had normal hearing or corrected-to-normal vision. 9 participants were a priori excluded from the initial samples (4 in the lab and 5 in the TGV experiment) due to their inability to discriminate the sound target or the trajectory. Hence, 31 participants (15 females, 19–33 y.o) with a laterality quotient (right-handedness) of M = 74% according to the Edinburgh test (Oldfield, 1971) in the lab were included in our analysis. In the TGV, 31 participants (19 females, 24–36 y.o) with a laterality quotient of M = 54% were included in this analysis. This study was not preregistered.

4.2.2. Procedure

The train journey met specific requirements to accommodate the testing of the task: an outbound of at least 2 h allowed enough time for training participants and running the full task, including breaks and unplanned issues (e.g., electrode impedance). A round-trip the same day minimized fluctuations in participants' physiological or psychological states. Each journey covered 629 km in 2 h 10 min with a mean speed of 200 km/h and a peak at 300 km/h (Fig. 1C). The task started precisely 20 min after the departure of the train, allowing precise alignment to the TGV speed schedule. According to the TGV velocity profiles, the nominal speed was reached 15 min after departure and plateaued for one hour (the approximate duration for the TTC task): the speed effectively varied from 270 km/h and 300 km/h during testing.

Participants were seated next to a window, equipped with an audio headset that limited the effect of noise from and with an mEEG setup at the train station before departure (Fig. 1A). Surrounding seats were free and reserved for experimental needs, thereby limiting risks of interference during the testing session. Participants could be facing forward or facing backward with the direction of travel (Fig. 1B).

4.2.3. Stimuli

Auditory stimuli were generated using Matlab R2019a (The MathWork, Massachusetts, USA) as full trajectories of equal distance, duration, and constant velocity then trimmed using Audacity© 2.3.0 software (Team, 2014).

4.2.4. Task

Participants were trained and informed that the goal of the task was to estimate when the auditory stimulus would reach them. They were asked to report their estimation of TTC by pressing the SPACE bar on a computer keyboard as accurately as possible. They heard 50% or 60% of a sound trajectory that should have lasted duration for 1 s, 1.75 s, 2.4 s, 4.2 or 7.35 s before actual contact. The same experiment was tested in the lab.

4.2.5. Statistical analysis

Statistical analyses were performed with R v3.5.0 (Team & others, 2013). Before entering the data into the statistical model, outlier trials defined as + /- 2 * standard error to the mean in pTTC were removed. Constant errors (CEs; difference between the sound arrival time and the participant's estimated arrival time) were computed separately for the five durations: $CE_i = \frac{\sum (Target \quad duration_i - pTTC_i)}{n_i}$. An ANOVA using 3-levels predictor (*Lab* vs. *Backward* vs. *Forward*) and 5-levels predictor (*Durations 1–5*) was performed.

4.2.6. EEG

The EEG data collection PC (HP Elitebook 820 G1 - Intel (R) Core i5–4300 U CPU@ 1.90 GHz) was equipped with Brain Vision Recorder Version 1.24 (Brain Products, GmbH). We used thirty-two electrodes actiCAP Snap (10–20 international system; Brain Products, GmbH) coupled with the 32-channel version of the LiveAmp amplifier (Brain Products, GmbH) for the recording of the EEG signals. We used MNE-Python (Gramfort et al., 2013). Visual inspection identified bad sensors (on average, less than 5% per dataset), which we interpolated. Raw data were bandpass filtered 0.1–40 Hz. Ocular artifact rejection used routine Independent Component Analysis. Raw EEG signals were epoched per condition from − 500–1200 ms.

4.2.7. Observations

The data collection in this task demonstrates the possibility to experiment with a psychophysical task under in situ design and to collect mEEG data with the same reliability of signal as in the traditional lab context. The following behavioural observations (Fig. 3) show the pTTC production for both Lab and TGV conditions and the effect of the factorized ecological variable (Forward vs. Backward) on the pTTC production.

Regarding mEEG signals, we reported all activities evoked by sound onset as a quality check (Fig. 4). The evoked signal amplitude for each condition (The Lab (Fig. 4A), the TGV(Fig. 4B), Forward (Fig. 4C) and backward (Fig. 4D)) was reported for all sensors. The mean evoked activities comparing Lab and TGV ((Fig. 4E) and Forward vs Backward (Fig. 4F) were reported over T7 and T8 sensors. The visual inspection allows us to confirm the close pattern of the evoked signal by sound onset in the *Lab* and the *TGV* and the effect of the ecological variable on evoked activities (Fig. 4F). In addition, the reported patterns of evoked activities also favor the possibility of collecting mEEg data in the real world during a psychophysics task such as the TTC task.

4.2.8. Interpretation and limitations

The current TTC task illustrates the modulation of time-to-contact estimation by the sense of optic flow generated by TGV displacement. The effect of optic flow is particularly relevant in the Forward condition, with a significant increase in CEs compared to Backward and Lab. The observation of this effect is an important proof of concept for in situ design and the possibility of including factorized ecological variables in the experimental design.

Real-world experiments also show several limitations. For instance, in the current In Situ design, the outdoor luminosity could not be controlled due to the substantial variability of the times, days, and seasons during which data were collected. Nevertheless, we limited this possible confound by conducting experiments between May and September, with a stable day duration during this period (in France).

Additionally, experimentation in the wild can be more intense than in the lab: the duration of a complete session was close to ten hours, from the departure to the return to original location (Paris). Participants were not tested continuously and were granted many breaks. Tests were diversified to ensure attention on each, prevent boredom and avoid cognitive fatigue. The well-being of participant was taken care at all times and a great emphasis was placed on listening to the needs of participants.

From the participant's standpoint, wearing an mEEG cap in public can be uneasy (modulo cultural and personal preferences). The level of social acceptability from one participant to another was quite variable and the experimenters had to provide individualized assistance for each. For instance, alternatives for setting up or seating the mEEG participants that could limit stress or anxiety for the participants.

At first sight, the will to introduce controls under an ecological situation feels contradictory with the multi-causality of the real-world situation. However, we argue that multi-causality is still present in the ecological variable. For example, in the current study, the optical flow (i.e. the ecological variable) was composed of many variations inherent to natural optical flow (visual, vestibular, luminosity, haptic, sound, etc.) able to modulate the behavior. Thus, all these sources of potential causality on target behavior (aka multi-causality) shared a common origin yet (the movement of the TGV). In this, the In situ design preserve the multi-causality through its ecological independent variable, despite control above the situation. .

4.3. Conformity of the tasks to the proposed In situ guidelines

Table 2 provides a summary of how both experiments conform to the criteria delineated in 3.1.

5. Discussion

Our goal was to discuss a new In situ design adapted to real world situations in the historical context of the lab dilemma. The critical message of the In situ design is its capacity to fit the experiment and the parameters with an ecological situation, while limiting uncontrolled variables from the real world. The current report shows the feasibility of running psychophysics tasks in the real world and operationalizing an ecological parameter (e.g. the optical flow generated by the TGV) as a controlled variable integrated into the experimental design. Herein, we propose that the In situ design improves scientific criteria for real world experiments regarding the reliability of the results because the choice and the control of the ecological situation limit the multiplicity of variables that can affect the targeted behavior. It also allows a better reproducibility level because the ecological situation through task parameters allows for the operationalization of behavior and brain processes (like a classic parameter in an experimental design). The In situ design partly solves the double standard issue and guarantees a better validity toward the generalization of the results.

The goal of the In situ design is not to tend toward a full generalization that seems unattainable, but rather to assign a limited degree of generalization to a finding based on the specific conditions of its study In

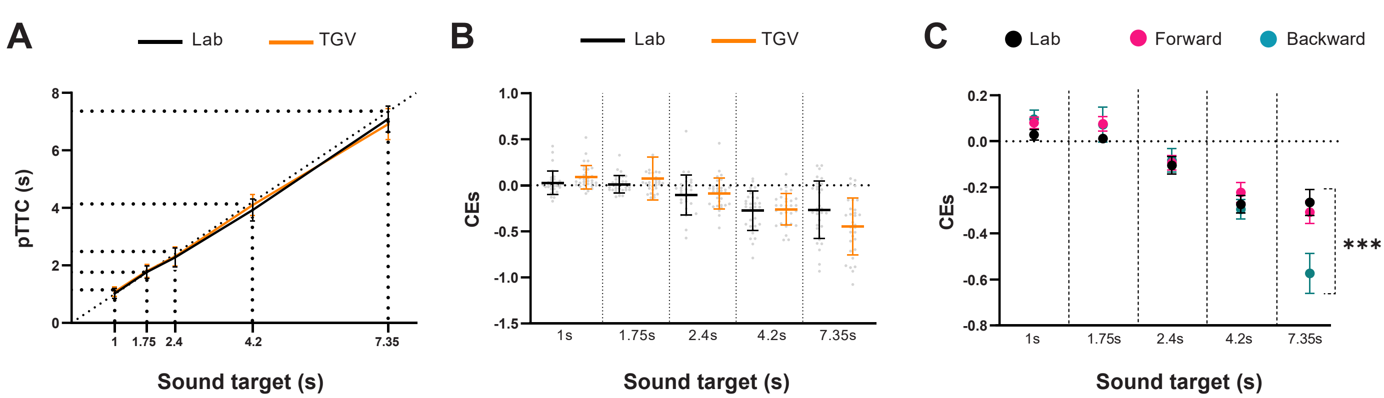

Fig. 3. : Behavioral observations in the TTC task. A: Mean produced time-to-contact (pTTC) as a function of target duration. In the both the lab (black) and in the TGV (orange), pTTC closely follow the identity line indicating that participants produced accurate responses for each of the five target durations. **B:** Mean Constant Errors (CEs) per target durations in the Lab (black) and in the TGV (orange). Grey dots are individual participants. CEs significantly varied with target duration (F (4,30) = 48.03, $p < .001$; $\eta_p^2 = .380$). A CE above 0 indicates that participants overestimated the TTC and were too late. A CE below 0 indicates that participants underestimated the TTC and were too early. Error bars are one standard deviation away from the mean. **C:** Effect of optical flow on pTTC. Mean CEs as a function of target duration and optical flow. No optical flow in the Lab (black). Participants facing forward (fuchsia) or backward (bleu) in the TGV. A significant interaction effect between Optical Flow-Forward × Target-7.35 s increased CEs. * ** correspond to significant contrast with Bonferroni correction (CEs = 0.26, SE = 0.074, Z = 3.56, p_{bonf}<.001) for the target duration 7.25 s. Error bars are one standard deviation away from the mean.

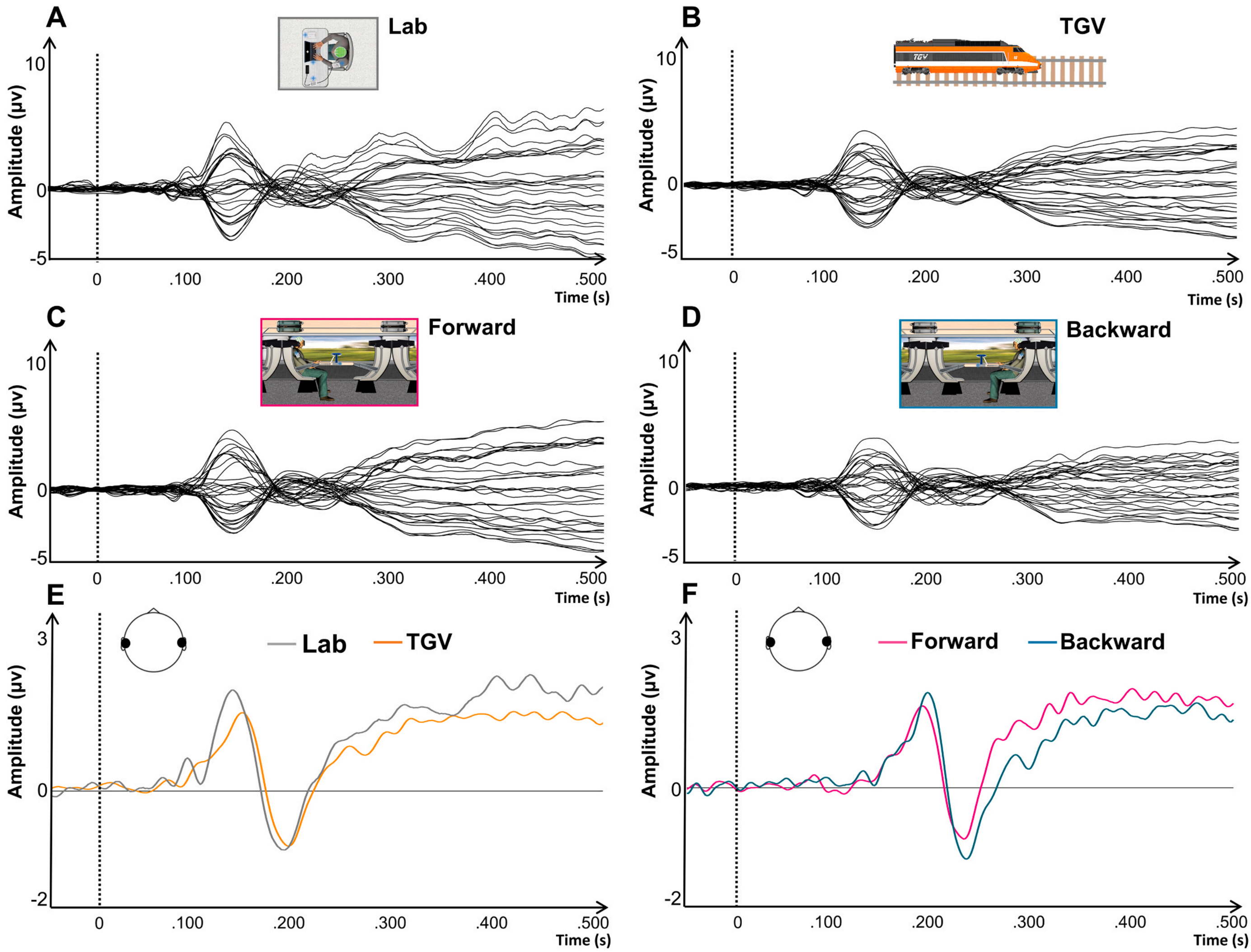

Fig. 4. : mEEG Evoked brain responses (mEEG). **A** & **B**: Activities in the Lab and in the TGV evoked by the sound onset. **C** & **D**: The panels display the mean evoked signal regarding the factorized ecological variable, respectively Forward (**C**) and Backward (**D**) condition. **E**: Evoked comparison activities in the Lab (grey) and TGV (orange) over sensors T7 and T8.

Table 1
Pairwise contrasts of ITIs across all episodes of a train journey. Results are given on the log scale. Confidence level used: 0.95. Confidence level and p-values were adjusted using the Tukey method for comparing a family of five estimates. All episodes significantly differed from each other.

Pairwise Contrasts	Estimate	SE	symp.LCL	asymp.UCL	z.ratio	p.value
Departure - Departure Station	-0.0297	0.00806	-0.0516	-0.00768	-3.681	0.0022
Departure Station - Travel	0.1619	0.00767	0.1409	0.18281	21.092	< .0001
Arrival - Departure Station	-0.2655	0.00757	-0.2862	-0.24486	-35.08	< .0001
Arrival Station - Departure Station	-0.0819	0.00771	-0.1029	-0.06083	-10.615	< .0001
Departure - Travel	0.1322	0.00766	0.1113	0.15311	17.259	< .0001
Arrival - Departure	-0.2358	0.00755	-0.2564	-0.21526	-31.256	< .0001
Arrival Station - Departure	-0.0522	0.00768	-0.0732	-0.03125	-6.797	< .0001
Arrival - Travel	-0.1036	0.00707	-0.1229	-0.08434	-14.653	< .0001
Arrival - Arrival Station	-0.1836	0.00719	-0.2033	-0.16402	-25.53	< .0001
Arrival Station - Travel	0.08	0.0073	0.0601	0.09993	10.954	< .0001

situ. This is how we can generalize the results obtained in an experimental setting, as long as the same conditions and an equivalent sample are replicated. The key difference here is that the effect of the dependent variable in the In situ design goes beyond the design and, because of its ecological nature, extends to certain aspects of the real world. Consequently, we can be confident that the behaviors influenced by these ecological variables will be replicated in any real-world scenarios that share the same characteristics as the In situ situation. For example, we could observe the same behaviors using an optical flow generated during a car journey. Given this, it would be quite straightforward to enhance the generalizability of the findings by considering the brightness of the optical flow for example. In order to achieve this, it would be adequate for the In situ design to incorporate an ecological luminosity variable, which would vary depending on the time of day, and include it in the statistical model. Consequently, we would be able to elucidate the results in terms of luminosity. It is important to note that we are operating

Table 2
Item validation for In situ design in the TGV.

Item	In situ	Finger Tapping Task	Auditory Time-To-Contact task
Ecological parameter	✓	Episodes during the train journey.	Optical flow given by motion comprising one factor with two levels (forward vs. backward).
Multi-causality	✓	Confounding factor of the chronology of episodes during the journey calling for a control experiment fully randomizing the time of testing.	The trains are generally quiet and stable regarding movement generated by the displacement. The occurrence of unaccustomed events is very low.
Reproducibility	✓	The experimentation takes place systematically for the same travel at the same day/hour.	
Replicability	✓	Paris-Bordeaux TGVs are available to all. The design is fully replicable.	Paris-Lyon TGVs are available to all. The design is fully replicable.
Safety	✓	No particular risks associated behavioral testing in high-speed trains. The protocol is also safe for people surrounding the experiment.	No particular risks associated with mEEG recording and behavioral testing in high-speed trains. The protocol is also safe for people surrounding the experiment.
Stability	✓	The layout of seat rows allows the possibility to deploy tables, useful for mEEG and behavioral apparatus (amplifier + laptop).	
Noise (behavior)	✓	The seats surrounding the experimentation were dedicated to the research protocol. Thus, the experimental situation was generally very quiet without human perturbation due to chatting or movements.	
Noise (mEEG)	✓	N/A	We systematically unplug the amplifier and laptop before starting mEEG recording.

within a purely hypothetico-deductive framework at this point, and any variable that is considered for integration into the design must be justified based on its relevance to the behavior being explained. Thus, the incorporation of variables into the In situ design follows the same principle as the experimental design. We view this approach as a middle ground between the complex nature of the real world, with its multi-causes on behavior, and the controlled nature of experimental design, which aims to isolate the effect of a specific variable on a targeted behavior.

The need for cognitive neurosciences to improve experimental designs for characterizing ecological behaviors is becoming critical. A recent review on executive control of stopping action has underlined the need to adapt tasks for the real world (Hannah and Aron, 2021). The authors state the possibility of experimenting outside the lab to explore neural activities related to behavior required in everyday life. The possibilities to improve the design of psychological tasks for real-world experimenting can be summed up as an imperative for naturalistic tasks and the ecological validity of contexts. However, these possibilities can lead to confusion. As pointed out by Holleman et al. (2020), most studies claiming ecological validity refer to the putative proximity between a lab task and the real world, or to whether results from the lab can be generalized to the real world. For example, when a study in the lab uses pictures of naturalistic scenes to explore attention during visual search (Seidl-Rathkopf et al., 2015; Zeni et al., 2020) or naturalistic behavior for planning ability (e.g., Phillips et al., 2006), the added value regarding ecological validity resides solely in the task-goal (aka naturalistic task). Indeed, neither the pictures of naturalistic scenes nor the natural planning ability in the experimental situation can warrant that the inference process can generalize to results in the real world. The second option is to move towards a real world scenario to improve the ecological validity regarding the context and the sensory cues.

In the best-case scenario, an experiment taking place in the real-world should be the combination of a naturalistic task with a sufficient level of internal validity in an ecological situation. It should also involve a factorized ecological variable, and control for other sources of effect. Indeed, the search for more proximity with the real world can be applied to the experimental context, the behavioral task, and the causality drawn from the electrophysiological and the behavioral measures. In all kinds of experimentations (lab, ecological, and VR), the context can be evaluated as a function of its proximity to the real world. For example, experimentation with spatial navigation in the lab does not have the same ecological value regarding the context as experimentation in real world navigation. In the same way, a behavioral task with solid internal validity is not equivalent to a naturalistic task regarding its generalizability. Last, it is logical that the measure following experimentation and task performance strongly related to the ecological validity of the task and the context. For a given targeted behavior, like the estimation of a sound trajectory and its time course exemplified with time to contact, each context can act for or against the ecological validity, the reliability of the measure, and its reproducibility. One possibility can be to combine the context of an experiment with each situation's specific features to understand a targeted behavior.

VR is often proposed as a solution to improve the ecological value of experiments, notably with the possibility to emulate the real world and implement controlled parameters. However, VR is also faced with some paradoxical issues of its own. While the use of VR to explore behavior could mediate the non-ecological approach of lab experiments and the multi-causal real world situation, it requires a deep understanding of the naturalistic behavior to parameterize the virtual environment in a relevant way for the behavior targeted in VR. Only a detailed knowledge about behavior and related brain mechanisms in the real world can ensure, by comparison, the reliability of a VR environment. Without this prerequisite, cognitive neuroscience loses the ecological value of VR. For example, in VR studies exploring the effect of optical flow on brain processes, the parameters related to optical stimulation generally use a moving peripheral grating or virtual environment (e.g., Lo Verde et al., 2019) without the possibility of being sure that the simulation generate the same behavior as the naturalistic stimulation. Furthermore, as argued by Krakauer et al. (2016), the possibility of VR to generate meaningful advances in cognitive neuroscience regarding human or animal behavior and related neural activities will require a strong emphasis on natural behaviors performed by individuals. Therefore, to ensure the ecological validity of its paradigms, VR is constrained to finding markers in the study of natural behavior.

5.1. In situ design as prerequisite to virtual reality?

The In situ design proposed herein precisely provides a means to find these prerequisites for experimentation in VR. As illustrated in the current article, using the In situ design for understanding and measuring the impact of ecological optical flow on specific behavior and then comparing it with a VR simulation is a possibility to ensure the ecological validity of VR studies. Assuming the quality of the virtual environment, we should find in VR the same pattern of behavioral responses during the TTC task as in the In situ design. Once the virtual reality setup is validated from behavior observed in the real world, the VR can assume the ecological validity of the experiment design and allows an infinite number of conditions and parameters to be implemented to study the targeted behavior (i.e., manipulate the relation between sensory cues). However, VR remains a simulation in which neither vestibular inputs nor optical flows can be precisely reproduced. Thus, the back and forth between lab to real world and real world to VR remains essential to implement ecological validity.

5.2. Why is advancing the ecological validity of experiments so crucial to cognitive neurosciences?

The greatest challenge for cognitive neuroscience is to understand complex behavior and associated neural activity in an environment with multi-causal properties. Multi-causality in the real world and the probabilistic causation of events is the primary deterministic feature that has driven evolution and its shaping of psychological processes. Many studies in humans and animals have shown that brain responses are more reliable under natural conditions than they are under artificial stimulation (Hasson et al., 2011). Responses to natural scenes in visual cortex viewing are more reliable than artificial ones, and they also largely differ from those measured under artificial lab conditions (Yao et al., 2007). The capacity of lab experiments to understand ecological behavior and their related brain processes may be currently limited. To improve cognitive neuroscience, variables with characteristics of the targeted ecological situation must be integrated into the experimental design one way or another.

5.3. Conclusions

Experimenting in the wild provides excellent new opportunities to advance traditional questions in cognitive neuroscience. The real world characteristics forces to invent new approaches, outside the traditional lab framework. Our main goal in the current article was to give some practical and theoretical insights on In situ experimental designs, from which experiment in the real world can generalize assuming scientific standards. The variety of real world situations generating ecological variables that can be integrated in an In situ design is near infinite. The take-home message is to search for these ecological situations to improve our understanding of animal and human behaviors in the wild.

Author contributions

W.V and V.vW designed the study. W.V performed the experiments. W.V and V.vW analyzed the data. W.V wrote the first draft of the manuscript with inputs and corrections from V.vW. W.V. V.vW reviewed and corrected the final version of the manuscript.

Ethics approval statement

CER-2018-034_UPSAY from the Comité d'Ethique pour la Recherche de l'Université Paris-Saclay.

Funding statement

The study was funded by the Project-ANR-18-CE22-0016 Wildtimes to V.vW.

Data Availability

The data that support the findings of this study are available on request from the corresponding author WV.

Acknowledgements

We thank Guillaume Lemaitre (SNCF) and Antoine Hone-Blanchet for their careful reading and advice on the manuscript. We thank Simone Morgagni for making possible experimentation in the TGV and Valérie Gyselinck for her inputs throughout the project. Finally, we thank Raphaël Bordas, Yvan Nédélec and, more broadly, all participants who took part in the experimentation campaigns and realized more than 100 train travels between Paris and Lyon or Bordeaux.

Declaration of interest

The authors declare no competing interests.

References

Anwyl-Irvine, A., Dalmaijer, E.S., Hodges, N., Evershed, J.K., 2021. Realistic precision and accuracy of online experiment platforms, web browsers, and devices. Behav. Res. Methods 53 (4), 1407–1425. https://doi.org/10.1037/t15170-000.

Araujo, D., Davids, K., 2009. Ecological approaches to cognition and action in sport and exercise: ask not only what you do, but where you do it. & others Int. J. Sport Psychol. 40 (1), 5.

Araújo, D., Davids, K., 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 (1), 69–78. https://doi.org/10.1080/10407410709336951.

Aspinall, P., Mavros, P., Coyne, R., Roe, J., 2015. The urban brain: analysing outdoor physical activity with mobile EEG. Br. J. Sports Med. 49 (4), 272–276. https://doi.org/10.1136/bjsports-2012-091877.

Bateson, A.D., Baseler, H.A., Paulson, K.S., Ahmed, F., Asghar, A.U.R., 2017. Categorisation of mobile EEG: a researcher's perspective. BioMed. Res. Int. 2017 https://doi.org/10.1155/2017/5496196.

Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., Poeppel, D., Dikker, S., 2019. Brain-to-brain synchrony and learning outcomes vary by student–teacher dynamics: evidence from a real-world classroom electroencephalography study. J. Cogn. Neurosci. 31 (3), 401–411. https://doi.org/10.1162/jocn_a_01274.

Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S.S., Muñoz, L.D., Mullinger, K.J., Tierney, T.M., Bestmann, S., Barnes, G.R., Bowtell, R., Brookes, M.J., 2018. Moving magnetoencephalography towards real-world applications with a wearable system. Nature 555 (7698), 657–661. https://doi.org/10.1038/nature26147.

Breiman, L. (2001). Statistical Modeling: The Two Cultures. 16(3), 199–231.

Brunswik, E., 1943. Organismic achievement and environmental probability. Psychol. Rev. 50 (3), 255.

Brunswik, E. (1952). The conceptual framework of psychology.(Int. Encycl. unified Sci., v. 1, no. 10.).

Brunswik, E., 1955. Representative design and probabilistic theory in a functional psychology. Psychol. Rev. 62 (3), 193–217. https://doi.org/10.1037/h0047470.

Brunswik, E., 1956. Perception and the representative design of psychological experiments. Univ of California Press.

Cade, D.E., Carey, N., Domenici, P., Potvin, J., Goldbogen, J.A., 2020. Predator-informed looming stimulus experiments reveal how large filter feeding whales capture highly maneuverable forage fish. Proc. Natl. Acad. Sci. USA 117 (1), 472–478. https://doi.org/10.1073/pnas.1911099116.

Carlson, T., Goddard, E., Kaplan, D.M., Klein, C., Ritchie, J.B., 2018. Ghosts in machine learning for cognitive neuroscience: moving from data to theory. NeuroImage 180 (February 2017), 88–100. https://doi.org/10.1016/j.neuroimage.2017.08.019.

Chang, C.J., Jazayeri, M., 2018. Integration of speed and time for estimating time to contact. Proc. Natl. Acad. Sci. USA 115 (12), E2879–E2887. https://doi.org/10.1073/pnas.1713316115.

Chaumon, M., Rioux, P.A., Herbst, S.K., Spiousas, I., Kübel, S.L., Gallego Hiroyasu, E.M., Runyun, L., Micillo, L., Thanopoulos, V., Mendoza-Duran, E., Wagelmans, A., Mudumba, R., Tachmatzidou, O., Cellini, N., D'Argembeau, A., Giersch, A., Grondin, S., Gronfier, C., Igarzábal, F.A., van Wassenhove, V., 2022. The Blursday database as a resource to study subjective temporalities during COVID-19. Nat. Hum. Behav. https://doi.org/10.1038/s41562-022-01419-2.

Collyer, C.E., Broadbent, H.A., Church, R.M., 1994. Preferred rates of repetitive tapping and categorical time production. Percept. Psychophys. 55 (4), 443–453. https://doi.org/10.3758/BF03205301.

Coull, J.T., Vidal, F., Goulon, C., Nazarian, B., Craig, C., 2008. Using time-to-contact information to assess potential collision modulates both visual and temporal prediction networks. Front. Hum. Neurosci. 2 (SEP), 1–12. https://doi.org/10.3389/neuro.09.010.2008.

David, S.V., Vinje, W.E., Gallant, J.L., 2004. Natural stimulus statistics alter the receptive field structure of v1 neurons. Journal of Neuroscience 24 (31), 6991–7006.

De Sanctis, P., Solis-Escalante, T., Seeber, M., Wagner, J., Ferris, D.P., Gramann, K., 2021. Time to move: brain dynamics underlying natural action and cognition. Eur. J. Neurosci. 54 (12), 8075–8080. https://doi.org/10.1111/ejn.15562.

De Vos, M., Gandras, K., Debener, S., 2014. Towards a truly mobile auditory brain–computer interface: exploring the P300 to take away. Int. J. Psychophysiol. 91 (1), 46–53.

Debener, S., Minow, F., Emkes, R., Gandras, K., Vos, M.D.E., 2012. How about taking a low-cost, small, and wireless EEG for a walk?, 49, 1449–1453. https://doi.org/10.1111/j.1469-8986.2012.01471.x.

Delaux, A., de Saint Aubert, J.B., Ramanoël, S., Bécu, M., Gehrke, L., Klug, M., Chavarriaga, R., Sahel, J.A., Gramann, K., Arleo, A., 2021. Mobile brain/body imaging of landmark-based navigation with high-density EEG. Eur. J. Neurosci. 54 (12), 8256–8282. https://doi.org/10.1111/ejn.15190.

Dhami, M.K., Hertwig, R., Hoffrage, U., Feigl, H., Go, K., Neurath, O., Schlick, M., 2004. The role of representative design in an ecological approach to cognition, 130 (6), 959–988. https://doi.org/10.1037/0033-2909.130.6.959.

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J.J., Ding, M., Poeppel, D., 2017. Brain-to-brain

synchrony tracks real-world dynamic group interactions in the classroom. Curr. Biol. 27 (9), 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002.
Edwards, D.J., Trujillo, L.T., 2021. An analysis of the external validity of eeg spectral power in an uncontrolled outdoor environment during default and complex neurocognitive states. Brain Sci. 11 (3), 1–26. https://doi.org/10.3390/brainsci11030330.
Eisenberg, I.W., Bissett, P.G., Zeynep Enkavi, A., Li, J., MacKinnon, D.P., Marsch, L.A., Poldrack, R.A., 2019. Uncovering the structure of self-regulation through data-driven ontology discovery. Nature communications 10 (1), 2319.
Fisher, R.A., & others. (1926). 048: The Arrangement of Field Experiments.
Fisher, R.A., 1936. Design of experiments. Br. Med. J. 1 (3923), 554.
Gomez-Marin, A., Paton, J.J., Kampff, A.R., Costa, R.M., Mainen, Z.F., 2014. Big behavioral data: psychology, ethology and the foundations of neuroscience. Nat. Neurosci. 17 (11), 1455–1462. https://doi.org/10.1038/nn.3812.
Gramann, K., Ferris, D.P., Gwin, J., Makeig, S., 2014. Imaging natural cognition in action. Int. J. Psychophysiol. 91 (1), 22–29. https://doi.org/10.1016/j.ijpsycho.2013.09.003.
Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., Hämäläinen, M., 2013. MEG and EEG data analysis with MNE-Python. Front. Neurosci. 7 (7 DEC), 1–13. https://doi.org/10.3389/fnins.2013.00267.
Hall, D.A., Moore, D.R., 2003. Auditory neuroscience: the salience of looming sounds. Curr. Biol. 13 (3), 91–93. https://doi.org/10.1016/S0960-9822(03)00034-4.
Hammerschmidt, D., Frieler, K., Wöllner, C., 2021. Spontaneous motor tempo: investigating psychological, chronobiological, and demographic factors in a large-scale online tapping experiment. Front. Psychol. 12 (June) https://doi.org/10.3389/fpsyg.2021.677201.
Hammond, K.R. (1996). Upon reflection. Thinking \& Reasoning, 2(2–3), 239–248.
Hammond, K.R., Stewart, T.R., 2001. The essential Brunswik: Beginnings, explications, applications. Oxford University Press.
Hannah, R., Aron, A.R., 2021. Towards real- world generalizability of a circuit for action-stopping. Nat. Rev. Neurosci. 22 (September) https://doi.org/10.1038/s41583-021-00485-1.
Hasson, U., Malach, R., & Heeger, D.J. (2011). Reliability of cortical activity during natural stimulation. 14(1), 1–17. https://doi.org/10.1016/j.tics.2009.10.011. Reliability.
Hölle, D., Meekes, J., Bleichner, M.G., 2021. Mobile ear-EEG to study auditory attention in everyday life: Auditory attention in everyday life. Behav. Res. Methods 53 (5), 2025–2036. https://doi.org/10.3758/s13428-021-01538-0.
Holleman, G.A., Hooge, I.T.C., Kemner, C., Hessels, R.S., & Gregory, N.J. (2020). The ' Real-World Approach ' and Its Problems: A Critique of the Term Ecological Validity. 11(April), 1–12. https://doi.org/10.3389/fpsyg.2020.00721.
Ju, U., Wallraven, C., 2023. Relationship between virtual reality balloon analogue risk task and risky decision-making. PLoS ONE 18 (2 February), 1–14. https://doi.org/10.1371/journal.pone.0282097.
Kingstone, A., Smilek, D., Eastwood, J.D., 2008. Cognitive ethology: a new approach for studying human cognition. Br. J. Psychol. 99 (3), 317–340. https://doi.org/10.1348/000712607X251243.
Kingstone, A., Smilek, D., Ristic, J., Friesen, C.K., Eastwood, J.D., 2003. Attention, researchers! It is time to take a look at the real world. Curr. Dir. Psychol. Sci. 12 (5), 176–180. https://doi.org/10.1111/1467-8721.01255.
Kirk, R.E. (1995). Experimental design: Procedures for the behavioral sciences (3rd edn.). Pacific Grove, CA: Brooks/Cole. Communicating Before a Crisis, 589.
Kitchin, R., 2014. Big Data, new epistemologies and paradigm shifts. Big Data Soc. 1 (1), 1–12. https://doi.org/10.1177/2053951714528481.
Koffka, K., 1936. Principles of gestalt psychology. Philosophy 11 (44).
Krakauer, J.W., Ghazanfar, A.A., Gomez-marin, A., Maciver, M.A., Poeppel, D., 2017. Perspective neuroscience needs behavior: correcting a reductionist bias. Neuron 93 (3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041.
Ladouce, S., Donaldson, D.I., Dudchenko, P.A., Ietswaart, M., 2017. Understanding minds in real-world environments. a Mob. Cogn. Approach 10 (January), 1–14. https://doi.org/10.3389/fnhum.2016.00694.
Ladouce, S., Donaldson, D.I., Dudchenko, P.A., Ietswaart, M., 2019. Mobile EEG identifies the re- allocation of attention during real- world activity. Sci. Rep. 1–10. https://doi.org/10.1038/s41598-019-51996-y.
Lange, L., Osinsky, R., 2021. Aiming at ecological validity—midfrontal theta oscillations in a toy gun shooting task. Eur. J. Neurosci. 54 (12), 8214–8224. https://doi.org/10.1111/ejn.14977.
Lee, D.N., Young, D.S., McLaughlin, C.M., 1984. A roadside simulation of road crossing for children. Ergonomics 27 (12), 1271–1281. https://doi.org/10.1080/00140138408963608.
Lejuez, C.W., Richards, J.B., Read, J.P., Kahler, C.W., Ramsey, S.E., Stuart, G.L., Strong, D.R., Brown, R.A., 2002. Evaluation of a behavioral measure of risk taking: the balloon analogue risk task (BART). J. Exp. Psychol.: Appl. 8 (2), 75–84. https://doi.org/10.1037/1076-898X.8.2.75.
Liebherr, M., Corcoran, A.W., Alday, P.M., Coussens, S., Bellan, V., Howlett, C.A., Immink, M.A., Kohler, M., Schlesewsky, M., Bornkessel-Schlesewsky, I., 2021. EEG and behavioral correlates of attentional processing while walking and navigating naturalistic environments. Sci. Rep. 11 (1), 1–13. https://doi.org/10.1038/s41598-021-01772-8.
Lo Verde, L., Alais, D., Burr, D.C., Morrone, M.C., MacDougall, H., Verstraten, F.A.J., 2019. Time dilation effect in an active observer and virtual environment requires apparent motion: No dilation for retinal- or world-motion alone. J. Vis. 19 (3), 1–8. https://doi.org/10.1167/19.3.4.
MacDougall, H.G., Moore, S.T., 2005. Marching to the beat of the same drummer: the spontaneous tempo of human locomotion. J. Appl. Physiol. 99 (3), 1164–1173. https://doi.org/10.1152/japplphysiol.00138.2005.
Makeig, S., Gramann, K., Jung, T.-P., Sejnowski, T.J., Poizner, H., 2009. Linking brain, mind and behavior. Int. J. Psychophysiol. 73 (2), 95–100. https://doi.org/10.1016/j.ijpsycho.2008.11.008.
Massaro, D.W., 2014. Speech perception by ear and eye: a paradigm for psychological inquiry. *Speech Percept. Ear Eye: A Paradig. Psychol. Inq.* https://doi.org/10.4324/9781315808253.
Mavros, P., Austwick, M.Z., Smith, A.H., 2016. Geo-EEG: towards the Use of EEG in the study of urban behaviour. Appl. Spat. Anal. Policy 191–212. https://doi.org/10.1007/s12061-015-9181-z.
Miller, L.C., Shaikh, S.J., Jeong, D.C., Wang, L., Gillig, T.K., Godoy, C.G., Appleby, P.R., Corsbie-Massay, C.L., Marsella, S., Christensen, J.L., Read, S.J., 2019. Causal inference in generalizable environments: systematic representative design. Psychol. Inq. 30 (4), 173–202. https://doi.org/10.1080/1047840X.2019.1693866.
Mustile, M., Kourtis, D., Ladouce, S., Learmonth, G., Edwards, M.G., Donaldson, D.I., Ietswaart, M., 2021. Mobile EEG reveals functionally dissociable dynamic processes supporting real-world ambulatory obstacle avoidance: evidence for early proactive control. Eur. J. Neurosci. 54 (12), 8106–8119. https://doi.org/10.1111/ejn.15120.
Nastase, S., 2021. Toward a more ecological cognitive neuroscience. Brunswik Soc Newsletter 36, 1–23.
Nastase, S.A., Goldstein, A., Hasson, U., 2020. Keep it real: rethinking the primacy of experimental control in cognitive neuroscience. NeuroImage 222 (December 2019), 117254. https://doi.org/10.1016/j.neuroimage.2020.117254.
Neisser, U. (1976). Cognition and reality San Francisco: VV H. Freeman.
Nenna, F., Do, C.T., Protzak, J., Gramann, K., 2021. Alteration of brain dynamics during dual-task overground walking. Eur. J. Neurosci. 54 (12), 8158–8174. https://doi.org/10.1111/ejn.14956.
Neuhoff, J.G., 2001. An adaptive bias in the perception of looming auditory motion. Ecol. Psychol. 13 (2), 87–110. https://doi.org/10.1207/S15326969ECO1302_2.
Phillips, L.H., Kliegel, M., Martin, M., 2006. Age and planning tasks: the influence of ecological validity. Int. J. Aging Hum. Dev. 62 (2), 175–184. https://doi.org/10.2190/EM1W-HAYC-TMLM-WW8X.
Piñeyro Salvidegoitia, M., Jacobsen, N., Bauer, A.K.R., Griffiths, B., Hanslmayr, S., Debener, S., 2019. Out and about: subsequent memory effect captured in a natural outdoor environment with smartphone EEG. Psychophysiology 56 (5), 1–15. https://doi.org/10.1111/psyp.13331.
Pleskac, T.J., Wallsten, T.S., Wang, P., Lejuez, C.W., 2008. Development of an automatic response mode to improve the clinical utility of sequential risk-taking tasks. Exp. Clin. Psychopharmacol. 16 (6), 555–564. https://doi.org/10.1037/a0014245.
Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data, 10 (2), 6–10. https://doi.org/10.1016/j.tics.2005.12.004.
Poulsen, A.T., Kamronn, S., Dmochowski, J., Parra, L.C., Hansen, L.K., 2017. EEG in the classroom: synchronised neural recordings during video presentation. Sci. Rep. 7 (1), 9. https://doi.org/10.1038/srep43916.
Protzak, J., Gramann, K., 2018. Investigating established EEG parameter during real-world driving. Front. Psychol. 9 (NOV), 1–11. https://doi.org/10.3389/fpsyg.2018.02289.
Redcay, E., Moraczewski, D., 2020. NeuroImage Social cognition in context: a naturalistic imaging approach. NeuroImage 216 (July 2019), 116392. https://doi.org/10.1016/j.neuroimage.2019.116392.
Reiser, J.E., Wascher, E., Arnau, S., 2019. Recording mobile EEG in an outdoor environment reveals cognitive-motor interference dependent on movement complexity. Sci. Rep. 9 (1), 1–14. https://doi.org/10.1038/s41598-019-49503-4.
Reiser, J.E., Wascher, E., Rinkenauer, G., Arnau, S., 2021. Cognitive-motor interference in the wild: assessing the effects of movement complexity on task switching using mobile EEG. Eur. J. Neurosci. 54 (12), 8175–8195. https://doi.org/10.1111/ejn.14959.
Rogers, M.L., 2021. Feasibility and acceptability of ecological momentary assessment in a fully online study of community-based adults at high risk for suicide. Psychol. Assess. 33 (12), 1215–1225. https://doi.org/10.1037/pas0001054.
Scanlon, J.E.M., Redman, E.X., Kuziek, J.W.P., Mathewson, K.E., 2020. A ride in the park: Cycling in different outdoor environments modulates the auditory evoked potentials. Int. J. Psychophysiol. 151 (February), 59–69. https://doi.org/10.1016/j.ijpsycho.2020.02.016.
Schmuckler, M.A., 2001. What is ecological validity? A dimensional analysis. Infancy 2 (4), 419–436. https://doi.org/10.1207/S15327078IN0204_02.
Seidl-Rathkopf, K.N., Turk-Browne, N.B., Kastner, S., 2015. Automatic guidance of attention during real-world visual search. Atten., Percept., Psychophys. 77 (6), 1881–1895. https://doi.org/10.3758/s13414-015-0903-8.
Shiffman, S., Stone, A.A., Hufford, M.R., 2008. Ecological momentary assessment. Annu. Rev. Clin. Psychol. 4 (February), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415.
Siddharth, Patel, A.N., Jung, T.P., Sejnowski, T.J., 2019. A wearable multi-modal bio-sensing system towards real-world applications. IEEE Trans. Biomed. Eng. 66 (4), 1137–1147. https://doi.org/10.1109/TBME.2018.2868759.
Smith, E.R. (2000). Research design. Handbook of Research Methods in Social and Personality Psychology, 17–39.
Steiner, M.D., Frey, R., 2021. Representative design in psychological assessment: a case study using the balloon analogue risk task (BART). J. Exp. Psychol.: Gen. 150 (10), 2117–2136. https://doi.org/10.1037/xge0001036.
Styns, F., van Noorden, L., Moelants, D., Leman, M., 2007. Walking on music. Hum. Mov. Sci. 26 (5), 769–785. https://doi.org/10.1016/j.humov.2007.07.007.

Wenczel, F., Hepperle, L., von Stülpnagel, R., 2017. Gaze behavior during incidental and intentional navigation in an outdoor environment. Spat. Cogn. Comput. 17 (1–2), 121–142. https://doi.org/10.1080/13875868.2016.1226838.

Wong, J., & Bronfenbrenner, U. (1977). Bronfenbrenner Toward an experimental ecology of human development 1977 Related papers Toward an Experimental Ecology of Human Development.

Wunderlich, A., Gramann, K., 2021. Eye movement-related brain potentials during assisted navigation in real-world environments. Eur. J. Neurosci. 54 (12), 8336–8354. https://doi.org/10.1111/ejn.15095.

Yao, H., Shi, L., Han, F., Gao, H., Dan, Y., 2007. Rapid learning in cortical coding of visual scenes, 10 (6), 772–778. https://doi.org/10.1038/nn1895.

Yarkoni, T., 2022. The generalizability crisis. Behav. Brain Sci. 45 https://doi.org/10.1017/S0140525X20001685.

Zeni, S., Laudanna, I., Baru, F., Heimler, B., & Melcher, D. (2020). *Increased overt attention to objects in early deaf adults: An eye-tracking study of complex naturalistic scenes. 194*(August 2019). https://doi.org/10.1016/j.cognition.2019.104061.

Zhao, S., Brown, C.A., Holt, L.L., Dick, F., 2022. Robust and efficient online auditory psychophysics. Trends Hear. 26 https://doi.org/10.1177/23312165221118792.

Zink, R., Hunyadi, B., Huffel, S.Van, Vos, M.De, 2016. Mobile EEG on the bike: disentangling attentional and physical contributions to auditory attention tasks. J. Neural Eng. 13 (4) https://doi.org/10.1088/1741-2560/13/4/046017.