Rethinking Human Morality in Real-Life Social Contexts

Eva Vives[1] & Emilie A. Caspar[1]

[1] Moral & Social Brain Lab, Department of Experimental Psychology, Center for Cognitive Neuroscience, Ghent University, Belgium

Corresponding authors

Eva Vives, Moral & Social Brain Lab, Department of Experimental Psychology, Ghent University, Henri Dunantlaan, 2, 9000 Ghent, Belgium. **Email :** eva.vives@ugent.be

Preprint: v1, 28.04.25.

Abstract

Psychological and neuroscientific research has traditionally prioritized internal validity and experimental control, often at the expense of ecological and external validity. This imbalance has produced models of moral cognition that may not reflect how moral decisions unfold in real-world, socially embedded contexts. In this perspective paper, we argue that the lack of attention to ecological and external validity limits the generalizability and relevance of findings—particularly in moral psychology, where decisions are deeply shaped by cultural and interpersonal dynamics. We highlight how variables such as social influence, environments and populations are regularly overlooked when moral cognition is studied in traditional settings, constraining our understanding of the factors influencing (im)moral behavior. Based on recent proof-of-concept studies, we propose actionable solutions to integrate naturalistic environments and diverse populations without sacrificing methodological rigor, calling for a more balanced approach to validity that better captures human moral behavior in its full complexity. We also propose a self-reflection table that researchers from all fields can use to openly reflect on the external and ecological validity of their experimental paradigms.

Keywords: Morality; Social Cognition; Behavior; Ecological validity; External validity

Rethinking Human Morality in Real-Life Social Contexts

Stanley Milgram's experimental approach to obedience is probably the most (in)famous study in the field of moral cognition and helps to illustrate the risks of building sweeping theories based on a single experimental setup. Many can recall its most striking finding: 65% of participants delivered the maximum 450-volt shock to the victim, and this result has been cited many times to explain participation in mass-atrocities (Blass, 2002; Staub, 2014). Yet numerous lesser-known variants of the experiment revealed dramatic shifts in obedience depending on situational factors—such as the location of the study, the proximity to the victim, or the status of the experimenter (Milgram, 1974). A televised replication further highlighted the importance of the authority figure's presence: 81% of participants delivered the final shock when the authority (a TV presenter) was present, while obedience dropped to 25% when they left the stage (the Game of Death, France Television, 2010). But the presence of authority does not seem to be the sole factor influencing disobedience. For instance, among those who refused to obey in the TV show, many cited personal values and experiences—compassion, opposition to Nazism or totalitarian regimes, and the belief that continuing the game went against their nature. This is consistent with a recent study with former genocide perpetrators tested in Rwanda who frequently resisted immoral orders, explaining that their past experiences with obedience made them unwilling to inflict further harm (Seyll et al., 2025). These examples demonstrate that no single experimental context can fully capture the complexity of human behavior, as both situational and dispositional factors also shape decision-making. For instance, the agentic state theory developed by Milgram to explain obedience has been criticized by some authors, as they argue it cannot explain the variability observed across different experimental variants (Haslam et al., 2015, 2017). It is widely accepted that human beings are inherently social, with our development and survival deeply intertwined with our interactions with others. Yet, most studies have sought to understand human moral cognition and behavior in isolated settings, using artificial stimuli and controlled environments that constrain the expression of our social nature within the confines of a laboratory. Furthermore, the vast majority of studies have been conducted on convenience samples from western countries with homogeneous profiles (Arnett, 2016). The main issue is that nearly all current models in moral cognition - and in the whole field of psychology and neuroscience - have only been validated using studies that fit this description.

Two main questions, based on these limitations, help illustrate better the core of the problem. First, do the models and theories we construct truly reflect human functioning, or are they merely artifacts shaped by specific experimental conditions? And second, are our results and interpretations generalizable across different settings, time periods, and/or populations? The first question has been debated since the early years of experimental psychology (Brunswik, 1949; Barker, 1968) and is known as the "real-world or the lab"- dilemma (Hammond & Stewart, 2001). It concerns how well a study mirrors real-life situations and has been closely linked to the notion of ecological validity, which suggests that greater naturalism—of tasks, stimuli, research context, and other factors (Holleman et al., 2020)—brings researchers closer to studying "real" cognition. Some scholars argue that there is an inherent tradeoff between ecological validity and experimental control: the greater the ecological validity of a task, the lower its experimental control (Fan et al., 2021). Critically—and perhaps due to the many challenges of incorporating naturalness into an experimental context—much of the existing literature has prioritized experimental control over ecological approaches. As a consequence, this has significantly limited the extent to which psychological and neuroscientific theories accurately reflect human moral behavior in its natural, interactive, and socially embedded context.

The second question pertains to the concept of external validity, which refers to the generalizability of findings beyond controlled research settings and homogeneous convenience samples (Campbell, 1957). Despite its importance, external validity has historically been less frequently prioritized in psychology and neuroscience compared to internal validity (Tebes, 2000), which has largely contributed to the ongoing "generalizability crisis" (Vazire et al., 2022; Yarkoni, 2022; Simons et al., 2017). One key reason for this is that practical and methodological constraints make external validity difficult to achieve. For instance, real-world environments (as opposed to laboratory settings) introduce uncontrolled variables, and traditional neuroscientific techniques were originally designed to capture brain activity under highly controlled conditions. Additionally, the ease of recruiting university students has led to a dominant reliance on this homogeneous sample, primarily from so-called WEIRD (Western, Educated, Industrialized, Rich, and Democratic) populations (Henrich et al., 2010). This reliance has reinforced a narrow, internally valid approach while neglecting external validity.

Ecological validity and external validity have sometimes been used interchangeably (Berkowitz & Donnerstein, 1982; Holleman et al., 2020), and although they are indeed distinct concepts, both aim to achieve a common goal: enhancing the naturalism of studies and ensuring that the theories that emerge hold across various contexts and populations. Some initiatives have recently emerged to address the issues associated with the lack of external validity. For instance, some journals now require constraints on generalizability statements regarding the population studied to ensure that researchers accurately reflect on the external validity of their paradigms (Simons et al., 2017). Nevertheless, this requirement often fails to consider some key aspects of ecological validity, such as the naturalism of stimuli, research environments, participants' movements, and social interactions—crucial elements that complement an externally valid approach. Currently, there remains a strong favoritism toward internal validity and a predominant focus on homogeneous populations in the literature. However, with recent methodological advancements and increased access to diverse populations, we argue that the trade-off between experimental control and an ecologically and externally valid approach is not as rigid as previously assumed (Fan et al., 2021).

In this perspective paper, we argue that social variables—often overlooked in moral psychology—are essential for developing more nuanced, ecologically valid models of (im)moral behavior. We focus on the field of moral cognition to provide a theoretical framework for our argument, although the broader issues of ecological and external validity should be considered across all fields of psychology and neuroscience. We focus on moral cognition because the lack of ecological and external validity in this domain may hinder our ability to obtain a reliable proxy for a "true" and culturally-sensitive moral response. Moral decisions often occur in dynamic, emotionally charged, and uncertain environments, which are difficult to replicate in controlled experiments. The role of social context has often been overlooked, despite the fact that moral decisions never occur in a social vacuum. The motivations to engage in morally right or wrong behaviors are always related to human social life. There is therefore a critical need to articulate the moral findings obtained in "cold," individual contexts, with the complexity of the "hot" social world, particularly in relation to well-known mechanisms such as conformity, social influence, or self- and other-perceptions (e.g., Van Bavel et al., 2015). Additionally, moral reasoning and values differ across cultures (Graham et al., 2016), yet most studies assume a universal moral framework based on Western participants, limiting generalizability. First, we present the main theories of moral cognition and examine the challenges posed by the lack of ecological and external validity. Then, we propose actionable solutions to address these issues without compromising experimental control. Specifically, we review various proof-of-concept studies that have extended laboratory paradigms to real-life social contexts by incorporating more naturalistic environments and stimuli, including underrepresented populations and large-scale international consortia. These approaches are, of course, relevant beyond the field of moral cognition—not only for researchers employing neuroscientific methods but also for those using other methodologies—and can be adapted to various research fields.

Current Theories In Moral Cognition: Do They Hold Up In The Real World?

The different complexities and issues leading to a lack of external validity in the current theories and models of moral cognition

Since 2001, the field of moral cognition has been largely shaped by the Dual-Process Model (DPM, Greene et al., 2004, 2008; Greene & Haidt, 2002), which posits that moral reasoning results from the interaction of two distinct systems: affective and cognitive processes (see Box 1). Similar to the Social Intuitionist Model (Haidt, 2001), the DPM emphasizes the role of affective processes, which had long been neglected in earlier models of moral cognition (e.g., Kohlberg, 1969; Turiel, 1983) that prioritized moral reasoning as a precursor to moral judgment (see Box 1). The DPM was initially supported by Greene and colleagues' (2001) fMRI study on moral dilemmas, which compared responses to "personal" and "impersonal" dilemmas (see Box 2). Their findings showed that resolving sacrificial dilemmas involved brain regions associated with both deliberative, controlled processes and emotional processing. In impersonal dilemmas—which involve indirect physical harm—participants showed higher rates of utilitarian responses, associated with deliberative reasoning and increased activation in frontal regions such as the medial prefrontal cortex (mPFC) and dorsolateral prefrontal cortex (DLPFC). In contrast, personal dilemmas elicited faster deontological responses linked to intuitive and automatic processes, engaging prefrontal and limbic regions, including the ventromedial prefrontal cortex (vmPFC) and anterior cingulate cortex (ACC).

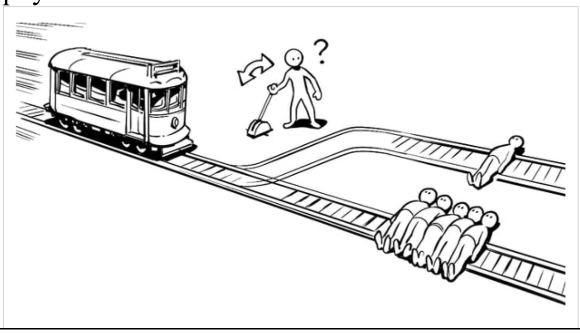
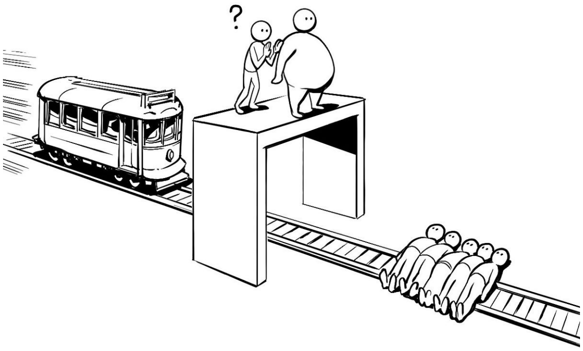
However, the DPM has received multiple criticisms over the years, particularly concerning its accuracy and specificity (Guglielmo, 2015). Notably, further analysis of the data supporting the personal/impersonal distinction revealed that the effects reported in Greene et al. (2001) were driven by idiosyncratic item characteristics, which represented only a small subset of stimuli in the dilemmas set. McGuire et al. (2009) indeed found that the observed interaction between dilemma type and response type was not due to generally longer reaction times for "appropriate" responses in personal dilemmas, as initially proposed. Rather, the effect was driven by unusually fast "inappropriate" responses in a small number of impersonal dilemmas, suggesting that the original findings may have been influenced by item-level artifacts. Additionally, although the DPM is supported by a large body of evidence, most of the data have been collected from WEIRD (Western, Educated, Industrialized, Rich, and Democratic) populations or large-scale industrialized societies. For example, Greene et al. (2001) included only 18 participants, all undergraduates; their subsequent studies included 41 and 82 Princeton undergraduates (Greene et al., 2004) and later studies used Amazon Mechanical Turk workers and Harvard students (Paxton et al., 2012). These narrow samples raise concerns about the generalizability of the findings.

Other models of moral judgment—also known as information-processing models (Guglielmo, 2015)—emphasize how individual moral judgement is shaped by the information available at the time of appraising a moral transgression. This includes the agent's degree of volition, responsibility, and intentionality (Shaver, 1985), the consequences of the transgression that may influence punishment (Cushman, 2008), and the causal relationship between the agent and the victim that informs harm perception (Schein & Gray, 2018). A distinct framework, the CNI model—positioned at the intersection of philosophy and psychology—proposes that deontological and utilitarian choices should be evaluated along three independent dimensions: sensitivity to consequences, sensitivity to moral norms, and a general preference for action versus inaction (Dale & Gawronski, 2023; Gawronski et al., 2017). This model challenges the normative claim that utilitarian judgments would be more reliable and potentially "morally superior."

While these theoretical contributions offer significant insights into third-party moral reasoning and individual's moral intentions, we argue they are less effective in predicting first-person moral decisions. This limitation arises, first, because the supporting data are largely derived from experiments using text-based scenarios—such as moral dilemmas and vignettes—which raises questions about their ecological validity and applicability to real-life moral decision-making. Second, the effects are typically observed in isolated testing environments, either in laboratories or online, with limited attention paid to the influence of peers, social conformity, and social dynamics on moral judgments. Overall, these models are task-dependent and differ from theories explaining moral behavior in other experimental contexts, such as obedience (e.g., the agentic state theory, Milgram, 1974) or the engaged followership theory (Haslam et al., 2015), and conformism (e.g., Social Impact Theory, Latané, 1981) (see Box 1). To date, no single model fully accounts for moral cognition and behavior, though existing ones share common limitations regarding both external and ecological validity. In the following sections, we expand on these issues by connecting research in the moral domain to existing work in psychology and neuroscience.

Box 1. Overview of the dominant models of moral behavior
Dual-process model (Greene et al., 2002; 2004; 2007) This model emphasizes the roles of both affective and cognitive processes during moral judgment. The recruitment of distinct brain areas depends on the nature of the dilemma: People endorse utilitarian responses if the emotional response is low, as observed in impersonal dilemmas (trolley-type) and endorse deontological responses if the emotional response is high, as observed in personal dilemmas (footbridge-type).
Social intuitionist model (Haidt, 2001) This model stresses the role of emotions, assuming that moral judgements are explained by automatic affective reactions (qualified as intuitions), that precede moral reasoning, incorporating social intuitions about events and people (i.e., others evaluations, maintenance of a positive self-image).
Causal-Intentional model (Cushman, 2008) This model, outcome-oriented, stresses the importance of causality and intentionality factors in moral judgement. It posits that mental states such as beliefs and desires, drive harmful intentions of actions, that in turn are sanctioned by wrongness judgements. In contrast, consequences and causes of one's agent's action are apprehended by blame and punishment. Additionally, empirical evidence suggests blame is both a function of mental states (intention) and actual consequences (incidence of harm).
CNI model (Gawronski et al., 2017) This mathematical model quantifies three determinants based on responses to moral dilemmas : sensitivity to consequences (C), sensitivity to norms (N), and general preference for inaction versus action (I). Utilitarian and deontological judgments are measured in an independent manner. This model has been shown to be sensitive to inter-individual differences on scales of empathy, psychopathy and need for cognition (Körner, Deutsch & Gawronski, 2020).
Theory of Dyadic Morality (Schein & Gray, 2018) This model posits that moral judgments are shaped by causal inferences about the mental states of a dyadic structure involving the perpetrator (agent causing damage) and the victim (vulnerable patient). Three core components are necessary to attribute moral responsibility: (1) an intentional agent, (2) a suffering victim, and (3) a clear causal association between the two. The TDM also challenges the idea of "harmless wrong" by proposing that actions labeled as victimless are still intuitively processed as harmful.
Agentic State Theory (Milgram, 1974) This theory proposes that individuals shift into an "agentic state" when they perceive themselves as acting on behalf of an authority figure. In this state, they suspend personal responsibility for their actions, viewing themselves as mere instruments executing orders. Following Milgram, this psychological shift enables ordinary people to carry out harmful acts they would normally resist.
Engaged Followership Theory (Haslam & Reicher, 2017) Rather than obeying blindly, this theory suggests people comply with authority because they identify with the authority's goals and believe in the legitimacy of the cause. Harmful actions arise not from passive obedience, but from active commitment to what is perceived as a meaningful collective mission.

Social Impact Theory (Latané, 1981)
This theory explains conformity as a function of social forces exerted by others. It posits that the impact of a group on an individual's behavior depends on three key factors: the strength of the influencing group (how important or authoritative they are), their immediacy (how physically or psychologically close they are), and their number (how many people are exerting the influence). While conformity generally increases as these factors increase, the theory also posits a diminishing effect—each additional person adds less influence than the one before. This model highlights how social influence is not just about group size, but about relational and contextual dynamics.

Box 2. Moral judgement terminology	
Sacrificial and hypothetical moral dilemmas vs. real-life dilemmas	
Sacrificial and hypothetical dilemmas are scenarios that elicit a moral conflict in participants, because they involve a choice between two mutually exclusive moral values, such as utilitarian or deontological outcomes. These scenarios often describe hypothetical situations that do not reflect real-life implications for participants.	Real-life moral scenarios are instruments that allow to elicit real moral behavior on the environment, by employing paradigms inducing realistic affective reactions and judgements. For instance, by using medical electroshocks that induce real-pain to confederates or participants, it is possible to measure first- (or third party) responses to harm.
Personal vs. Impersonal dilemmas	
Personal dilemmas, also referred to as 'Incidental' dilemmas (Lotto et al., 2014), are scenarios in which the death of one person is a foreseen but unintended consequence of the action aimed at saving more people, like in the Trolley dilemma. They usually involve indirect physical harm. 	Impersonal dilemmas, also referred to as 'Instrumental' dilemmas (Lotto et al., 2014), are scenarios in which the death of one person is a means to save more people, like in the Footbridge dilemma. They involve direct physical harm. 
Utilitarian vs. Deontological Judgements	
Utilitarian (consequentialist) judgements refer to responses that maximize resources for the greater good (e.g., lives, money). It focuses on the outcomes and consequences of the moral action, maximizing benefits and minimizing costs.	Deontological judgements refer to axiomatic rejection of harm, following ethical universal social rules such as the obligation of not killing, 'thou shalt not kill'. It absolutely prohibits causing harm of any kind, whatever the consequences of the moral action.

"No person is an island" - The lack of social interactions

The adage "No person is an island" serves as a crucial reminder that human cognition and behavior cannot be fully understood in isolation. Yet, much of psychology and neuroscience continues to focus on individuals as though their thoughts, decisions, or emotions exist independently of social contexts. Traditional experimental paradigms have largely prioritized controlled environments where participants complete tasks alone, minimizing external influences to ensure internal validity (e.g., Kahneman & Tversky, 1979; Miller, 1969). While this approach has yielded critical insights into various processes, it does not fully capture the complexities of human behavior, which is inherently shaped by social interactions (Baumeister & Leary, 2017; Schilbach et al., 2013), even the most basic (and unconscious) ones with the experimenter (Doyen et al., 2012). Numerous studies have shown that our behaviors, decisions, and cognition are influenced by others—ranging from simple actions, such as adjusting our food choices based on social influence (Higgs, 2015), to more complex decisions, like financial investments shaped by peer recommendations and market trends (Zheng et al., 2021), and even to extreme cases, such as joining violent movements (Mironko 2004).

There are various ways in which we interact with others or are influenced by them, ranging from basic social interactions to joint actions aimed at achieving a common goal, and extending to different forms of social influence, such as peer presence, conformity, compliance, and obedience (Tricoche & Caspar, 2024). Social interactions refer to reciprocal exchanges between two or more individuals that shape their behaviors, emotions, and cognitions. These interactions can be direct (e.g., face-to-face conversations, cooperative tasks) or indirect (e.g., engaging in online discussions). A specific subset of social interactions, joint actions, occur when two or more individuals coordinate their behaviors to achieve a shared goal (Sebanz, Bekkering, & Knoblich, 2006). Examples include playing music in an ensemble, carrying furniture together, or passing a ball in a team sport. Social influence refers to the ways in which individuals' thoughts, feelings, and behaviors are affected by the real or imagined presence of others (Cialdini & Goldstein, 2004). The literature has mostly described three main forms of social influence, that are conformity, compliance and obedience to authority. Conformity occurs when individuals adjust their behaviors, attitudes, or beliefs to align with group norms (Asch, 1951). This can be informational, where people conform because they believe the group has accurate knowledge, or normative, where they conform to fit in and avoid rejection (Toelch & Dolan, 2015). Obedience, in contrast, involves following direct orders from an authority figure and is often driven by perceived legitimacy and power (Milgram, 1963). Unlike conformity, which is influenced by peer dynamics, obedience relies on an explicit command from someone in a position of authority. Compliance, on the other hand, refers to changing one's behavior in response to a direct request from another person or group, even when authority is not involved (Cialdini, 2001). Social presence has also been considered as a form of social influence, even though its influence would be more discreet (Tricoche & Caspar, 2024). It has notably been shown that the mere presence of others can affect performance and behavior, a phenomenon known as the social facilitation effect (Zajonc, 1965). For example, people may perform better on simple tasks but worse on complex tasks when others are watching (Uziel, 2007).

The consideration of others' presence or our interactions with them has been referred to as a second-person approach in the literature (Schilbach et al., 2013). This approach involves studying behaviors, decisions, or brain mechanisms during real-time, interactive social experiences rather than observing individuals in isolation (Redcay & Schilbach, 2019). Unlike traditional first-person approaches (i.e., the experience of mental states in first-person perspective) or third-person approaches (i.e., the passive observation of others), second-person approaches examine reciprocal, dynamic social interactions where two or more individuals actively engage with each other (Hari et al., 2015; Northoff & Heinzel, 2006). These various forms of social interactions have rarely been considered in the development of models of moral cognition, where most participants must make decisions alone in front of a computer screen. Yet, numerous studies on social interactions and influence have demonstrated that our interactions with others can shape behavior (Milgram, 1974; Asch, 1951) and even modifies the brain's response (Caspar et al., 2016; Beyer et al., 2017; Caspar, Rovai et al., 2025; see Tricoche & Caspar, 2024, for a review).

The lack of representativeness of the populations

A persistent challenge in psychology and neuroscience is the overreliance on WEIRD (Western, Educated, Industrialized, Rich, and Democratic) populations (Arnett, 2016), which significantly limits the generalizability of research findings. A recent study by Thalmayer et al. (2021) showed that, despite some progress, 89% of the world's population remains underrepresented in psychological research. Focusing solely on these so-called WEIRD populations, which constitute only 12% of the global population (Henrich et al., 2010), severely restricts our ability to determine whether a given psychological phenomenon is universal or culture specific. This issue is further compounded by the widespread reliance on convenience samples, such as university students, who represent only a narrow segment of human experience and cognition. Consequently, the current models developed in moral psychology (see Box 1)—as well as psychological and neuroscientific theories more broadly—may be "generalizable" only to a small portion of the world's population.

Yet, numerous studies have demonstrated significant cultural differences across various domains of human behavior and cognition, including visual processing, motivation, self- and social perceptions, or arithmetic processing (see reviews and meta-analyses by Ames & Fiske, 2010; Falk et al., 2013; Han & Ma, 2014; Kim & Sasaki, 2014). In the field of moral cognition, it is well-established that moral concepts and their development vary across cultures, shaped by distinct social systems and values (Bentahila et al., 2021), which may lead to a re-assessment of the generalizability of existing theories. Additionally, limiting research to homogeneous samples prevents us from understanding how individuals with diverse life experiences and environmental backgrounds may refine our understanding of human cognition and their models (e.g., Pechtel & Pizzagalli, 2011; Malti & Latzko, 2010). For instance, Milgram's famous obedience studies have largely been used to explain participation in genocides (Blass, 2002; Staub, 2014). However, these studies never included individuals who had actually participated in a genocide, leaving an important gap in understanding whether Milgram's theory (Milgram, 1974)—or subsequent theories (e.g., Haslam et al., 2015) —would hold true for individuals with such life experiences. Notably, recent experiments on resistance to immoral orders showed that former genocide perpetrators recruited in Rwanda displayed a high rate of resistance (Seyll et al., 2025), whereas younger individuals born after the genocide almost never disobeyed (Caspar, Gishoma, et al., 2022), revealing how dispositional factors can influence the results.

Increasing population diversity in research is thus a critical step to develop more accurate models of moral cognition —and, more broadly, most existing models of human cognition.

The problem of unnatural stimuli and environments

The question of the naturalness of stimuli from various modalities presented to participants has been extensively debated in the literature (Hamilton & Huth, 2018; Fan et al., 2021). For visual stimuli, static stimuli represent the most basic level of naturalness and are often considered simplistic, not fully effective, and incapable of conveying realistic information. Fan and colleagues (2021) describe five levels of naturalness, ranging from static images (e.g., a static face without contextual background—level 1) to dynamic images (e.g., video playback—level 2), real-time videos (level 3), real-life social stimuli in laboratory environments (level 4), and real-life social stimuli in real-world environments (level 5). Importantly, these authors note that despite being the least natural, static visual stimuli remain the most commonly used in psychology and neuroscience, even though the naturalism of the stimuli presented influences various brain's responses. The problem of unnatural stimuli has also been noted in the auditory modality and in psycholinguistics, whereas the most common methods for studying how the brain processes speech or understands language rely on isolated words or simple sentences, overlooking the richness of the human language (Hamilton & Huth, 2018). In studies relying on sacrificial moral dilemmas (see Box 2), scenarios were originally presented on paper or displayed on a screen, without providing any natural context (see Christensen & Gomila, 2012, for a review). For this reason, the use of hypothetical and sacrificial dilemmas has been criticized as such dilemmas suffer from poor external validity. In many cases, the scenarios are not adapted to conduct empirical behavioral science, as they can be seen as amusing, absurd and unrealistic by participants, entailing an increased risk of not engaging the same psychological processes as other moral situations (Bauman et al., 2014; Kahane et al., 2018).

The question of the environment can be approached from the perspective of both external validity and ecological validity. From an ecological validity standpoint, the key concern is whether the environment meaningfully recreates real-life contexts and whether participants interact with it in a naturalistic manner. Traditional neuroscience and psychology studies typically take place in laboratory settings, where participants are required to remain seated in front of a computer screen in a lab or lay still in a scanner, limiting real-world movements (Reis et al., 2014; Stangl et al., 2023). However, there are a growing number of studies showing that various brain states differ during movement compared to remaining seated (Ladouce et al., 2017; Aghajan et al., 2017; Bohbot et al., 2017; Richer et al., 2024).

In addition, and in complement to the question of naturalistic movements, prominent research also suggests that moral behaviors may change depending on whether individuals respond in a laboratory setting or a natural environment. From an external validity perspective, the question is whether a study's lab setting allows results to be generalized to real-world situations and whether findings would replicate across different environments. The study of Milgram has for instance famously shown different results on moral behaviors depending on the location where his study was conducted (Milgram, 1974). Classically, conducting research in naturalistic and/or various environments has been argued to pose several challenges, particularly regarding the use of neuroscience equipment. However, with the increasing portability of neuroimaging methods, researchers can now conduct studies in real-world settings outside academic institutions. This shift not only enhances the external validity of studied phenomena but also increases sample diversity (Caspar, 2024), thereby improving the generalizability of findings. These mobile techniques were described as next-generation neuroscientific technologies (Lau-Zhu et al., 2019). Despite their many advantages, studies conducted in real-world environments remain relatively scarce and are still the exception rather than the norm in the scientific literature.

How To Enhance External And Ecological Validity? Experimental Approaches And Proof-Of-Concept Studies

Following the above-mentioned limitations, in this section, we review proof-of-concept studies that demonstrate the feasibility of enhancing ecological and external validity in psychology and neuroscience research, with a focus on moral cognition. We focus on five main elements: (1) testing individuals during various forms of social interactions, (2) increasing the diversity of the population, (3) using more natural stimuli, (4) conducting studies in (various) real-life environments, and (5) allowing participants to make real movements and assume a natural physical position.

No single element is inherently more important than the others in enhancing the external and ecological validity of an experimental task; rather, a combination of these elements contributes to strengthening the validity of the theories derived from empirical research. For instance, some of the proof-of-concept studies we review in the following sections have successfully integrated various methodological and conceptual advancements, with or without the use of neuroscience equipment, and are therefore discussed across different sections. At the end of the paper, we propose a self-reflection table that encourages researchers to examine both the ecological and external validity of their experiments, following these five main categories.

"Every person is a society" - Integrating and controlling for social presence and interactions

Past research on moral judgement has predominantly focused on individuals' preferences measured in isolated settings, overlooking real-life social dynamics. However, people's moral judgements are inherently social, being influenced by norms, politics, religious beliefs (Graham et al., 2009, 2011) and more largely, by others. In a literature review encompassing over 70 years of moral psychology research (n = 419 *studies*), Ellemers et al. (2019) questioned whether social influences were adequately considered. Although some studies incorporated social variables through correlational designs or artificial scenarios, only a small proportion directly examined the impact of norm specificity (7%) and group identity (less than 1%) on moral cognition. This suggests that the socially shared nature of moral guidelines remains largely underexplored in the moral psychology literature. There is therefore a critical need to measure and control for moral appraisals and behaviors occurring in social contexts. For instance, famous psychological effects have been shown to diminish when reduced to the mere presence of an experimenter and their expectations (Doyen et al., 2012). One solution, derived from social psychology, is to control for social presence and audience effects during ongoing experimental tasks.

Seminal work in psychology has demonstrated that being observed by others can influence an individual's task performance (Allport, 1924; Triplett, 1898; Zajonc, 1965). In Triplett's pioneering experiment (1898), cyclists rode faster in the presence of a counterpart than when alone, providing the first evidence of the social facilitation effect. Consistent with this, numerous studies have shown that the mere presence of others can enhance individual performance through social encouragement or social comparison processes (see Guerin, 2010, for a review). Social presence also has significant effects on cognitive functioning, either impairing or facilitating performance on tasks that involve attentional and cognitive control processes (Belletier, Normand, & Huguet, 2019), such as working memory (Belletier et al., 2015; Belletier, Normand, Camos, et al., 2019), cognitive control (Sharma et al., 2010), and early-acquired skills like numerosity and phonological comparison (Tricoche et al., 2023). This is particularly relevant given that the dominant dual-process model of moral judgment emphasizes the roles of cognitive control and attentional resources in shaping moral preferences (Greene et al., 2002, 2004). We therefore argue that the effects of social presence are likely to extend to higher-order processes, such as moral decision-making and moral reasoning.

In addition to its effects on behavior, social presence influences brain functioning. The mere presence of an outgroup member is sometimes sufficient to affect basic neural reinforcement learning signals, particularly in response to feedback signals such as rewards and errors. In outcome evaluation tasks for instance, social presence effects have been reported on feedback monitoring (Hobson & Inzlicht, 2016) when an outgroup member was observing participants' ongoing task, while a diminution in feedback-related negativity (FRN) activity was observed, in comparison to ingroup member presence context. FRN amplitudes were also reported to be augmented on error feedback trials when participants were observed by two strangers (Huang & Yu, 2018) or a passive audience (Tian et al., 2015). It is important to note that several studies have reported no effect of social presence on moral judgement, although related brain activity appears to be affected, as revealed by EEG and fMRI data (Chen et al., 2020; Hobson & Inzlicht, 2016; Huang & Yu, 2018; Tian et al., 2015). These findings therefore support the use of pairing systems to assess both behavioral and neurophysiological responses.

Regarding generalizability, social presence effects - while not on moral behaviors - have also been observed in non-human primates. Studies have shown that the mere presence of a dominant conspecific can alter cognitive control performance in baboons (e.g., Huguet et al., 2014). Research conducted with rhesus monkeys (Demolliens et al., 2017) also showed that passive presence of a conspecific facilitated performance on a visuomotor task. More importantly, this work suggested the existence of 'social neurons', localised in prefrontal regions, that were preferentially active under social presence in contrast to 'asocial neurons' populations, that were discovered to be active under social isolation. Taken together, these studies suggest that the presence of conspecifics induces distinct neural activity in monkeys, highlighting the need to consider social context in human research. Given the close phylogenetic relationship between humans and monkeys, as well as the shared fundamental cognitive mechanisms (e.g., Fagot et al., 2019), incorporating social context is essential for a more comprehensive understanding of human (moral) cognition and behavior. In primatology and ethology, advancements in research methodologies and construction of research centers dedicated to the study of behavior in real-life environments have enabled animals to participate in experiments individually, while remaining socially connected to their group (Cronin et al., 2017; Fagot & Bonté, 2010; Testard et al., 2021). The reasons why such social settings are not largely generalized to the study of human cognition remain unclear, as feasibility is no longer a challenge due to large advances made in the field of social neurosciences to measure and control real-life brain dynamics.

To the best of our knowledge, only a limited number of studies have experimentally manipulated social contexts – albeit indirectly - by using artificial scenarios or paradigms to measure moral cognition. The tipping point here seems to be the direct or indirect nature of the social context, since distinct effects are reported in the literature, depending on the realism of the experimental manipulation. For instance, Bostyn and Roets (2017) investigated the effect of induced conformity pressures on responses to moral dilemmas. They observed that individuals were more likely to conform to deontological judgements when exposed to a majority of deontological opinions, but this was not the case when exposed to utilitarian-consequentialist judgements (see Box 2). This research indicates that participants were prone to conform themselves in a selective manner, favoring deontological to utilitarian opinions, possibly to avoid a morally aversive self-image and possible concerns about mutualistic partner choice preferences. Accordingly, other studies have shown people tend to prefer and trust moral agents who reject harm as close social partners (Everett et al., 2016; Lee et al., 2018). Similarly, a series of studies by Lucas and Livingston (2014) showed that social motivations impacted individuals' moral judgement. The authors manipulated social connectedness by placing participants physically together (Studies 1-2) or asking them to think about close or distant others (Study 3), then asked them to solve high-conflict moral dilemmas. In contrast to Bostyn and Roets (2017), Lucas and Livingston (2014) found increased levels of utilitarian responses when participants were feeling socially connected. A study by Lee et al. (2018) conducted among Korean undergraduates also showed that when observed by the experimenter, participants were more likely to endorse deontological responses, even on personal moral dilemmas. For the authors, this is explained by individuals' tendencies to maintain positive image and good self-reputation. This pattern of results therefore accentuates the need to conduct moral experiments using naturalistic settings, as intentions of actions collected in individual and isolated settings may vary from real-life social interactions with peers.

In brief, social presence can be successfully manipulated in various ways, including the presence of a familiar peer, an ingroup or outgroup member, an experimenter, or even an anthropomorphized robot (Spatola et al., 2018), with evidence of effects on behavior and brain activity.

Studying people in interaction is no longer difficult, as recent techniques in neuroscience, such as hyperscanning (simultaneous brain recordings from multiple participants) and interactive paradigms, allow researchers to explore neural synchrony, social coordination and mutual understanding (Schilbach et al., 2013). Unlike traditional methods that focus on single-subject recordings, hyperscanning allows researchers to assess synchronized neural activity across multiple brains, providing insights into the neural dynamics underlying social interactions (F. Babiloni & Astolfi, 2014). Hyperscanning can be conducted using various techniques, including functional near-infrared spectroscopy (fNIRS), electroencephalography (EEG), magnetoencephalography (MEG), and functional magnetic resonance imaging (fMRI) (for a review, see Czeszumski et al., 2020). This approach facilitates the study of interpersonal dynamics related to social influence, social interactions, social learning, and social presence effects. In this paper, we primarily focus on hyperscanning EEG studies, as EEG offers a real-time, dynamic assessment of brain activity, is non-invasive, and is highly portable, making it well-suited for naturalistic experimental paradigms.

The “social EEG” (i.e., the use of multiple EEG during naturalistic social interaction) has already yielded promising findings in measuring interpersonal synchrony among dyads, such as parent-toddler (Norton et al., 2023) and parent-adolescent pairs (Deng et al., 2024). Additionally, research has demonstrated its applicability to larger audiences. In the educational context, Dikker et al. (2017) investigated school engagement in a classroom setting composed of 12 students (see Figure 1), during 11 classes. Their findings showed that students’ brain-to-brain group synchrony was predictive of classroom engagement and classroom social dynamics. Another study by Nozawa et al. (2019), also reported that preceding physical synchrony positively influenced social bonding between teacher and learner. Interestingly, these findings also extended to non-traditional setups such as museum visits (Dikker et al., 2021) or live music experiences (Chabin, Gabriel, Comte, & Pazart, 2022; Chabin, Gabriel, Comte, Haffen, et al., 2022). Because of its high portability and reduced costs of use, the social EEG renders possible studies testing non-WEIRD populations. In sum, hyperscanning offers new opportunities to precisely record human interactions to investigate group social dynamics (Czeszumski et al. 2020).

No moral principle is universal - Increasing the diversity of the population

One of the most rapidly evolving approaches in cognitive science is the use of large scientific consortia, which facilitate data collection across multiple continents (Moshontz et al., 2018). Additionally, digital platforms and online experiments have enabled large-scale, cross-cultural data collection (e.g., Doell et al., 2024; Olson et al., 2022). For instance, in the field of moral cognition, the trolley dilemma has been tested across 45 countries (Bago et al., 2022), revealing that some effects, such as the personal force effect (see Box 2), have been consistently replicated, while others, such as the interaction between personal force and intention, vary across Western, Southern, and Eastern cultural clusters. Similarly, an online study examining the moral dilemmas faced by autonomous vehicles collected 40 million decisions in ten languages from millions of participants across 233 countries and territories (Awad et al., 2018), further demonstrating the influence of cross-cultural factors on moral preferences. These large-scale studies thus enable the collection of extremely large sample sizes and ensure greater population diversity. More importantly, they allow for the identification of processes that are universally shared, culturally specific, or variable across populations.

It is however important to note several limitations associated with these approaches, including challenges in ensuring data quality, obtaining ethical approvals, dealing with unsupervised samples, and maintaining transparency (see Newman et al., 2021, for a review). Additionally, field acquisition in offline studies requires the involvement of local research teams, which can be difficult to assemble in certain regions due to limited funding to support local researchers. Online studies, on the other hand, often target only populations with access to smartphones or computers, which remain an unaffordable resource for a significant portion of the global population. As a result, large-scale or online studies often neglect certain regions, despite some authors claiming they assess the "universality" of their concepts (Bago et al., 2022). For example, Africa—particularly its sub-Saharan region—remains significantly underrepresented (Awad et al., 2018; Bago et al., 2022; Olson et al. 2022), despite accounting for approximately 18.9% of the world's population.

Many scholars have addressed this issue of generalizability across cultures. Drawing on frameworks such as the collectivist/individualist distinction (Hofstede, 1991) and the independent/interdependent self-construal (Markus & Kitayama, 1991), cross-cultural studies have compared Western (e.g., U.S.) and Eastern (e.g., Chinese) populations, sometimes identifying similarities (e.g., Moore et al., 2011), and often revealing significant differences (see Bentahila et al., 2021, for a review on cultural diversity in moral reasoning). While these frameworks are interesting for studying this issue, they should be used with caution as sociocultural landscapes are constantly evolving. For example, China is increasingly exhibiting a mix of collectivist and individualist traits due to globalization and economic development (see Parker et al., 2009). Other studies have also challenged the assumption of universal moral principles (Graham et al., 2016; Barrett et al., 2016). Findings related to intentionality, moral foundations, or moral argumentation—often considered part of a "universal moral grammar"—may instead reflect evolving norms within Western, industrialized societies (Bentahila et al., 2021). Local and regional contexts must also be considered, as they relate to the beliefs and historical events of the culture. For instance, using participants samples from eight small-scale societies (i.e., hunter-horticulturist, pastoralist, rural agricultural), Barrett et al. (2016) showed that interpretations of intentionality and the moderating effect of mitigating circumstances in moral transgression scenarios varied substantially across societies, affecting the severity of the moral judgement. Such findings emphasize the importance of being cautious when proposing “universal rules” for moral behavior, as cultural, regional, and societal norms often shape moral judgments in ways that diverge from those observed in Western contexts. Several Milgram-like studies were conducted in various countries, notably in India (Gupta, 1983), Jordan (Shanab & Yahya, 1978), and South Africa (Edwards et al., 1969). These studies revealed both variability in obedience across countries (e.g., 42.5% in India, 87.5% in South Africa, and 62.5% in Jordan) and commonalities in responses within this experimental setup (Blass, 2012). Such cross-cultural comparisons are essential for more precisely defining a theory of moral behavior.

Several other studies, whether using neuroscience equipment or not, have also demonstrated the feasibility of reaching rare and remote populations, opening the path to a wide range of new findings. A well-known example is research conducted by Paul Ekman, who investigated the universality of facial expressions and emotions across cultures (Ekman & Friesen, 1971). His fieldwork included the Fore tribe in Papua New Guinea, a remote society with limited exposure to Western culture. In neuroscience research, notable proof-of-concept studies with portable fNIRS and EEG have demonstrated a feasible access to underrepresented populations, despite various field constraints. For instance in the non-moral domain, in a study using portable fNIRS, a team of researchers examined children's reading development in a high-risk illiteracy environment in rural Ivory Coast (Jasinka & Guei, 2018). Another research team used portable EEGs to study survivors of the Khmer Rouge regime and their offspring across various rural villages in Cambodia, investigating the impact of genocide-related trauma on neural functioning (Caspar et al., 2025). Some of their results contrasted with those from studies conducted primarily on Western samples. In the field of moral cognition, notable examples include EEG studies conducted in rural Rwanda with former genocide perpetrators and survivors to examine intergroup biases (Caspar & Pech, 2023; Pech et al., 2024), as well as studies on genocide rescuers and perpetrators to investigate the neural mechanisms of (dis)obedience to immoral orders (Seyll et al., 2025). Other studies have tested inmates in prison (Kox & Caspar, 2025) and military privates on bases (Caspar et al., 2020) to investigate whether the effects observed in classic convenience samples during free and coerced moral decisions (Caspar et al., 2016; 2018) would replicate in populations with different daily life experiences. The authors notably observed that the sense of agency during free moral decision-making is reduced in individuals whose daily lives are marked by restricted autonomy (Caspar, Lo Bue, et al., 2020; Kox & Caspar, 2025). Francis et al. (2018) tested the trolley dilemma in a VR version to evaluate if a previously observed effect among convenience samples (Francis et al., 2016) would replicate among helping professionals (paramedics and fire commanders), two groups with experience in saving lives. Notably, they observed that trained professionals reported less regret about their moral actions compared to control participants, suggesting that specialized training provides post hoc resilience following utilitarian decisions. Decety et al. (2013) also tested incarcerated psychopaths from a North American medium-security center, employing a 1.5 Tesla Siemens Magnetom Avanto mobile scanner, equipped with multiple head coils (see Figure 1). This illustrates remote populations can be reached and eventually, the possible mobility of fMRI research.

Additionally, recruiting experienced populations—such as professionals in care-oriented careers who frequently encounter ethical dilemmas—is particularly valuable for gaining deeper insights into mental states and post hoc rationalizations. Because these populations may develop a "rescue personality" (Wagner et al., 2009), adhere to a professional code of conduct, or simply accumulate years of exposure to (im)moral situations, expanding participant samples beyond WEIRD populations is essential to get a better understanding of the mechanisms underlying moral behavior. Another study compared how different forms of social influence—namely obedience, conformity, and compliance—affect prosocial behaviors across Rwanda, Belgium, and Cambodia (Pech & Caspar, in press). The study found distinct patterns across behavioral, implicit, and neural measures. These studies thus successfully included underrepresented and diverse populations, allowing for a better assessment of whether the results extend to the broader population.

These examples, although specific to the researchers' scientific questions, demonstrate that reaching more diverse populations is largely possible, with or without the use of neuroscience equipment. Importantly, beyond the practical challenges associated with using equipment in complex environments, conducting such research must adhere to rigorous ethical and human considerations. These include collaborating with local researchers, obtaining ethics approval from local institutions, and addressing issues related to cross-border data transfer. These aspects are detailed in Caspar (2024) and are also a key focus of the TRUST Code (TRUST, 2018), an important resource adopted by various institutions and publishers worldwide. The TRUST Code highlights how equitable partnerships between high-income and lower-income countries can benefit both parties.

Beyond the lab: Expanding research into natural and diverse environments

Real-life environments can be reached by various disciplines, but the challenges often depend on the equipment and materials required. For instance, Ekman’s study in Papua New Guinea was conducted in the late 1960s, long before the widespread use of computers in psychological research. He used printed photographs on paper, which were easily transportable. While accessing real-life contexts may be relatively straightforward in disciplines that rely on simple materials like paper and pens (LoSchiavo & Shatz, 2009), neuroscience faces additional challenges due to the complexity of its equipment (Caspar, 2024). In this section, we review proof-of-concept studies that have successfully implemented complex neuroscience tools in real-world settings, demonstrating that such approaches are feasible. By extension, this suggests that research with simpler equipment could also be conducted in similar environments.

Electroencephalography (EEG) is likely the most accessible neuroimaging technique for testing in real-life environments. Due to modern compact, wireless, and quick-to-set-up systems, EEG is easy to transport and use in real-world settings, enabling flexible data collection outside laboratory environments, whether using gel or dry electrodes (see Figure 1). In non-moral domains, several proof-of-concept EEG studies include recording participants' brain activity while they choose a book from a library shelf (Ladouce et al., 2022), walk through an art exhibit in a public museum (Cruz-Garza et al., 2017; Dikker et al., 2021), or are in a classroom setting (Davidesco et al., 2021; Dikker et al., 2017; Poulsen et al., 2017). EEG has also been used to examine neural synchrony in musical ensembles (Babiloni et al., 2011), cognitive responses during sports activities (di Fronso et al., 2019), memory effects in natural environments using smartphone EEG (Piñeyro Salvidegoitia et al., 2019), and in rural Cambodia to assess the impact of the Khmer Rouge genocide on survivors' neural functioning (Caspar et al., 2025).

In the moral domain, such EEG studies are less common, yet several existing research demonstrates their feasibility in examining moral decision-making across diverse environments and populations. For instance, in research on obedience and resistance to immoral orders, proof-of-concept EEG studies have been conducted on military bases to investigate the cognitive mechanisms underlying obedience in soldiers (Caspar et al., 2020) inside prisons to examine the effects of incarceration on inmates' decision-making when following free or instructed choices (Kox & Caspar, 2025), and in rural Rwanda, utilizing local facilities such as churches and bars as testing sites, to investigate the neural mechanisms of resistance to orders among former genocide perpetrators and rescuers (Seyll, et al., 2025). Other studies have been conducted in various locations in Cambodia, Rwanda, and Belgium, transporting the same EEG equipment and computers across those countries to test prosocial behaviors or the intergroup empathy bias across various cultures (Pech & Caspar, in press; Caspar et al., 2024). Conducting EEG research in real-world settings presents of course additional challenges compared to traditional laboratory environments at academic institutions. These challenges include for instance power line noise interference, movement artifacts, battery maintenance, and environmental risks such as dust and humidity. However, many papers provide concrete solutions to overcome these obstacles, demonstrating that such projects remain feasible and methodologically sound (e.g., Caspar, 2024; di Fronso et al., 2019; Davidesco et al., 2021; Ronca et al., 2024; Matthewson et al., 2024).

While other portable neuroimaging methods—such as functional near-infrared spectroscopy (fNIRS) (e.g., Balconi & Fronda, 2020) and functional magnetic resonance imaging (Greene et al., 2001; Young & Saxe, 2009)—have been used to study moral cognition, we are not aware of studies using these techniques in naturalistic environments, highlighting a critical gap in the existing literature on morality. However, research in other non-moral domains demonstrates the feasibility of such projects. Notably, when combined with advanced motion artifact correction techniques, fNIRS has proven itself to be suitable for portable setups (Stangl et al., 2023). Notable examples include a study measuring neural responses to persuasive influence in Jordan (Burns et al., 2019) and a study on child development conducted in rural Ivory Coast where they used waterproof tents as experimental rooms (Jasinska & Guei, 2018). Challenges in using fNIRS in real-world settings include reducing noise, transporting equipment, and managing battery power, particularly in remote areas. However, several resources (Jasinska & Guei, 2018; Pinti et al., 2017) offer practical guidelines for overcoming these obstacles, ranging from artifact detection techniques to the use of solar generators, enabling the successful implementation of fNIRS-based research in naturalistic environments.

An important consideration is that while some neuroimaging methods, such as MRI and Magnetoencephalography (MEG), are less portable and not easily adaptable to naturalistic environments, they can still be integrated with virtual reality (VR) or augmented reality (AR) to immerse participants in more realistic scenarios. Numerous studies have demonstrated that most neuroimaging and electrophysiological techniques can be successfully combined with these setups, including MEG and (Roberts et al., 2019), EEG and VR (Tromp et al., 2018), mobile EEG and AR (Stringfellow et al., 2024), fMRI and VR (Cheetham et al., 2009; Clemente et al., 2014), and fNIRS and VR (Zapała et al., 2022).

Beyond these real-life environments, being able to test a theory or a model in various social contexts and with different social information can also be crucial. A famous example involves Milgram's work, in which he conducted his well-known studies at Yale University or in a run-down office building in Bridgeport. In the latter, obedience rates were drastically lower (Milgram, 1974). Another study showed that placing signs prompting people to put the cigarette butts in an ashtray reduced the number of cigarettes thrown (i.e., conformity), but that placing those signs nearby a location associated with authority (i.e., obedience) compared to a supermarket reduced even more the number of cigarettes thrown on the ground (Pascual et al., 2014). Thus, testing in various contexts, even less extreme than those mentioned previously, have many advantages for better understanding the complex interplay between the environment and (moral) cognition. The question of the environment and its diversity can also be approached following the enclothed cognition theory, which suggests an effect of wearing specific clothing on various psychological processes (Adam & Galinsky, 2012). For instance, Adam and Galinsky (2012) found that wearing a doctor's uniform improved performance on a

Stroop task compared to wearing a painter's uniform. Other studies have found that wearing certain clothes can influence mental abstraction (Burger & Bless, 2017), problem-solving (Van Stockum & DeCaro, 2014), and prosocial behavior (López-Pérez et al., 2016). In the moral domain, in a study where participants could freely decide or were ordered by an experimenter to send real, mildly painful electric shocks to another participant in exchange for a small monetary reward, the authors observed that wearing a classic civilian outfit, a military outfit, or a Red Cross uniform had different effects on behavior (Pech & Caspar, 2023). Findings showed that wearing the Red Cross uniform led to more prosocial behavior, that is, sending less shocks, compared to civilian clothing. The Red Cross uniform also increased neural responses to pain when participants witnessed shocks compared to civilian or military clothing. These results show how different social information and environments can affect behavioral and related brain mechanisms.

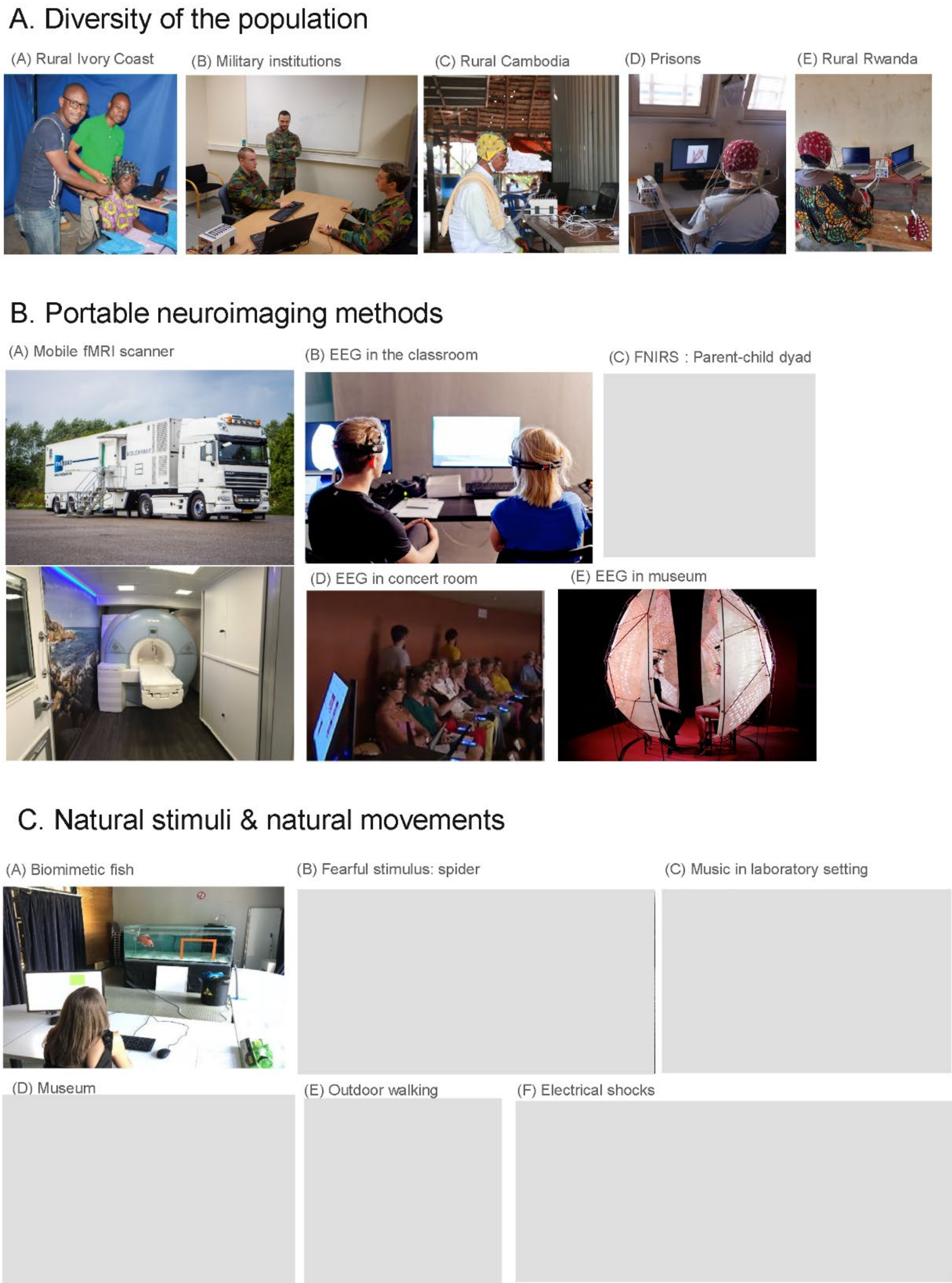


Figure 1. External and ecological validity in proof-of-concept studies. A: Diversity of the population. B: Portable neuroimaging methods. C: Natural stimuli and natural movements. For masked images, the rights acquisition process is ongoing.

Enhancing the naturalness of the stimuli

Many areas of behavioral social science and neuroscience have shown that, even if there is some overlap between hypothetical and real choices in certain instances (e.g., Nickel et al., 2009), there are also notable differences at both the behavioral and neural levels, including distinct activations and varying intensities (see Camerer & Mobbs, 2017, for a review). Beyond decision-making, some studies have also directly compared the realism of presented stimuli, clearly demonstrating the advantages of using more naturalistic materials. Naturalistic stimuli mimic experiences from everyday life and probe complex multimodal integration (Sonkusare et al., 2019). Overall, less natural stimuli tend to lack realism—they are abstract, decontextualized, and static (Fan et al., 2021), such as drawings or cartoons. It is therefore not surprising that several EEG and fMRI studies have shown that viewing real images, as opposed to cartoons depicting painful situations, elicits stronger neural empathic responses (Fan & Han, 2008; Gu & Han, 2007). However, the majority of studies use very abstract stimuli, which reduces ecological validity (Sonkusare et al., 2019). To enhance the naturalness of stimuli, various methods can be applied—individually or in combination—including increasing the realism of the stimuli, embedding them in meaningful contexts, or improving their dynamic properties.

For instance, to study threat processing, rather than relying on pictures (e.g., a picture of a spider or a snake), some authors have used real stimuli (e.g., a real spider; Mobbs, Yu, Rowe et al., 2010; Chouinard & Steward, 2020). In the domain of face processing, some databases have been developed to depict naturalistic emotional faces, incorporating variations in age, ethnicity, and gender (Ebner et al., 2010), in contrast to other databases that rely on static emotional displays without naturalistic emotions (e.g., Langner et al., 2010). This static aspect can be improved by involving dynamic stimuli. For instance, recent studies employed the Dynamic Faces system (Holland et al., 2019), a database of morphed video, to display condemning moral emotions in response to moral violations in third-party moral appraisal tasks (Thébault Guiochon et al., 2025; Vives et al., 2025). Other studies have used videos displaying painful or non painful stimulations to study empathy for pain (e.g., Benuzzi et al., 2018; Iompa et al., 2024), or a social situation in a street to assess the bystander effect (Hortensius & de Gelder, 2014). Using dynamic stimuli is crucial for enhancing naturalism, as several studies have shown that dynamic videos elicit stronger emotional responses than static images. For instance, amygdala and fusiform gyrus activity was heightened in response to dynamic emotional expressions compared to static ones (Freyd, 1987; LaBar et al., 2003). It has also been shown that adding context to emotional expressions also changes the performance of participants. For instance, contextualizing static faces has been shown to enhance emotion recognition compared to non-contextual images (Noh & Isaacowitz, 2013).

A key technology for increasing the authenticity of stimuli is virtual reality (VR) or augmented reality (AR), as these techniques allow participants to be immersed in various contexts and interact with realistic environments and stimuli. In moral judgment research, dilemmas are traditionally presented on a computer screen or paper, often without a specific contextual background (with the exception of historical events dilemmas set by Körner & Detusch, 2023). However, studies using VR have consistently shown that individuals are more likely to make harmful decisions (utilitarian responses) in a virtually realistic context compared to non-VR context (Francis et al., 2016, 2017; Patil et al., 2014), which has refined some of the existing theories. These studies reveal a significant discrepancy between VR-based and text-based moral dilemmas. In Patil et al. (2014), about 76% of participants preferred to save five people in the trolley dilemma, endorsing the utilitarian option, while only 20% did so in the text-based condition. Similarly, Francis et al. (2016) reported that 70% of participants endorsed the utilitarian action in the VR version of the footbridge dilemma. At first glance, this result might seem counterintuitive. According to the authors, this pattern occurs because the visual saliency of VR highlights the negative outcomes associated with inaction. VR has also been used to replicate Milgram's study, with participants knowingly administering shocks to an avatar (Statler et al., 2006). In this study, participants were assigned to one of two conditions: in the visible condition, they could see and hear the avatar throughout the shock administration, whereas in the hidden condition, they were only shown the avatar at the beginning of the study and later received feedback through text. The authors found that obedience was highest in the hidden condition, with all participants administering the maximum shocks (20/20), whereas this number was reduced in the visible condition. Physiological measures, including Skin Conductance Level, Skin Conductance Responses, Heart Rate, and Heart Rate Variability, were generally higher in the visible condition than in the hidden condition. This suggests that directly observing the consequences of one's actions heightened the stress response compared to less natural stimuli.

However, a key limitation of both images, static or not, and VR/AR is that participants remain aware they are engaging with a fictional scenario, which may influence their behaviors. For instance, overall obedience rates in the virtual adaptation of Milgram's study were generally higher than in the original experiment, even in the visible condition (Statler et al., 2006). In the VR study, 85% of participants administered the highest shock, compared to 65% in Milgram's original study (Milgram, 1963). To enhance realism, some studies have employed highly realistic videos or real stimuli (e.g.,Iompa et al., 2024; Crockett, Kruth-Nelson, Siegel et al., 2014), auditory stimuli (Milgram, 1974), or physical objects (e.g., Bègue et al., 2022; Maertens et al., 2007), sometimes in combination with deception. In studies using deception, the realism of the stimuli and the accompanying narrative can lead participants to believe that what they are witnessing is actually happening (Bègue et al. 2022; Iompa et al., 2024). For example, Bègue and colleagues (2022) designed a robotic fish placed in an aquarium to study obedience to orders to harm an animal. However, participants were not informed that the fish was robotic, reinforcing the illusion of real consequences. A general issue with using deception, even when necessary for ethical reasons—such as in studies involving harm to animals—is the challenge of ensuring that participants fully believe the cover story. This is particularly difficult when studies are conducted on convenience samples, as these participants may already be familiar with the frequent use of deception in psychological research. Further, deception has been discussed as increasing stress among volunteers (Herrera, 2001; Christensen, 1988; Boynton et al., 2013).

To avoid the use of deception and enhance the interpretation of findings, other research has taken this a step further by incorporating real stimuli, such as electric shocks, or real-time videos displaying these shocks, thereby eliminating deception and maximizing ecological validity (Caspar et al., 2020; Singer et al., 2004; Crockett et al., 2014; Lockwood et al., 2020; Bostyn et al., 2025). For instance, a novel experimental approach was developed by (Caspar, 2021) to study disobedience to authority in ecological setup, without employing confederates nor cover stories, where two volunteers act by turns, as moral agent or victim. This method relies on the use of a shock simulator (e.g., Digitimer) that delivers mild pain-inducing shocks to the back of the victim's hand when the agent sends electrical signals. The shock intensity is predetermined in collaboration with the participants to ensure it stays within their individual pain threshold. This paradigm has been successfully used in several studies, to examine the effect of obedience on various processes such as the sense of agency (Caspar et al., 2016, 2018), empathy for pain (Caspar, Ioumpa, et al., 2020, 2022) and moral conflicts (Caspar & Pech, 2024). Similarly, real-pain shocks have been used to induce physical harm in real-life moral dilemmas (Bostyn et al., 2018). The results showed that responses to hypothetical dilemmas were not predictive of real-life dilemma behavior, but were predictive of affective and cognitive aspects of the real-life decision. In a recent study, Bostyn et al. (2025), asked participants to privately solve hypothetical dilemmas, before engaging in real-life dilemmas. In these realistic situations, participants had to decide whether to redirect a painful electroshock to a single confederate instead of two confederates, then to provide motivations for their choice. To redirect the shock, participants had to press a button connected to the arm of the confederate, either directly (by pressing the button attached to the arm of the single confederate) or indirectly (by pressing a button placed in the participant room). Interestingly, their results showed that the choices made in the hypothetical dilemmas were predictive of the ones made in the realistic dilemmas, even though real harm was inflicted to confederates. This contrasts with previous results observed in text-based paradigms that reported discrepancies between intentions of actions and choice of actions (Tassy et al., 2013), with lower levels of utilitarian responses observed on personal and impersonal dilemmas. Overall, it appears that hypothetical-dilemma research does not fully reflect real-life dilemmas (Bostyn et al., 2018).

Using realistic paradigms with direct actions in the environment thus allows researchers to bridge the gap between hypothetical judgments and real-life moral behavior, while also measuring the likelihood of actions in personal and impersonal situations (see Box 1).

Integrating real movements and positions

In the field of moral cognition, most paradigms rely on two-alternative forced choice tasks using a keyboard. This dichotomous thinking helps reduce the complexity of the world (Berlin, 1990), but also prevents researchers from studying moral decision-making with nuance (Master et al., 2012). Several techniques, however, offer the possibility to precisely track participants' movements — from systems that capture hand movements using a mouse or touchscreen, to more elaborate setups using multiple infrared cameras and reflective markers on tight-fitting suits to capture full-body 3D motion (Harvey et al., 2024).

Very few studies have however investigated how tracking natural movements can reveal more precise patterns in moral decision-making (Gautheron et al., 2023, 2024). In 2023, Gautheron and colleagues explored how different response modes—binary (two-alternative forced choice) versus continuous (mouse-tracking)—affect the dynamics of moral decision-making. Participants were presented with moral dilemmas and responded using either a binary choice or a continuous rating scale. The findings revealed that the response mode significantly influenced early decision processes, suggesting that the method of response can constrain moral decision-making dynamics. To the best of our knowledge, more sophisticated full-body tracking systems have not yet been used in moral or social decision-making, leaving a significant opportunity for future research.

Eye tracking and pupillometry, which enable the measurement of eye gaze and pupil dilation, are also promising techniques for studying moral behavior. For example, previous studies have shown that eye gaze can predict participants' moral decisions (Pärnamets et al., 2015), provide insights into cognitive load and emotional involvement during VR moral dilemmas (Skulmowski et al., 2014), and that pupil size serves as a reliable physiological indicator of emotional arousal in response to purity violations (Kaspar et al., 2015).

Natural movements in combination with neuroimaging equipment can be complex due to the artifacts introduced by such movements. However, fNIRS and EEG are suitable for studying naturalistic movements and locomotion (Richer et al., 2024), and solutions for removing motion artifacts have been developed (Nathan & Contreras-Vidal, 2016). Notably, many studies have demonstrated that reliable EEG data can be collected even while participants are walking or engaging in sports activities (e.g., Jacobsen, Blum et al., 2020; Debener et al., 2012; di Fronso et al., 2019). Additionally, research has shown that fNIRS can be used while walking (McKendrick et al., 2016; Piper et al., 2014) or during a table tennis task (Balardin et al., 2017), broadening its applicability across diverse environments (see Pinti et al., 2017, for a review) and natural movements.

Discussion And Future Directions

Analogous to the French philosopher Henri Bergson (1900), who posits laughter is inherently social and rarely occurs in complete isolation, we argue that moral behavior is fundamentally social in nature. Consequently, empirical research on morality must more systematically consider social variables as critical predictors of individuals' moral intentions, judgments, and actions. Drawing on over two decades of research in moral psychology and the neuroscience of morality, we advanced three key perspectives, intrinsically connected to external and ecological validity concerns. First, the study and experimental manipulation of psychosocial variables—such as social identity, group dynamics, and the normative influence of others—can be meaningfully addressed through ecological, real-life paradigms. Because others play a central role in our moral decisions—conveying norms, acting as victims or perpetrators of moral transgressions—such decisions cannot be separated from the presence of others, whether perceived or real. Second, it appears that many studies in the field of moral cognition have tackled the problem of population diversity in order to better define cross-cultural behavioral differences, challenging the idea of a universal morality. Increasing this diversity by targeting non-WEIRD participants or reaching remote societies is now feasible through international academic collaborations and mobile equipments that allow conducting studies in person, online through the use of crowdsourced samples (i.e., Prolific, Amazon's MTurk), or large-scale intercultural samples. Third, we encourage authors to use less artificial paradigms and deception in their experiments, as discrepancies between hypothetical and realistic situations have been reported. Specifically, the use of naturalistic stimuli may more accurately capture individuals' "true" moral responses by approximating the real-world processes underlying moral behavior, thereby contributing to more robust and ecologically valid models of moral cognition.

In this paper, we proposed actionable solutions and presented proof-of-concept studies that illustrate how experimental designs can achieve greater ecological and external validity for investigating moral cognition (Ladouce et al., 2022; Gert et al., 2022). While some progress has been made in recognizing the importance of external validity—such as the inclusion of a constraint of generalizability statement in some manuscripts—this addresses only one aspect of the many ways such validity can be improved. As presented, external and ecological validity can be enhanced through multiple approaches, including population diversity, natural stimuli and environments, real-world movements, and social interactions. These elements can be combined to maximize the validity of theories emerging from such studies. To enhance reflection on these aspects of experimental protocols, we propose Table 1 as a tool for researchers to evaluate their own projects. This table presents examples from key studies presented in the present paper, illustrating how each of them integrates one or various aspects of ecological and external validity. With this table, and the technological solutions proposed above, we encourage researchers to strive for the highest possible external and ecological validity, while maintaining their desired level of experimental control.

How to make a study more ecological and externally valid?						
Study	**Topic**	**Diversity of the population**	**Various and real-life environments**	**Natural stimuli**	**Social interactions**	**Real movements and positions**
Self-reflection questions →		***Do my results extend to other populations?***	***Do my results extend to other locations? Was the environment natural?***	***Do my stimuli most naturally represent the expected behavior?***	***Do I evaluate the effect of presence of/interactions with others?***	***Did my participants made natural movements?***
Kox & Caspar, 2025	Moral decision-making in free and coerced conditions	YES Comparing inmates from different detention regimes and matched individuals	YES Inside various prisons with different security levels (intra muros).	YES Moral decisions involving real money taken from another real participant.	YES Condition: Alone versus in presence of an authority figure	NO Sitting in front of a computer with a keyboard.
Pech, Gishoma & Caspar, 2023	Intergroup bias in the aftermath of a genocide and across generations	YES Testing former genocide perpetrators, survivors, and their offsprings	YES Across various rural regions in Rwanda	Partly Imagined prosocial decisions; Static cartoon images with a context	NO People making only their decision in isolation	NO Sitting in front of a computer with a keyboard.
Gautheron, Quinton, Muller & Smeding (2023)	Moral decision-making dynamics	NO Only convenience (university) sample	NO Conducting tests only in the same lab environment	NO Judging moral statements	NO People making only their decision in isolation	Partly Mouse-tracking system compared to 2AFC, but no full body
Beauvois, Courbet & Oberlé (2012)	Obedience to authority during a TV show	Partly People from the same country, but with diversified profiles	YES On a TV platform and compared to the classic studies of Milgram at the university	YES Natural auditory feedback + cover story for (fake) electric shocks	YES Precise control of the experimenter's effect and of the audience' effect	YES They were moving, looking around, and using various levers
Bostyn et al. (2025)	Group effect on moral decisions	NO Only convenience (university) sample	NO Conducting tests only in the same lab environment	YES Moral decisions with real electric shocks delivered to another person.	YES Groups of 3 to 11	NO Sitting in front of a computer with a keyboard.
Hortensias & Gelder	Bystander effect and attentional processes	NO Only convenience (university) sample	NO Conducting tests only in the same MRI environment	YES Dynamic stimuli represrnting a bystander situation	YES Several experimental conditions with different number of bystanders	NO Sitting in front of a computer with a keyboard.
Tricoche, Lovai, lo Bue, & Caspar (2025)	Disobedience to immoral orders	YES Comparison between civilians and military	NO Conducting tests only in a lab environment	YES Moral decisions involving real electric shocks delivered to another real participant.	Partly Receive orders from an authority figure and witness the consequences on a « victim »	NO Laying down in a MRI scanner without moving
Bago, Kovacs et al. (2022)	Moral dilemma	YES Various participants from 45 countries	Partly People in various countries, but within their own immediate environment	NO Written moral judgements	NO People making only their decision in isolation	NO Sitting in front of a computer with a keyboard.
Lee, Sum & Kim (2018)	Moral dilemna	NO Only convenience (university) sample - but in Korea	NO Conducting tests only in a lab environment	NO Written moral judgements	YES Condition: Alone versus presence of an experimenter	NO Sitting in front of a computer with a keyboard.

Table 1. Examples of studies that have integrated one or more aspects to enhance external and ecological validity. We propose several reflection questions that should enhance reflexivity around the ecological and external validity of research designs.

Of course, we acknowledge the challenges associated with some of the proof-of-concept studies presented. For instance, conducting research with portable EEGs in rural Rwanda to study former genocide perpetrators, survivors, and rescuers (Pech et al., 2024; Seyll et al., 2025) is inherently more complex than conducting the same study in a controlled laboratory setting with a convenience sample. Such projects require understanding new cultures, gaining the trust of sensitive populations, and adhering to ethical principles specific to each country, among other considerations. Yet, these studies significantly enhance our understanding of key questions in the field. They also highlight how many current experimental protocols and methods are poorly suited to diverse global populations. For example, the use of reversed items in questionnaires assumes specific cognitive and linguistic skills, while some highly complex tasks may not even be conducted by individuals without an expertise in computer use. This raises the question of whether the observed effects truly reflect underlying cognitive processes or are merely artifacts of task or questionnaire design. Although awareness of these limitations has grown, some researchers still overlook the generalizability of their methods, failing to recognize that the instruments they use are tailored to only a small portion of the world's population (Corneille & Gawronski, 2024.

Ecological and external validity must also be accompanied by careful ethical reflection—particularly in fields such as moral cognition, trauma, or addiction research. For example, while highly realistic scenarios can enhance ecological validity, they may also increase psychological risks for participants. Studying the neural effects of trauma in victims of a traumatic event, for instance, would make it ethically unacceptable to place them in a VR or real environment that resembles the events they experienced, due to the risk of trauma reactivation. Similarly, in moral cognition research, placing participants in ultra-realistic scenarios—even in VR—where they must engage in severely harmful actions can induce high levels of stress. In studies investigating obedience to immoral orders with sensitive populations such as inmates (Kox & Caspar, 2025) or former genocide perpetrators (Seyll et al., 2025), researchers did not employ the classic paradigm involving real electric shocks administered to a victim in a convenience sample (Caspar et al., 2016). Instead, they used a modified paradigm involving financial losses imposed on a victim, thereby reducing the risk of psychological stress. For example, the replication of Milgram's studies for television with a live audience has been condemned by the American Psychological Association (APA) due to the high level of stress caused to participants. Recently, a famous YouTuber asked volunteers to face various moral dilemmas—most notably, a very realistic trolley dilemma—in which each participant's decision involved a potential loss of millions of dollars and caused significant psychological stress. A key point is that while ecological realism is crucial, it must be carefully balanced with ethical considerations.

In conclusion, having demonstrated the feasibility of various research approaches and proposed concrete solutions to enhance ecological and external validity, we urge the field of moral cognition to critically reflect on these aspects of experimental protocols. This includes the more systematic integration of social influence, social contexts, and real-life social dynamics, in order to enrich our understanding of the psychological and neurocognitive processes underlying moral behavior and to better inform current models of morality. We call for a transparent discussion regarding the rationale for including or excluding these elements, as well as the extent to which studies achieve strong ecological and external validity alongside internal validity.

References

Adam, H., & Galinsky, A. D. (2012). Enclothed cognition. *Journal of Experimental Social Psychology*, *48*(4), 918–925. https://doi.org/10.1016/j.jesp.2012.02.008

Aghajan, Z. M., Schuette, P., Fields, T. A., Tran, M. E., Siddiqui, S. M., Hasulak, N. R., ... & Suthana, N. (2017). Theta oscillations in the human medial temporal lobe during real-world ambulatory movement. *Current Biology*, *27*(24), 3743-3751. https://doi.org/10.1016/j.cub.2017.10.062

Allport, F. (1924). *Social psychology*. New York: Houghton Mifflin.

Ames, D. L., & Fiske, S. T. (2010). Cultural neuroscience. *Asian Journal of Social Psychology*, *13*(2), 72–82. https://doi.org/10.1111/j.1467-839X.2010.01301.x

Arnett, J. J. (2016). The neglected 95%: Why American psychology needs to become less American. *American Psychological Association*. https://doi.org/10.1037/14805-008

Asch, S.E. (1951). Effects of group pressure upon the modification and distortion of judgment. In: Guetzkow, H., editor. *Groups, Leadership and Men*. Pittsburgh, PA: Carnegie Press.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The Moral Machine experiment. *Nature*, 563(7729), 59–64. https://doi.org/10.1038/s41586-018-0637-6

Babiloni, C., Vecchio, F., Infarinato, F., Buffo, P., Marzano, N., Spada, D., Rossi, S., Bruni, I., Rossini, P. M., & Perani, D. (2011). Simultaneous recording of electroencephalographic data in musicians playing in ensemble. *Cortex*, 47(9), 1082–1090. https://doi.org/10.1016/j.cortex.2011.05.006

Babiloni, F., & Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: Past, present and future. *Neuroscience and Biobehavioral Reviews*, *44*, 76–93. https://doi.org/10.1016/j.neubiorev.2012.07.006

Bago, B., Kovacs, M., Protzko, J., Nagy, T., Kekecs, Z., Palfi, B., Adamkovic, M., Adamus, S., Albalooshi, S., Albayrak-Aydemir, N., Alfian, I. N., Alper, S., Alvarez-Solas, S., Alves, S. G., Amaya, S., Andresen, P. K., Anjum, G., Ansari, D., Arriaga, P., … Aczel, B. (2022). Situational factors shape moral judgements in the trolley dilemma in Eastern, Southern and Western countries in a culturally diverse sample. *Nature Human Behaviour*, *6*(6), 880–895. https://doi.org/10.1038/s41562-022-01319-5

Barker, R. G. (1968). *Ecological Psychology: Concepts and Methods for Studying the Environment of Human Behavior*. Stanford, CA: Stanford University Press.

Balardin, J. B., Zimeo Morais, G. A., Furucho, R. A., Trambaiolli, L., Vanzella, P., Biazoli Jr, C., & Sato, J. R. (2017). Imaging brain function with functional near-infrared spectroscopy in unconstrained environments. *Frontiers in Human Neuroscience*, *11*, 258. https://doi.org/10.3389/fnhum.2017.00258

Balconi, M., & Fronda, G. (2020). Morality and management: an oxymoron? fNIRS and neuromanagement perspective explain us why things are not like this. *Cognitive, Affective, & Behavioral Neuroscience*, *20*(6), 1336–1348. https://doi.org/10.3758/s13415-020-00841-1

Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M. T., & Fitzpatrick, S. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences*, *113*(17), 12–14. https://doi.org/10.1073/pnas.1522070113

Bauman, C. W., Mcgraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, *8*(9), 536–554. https://doi.org/10.1111/spc3.12131

Baumeister, R. F., & Leary, M. R. (2017). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Bègue, L., & Vezirian, K. (2022). Sacrificing Animals in the Name of Scientific Authority: The Relationship Between Pro-Scientific Mindset and the Lethal Use of Animals in Biomedical Experimentation. *Personality & social psychology bulletin*, *48*(10), 1483–1498. https://doi.org/10.1177/01461672211039413

Belletier, C., Davranche, K., Tellier, I. S., Dumas, F., Vidal, F., Hasbroucq, T., & Huguet, P. (2015). Choking under monitoring pressure: being watched by the experimenter reduces executive attention. *Psychonomic Bulletin and Review, 22*(5), 1410–1416. https://doi.org/10.3758/s13423-015-0804-9

Belletier, C., Normand, A., & Huguet, P. (2019). Social-Facilitation-and-Impairment Effects: From Motivation to Cognition and the Social Brain. *Current Directions in Psychological Science, 28*(3), 260–265. https://doi.org/10.1177/0963721419829699

Belletier, C., Normand, A., Camos, V., Barrouillet, P., & Huguet, P. (2019). Choking under experimenter' s presence: Impact on proactive control and practical consequences for psychological science. *Cognition, 189*, 60–64. https://doi.org/10.1016/j.cognition.2019.03.018

Bentahila, L., Fontaine, R., & Pennequin, V. (2021). Universality and Cultural Diversity in Moral Reasoning and Judgment. *Frontiers in Psychology*, *12*, 764360. https://doi.org/10.3389/fpsyg.2021.764360

Benuzzi, F., Lui, F., Ardizzi, M., Ambrosecchia, M., Ballotta, D., Righi, S., Pagnoni, G., Gallese, V., & Porro, C. A. (2018). Pain Mirrors: Neural Correlates of Observing Self or Others' Facial Expressions of Pain. *Frontiers in psychology, 9*, 1825. https://doi.org/10.3389/fpsyg.2018.01825

Bergson, H. (1900). *Le rire. Essai sur la signification du Comique* [Laughter: An Essay on the Meaning of the Comic]. Bibliothèque de philosophie contemporaine. Revue Philosophique de la France et de l'étranger 50:665-670.

Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist, 37*(3), 245–257. https://doi.org/10.1037/0003-066X.37.3.245

Berlin, S. B. (1990). Dichotomous and complex thinking. *Social Service Review, 64*(1), 46-59.

Beyer, F., Sidarus, N., Bonicalzi, S., & Haggard, P. (2017). Beyond self-serving bias: diffusion of responsibility reduces sense of agency and outcome monitoring. *Social cognitive and affective neuroscience, 12(*1), 138–145. https://doi.org/10.1093/scan/nsw160

Blass, T. (2002). Perpetrator behavior as destructive obedience: An evaluation of Stanley Milgram's perspective, the most influential social-psychological approach to the Holocaust. In L. S. Newman & R. Erber (Eds.), *Understanding genocide: The social psychology of the Holocaust* (pp. 91–109). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195133622.003.0005

Blass, T. (2012). A cross-cultural comparison of studies of obedience using the Milgram paradigm: A review. *Social and Personality Psychology Compass*, *6*(2), 196–205. https://doi.org/10.1111/j.1751-9004.2011.00417.x

Bless, H., & Burger, A. M. (2017). Mood and the regulation of mental abstraction. Current Directions in Psychological Science, 26(2), 159–164. https://doi.org/10.1177/0963721417690456

Bohbot, V. D., Copara, M. S., Gotman, J., & Ekstrom, A. D. (2017). Low-frequency theta oscillations in the human hippocampus during real-world and virtual navigation. *Nature Communications*, *8*, 14415. https://doi.org/10.1038/ncomms14415

Bostyn, D. H., Gouwy, M. C., De Craene, E., Scheirlinckx, J., Tissot, T., Van Severen, R., ... & Van, S. *Beyond Hypotheticals: Moral Choices and Motivations in a Real-life Sacrificial Dilemma.* PsyArXiv. https://doi.org/10.31234/osf.io/sgznt

Bostyn, D. H., & Roets, A. (2017). An Asymmetric Moral Conformity Effect: Subjects Conform to Deontological But Not Consequentialist Majorities. *Social Psychological and Personality Science, 8*(3), 323–330. https://doi.org/10.1177/1948550616671999

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of Mice, Men, and Trolleys: Hypothetical Judgment Versus Real-Life Behavior in Trolley-Style Moral Dilemmas. *Psychological Science*, *29*(7), 1084–1093. https://doi.org/10.1177/0956797617752640

Boynton, M. H., Portnoy, D. B., & Johnson, B. T. (2013). Exploring the ethics and psychological impact of deception in psychological research. *IRB, 35*(2), 7–13.

Brunswik, E. (1949). Discussion: remarks on functionalism in perception. *Journal of Personality, 18*(1).

Burns, S. M., Barnes, L. N., McCulloh, I. A., Dagher, M. M., Falk, E. B., Storey, J. D., & Lieberman, M. D. (2019). Making social neuroscience less WEIRD: Using fNIRS to measure neural signatures of persuasive influence in a Middle East participant sample. *Journal of Personality and Social Psychology, 116*, e1--e11. https://doi.org/10.1037/pspa0000144

Camerer, C., & Mobbs, D. (2017). Differences in Behavior and Brain Activity during Hypothetical and Real Choices. *Trends in Cognitive Sciences, 21*(1), 46–56. https://doi.org/10.1016/j.tics.2016.11.001

Campbell, D. T. (2017). Factors relevant to the validity of experiments in social settings. *Sociological methods*, 243-263.

Carron, R., Blanc, N., & Brigaud, E. (2022). Contextualizing sacrificial dilemmas within Covid-19 for the study of moral judgment. *PLoS ONE, 17*, 1–19. https://doi.org/10.1371/journal.pone.0273521

Caspar, E. A. (2021). A novel experimental approach to study disobedience to authority. *Scientific Reports, 11*(1), 22927. https://doi.org/10.1038/s41598-021-02334-8

Caspar E. A. (2024). Guidelines for Inclusive and Diverse Human Neuroscience Research Practices. *The Journal of Neuroscience, 44*(48), e1971242024. https://doi.org/10.1523/JNEUROSCI.1971-24.2024

Caspar, E. A., Christensen, J. F., Cleeremans, A., & Haggard, P. (2016). Coercion Changes the Sense of Agency in the Human Brain. *Current Biology, 26*(5), 585–592. https://doi.org/10.1016/j.cub.2015.12.067

Caspar, E. A., Cleeremans, A., & Haggard, P. (2018). Only giving orders? An experimental study of the sense of agency when giving or receiving commands. *PLOS ONE, 13*(9), e0204027. https://doi.org/10.1371/journal.pone.0204027

Caspar, E. A., Gishoma, D., & Magalhaes de Saldanha da Gama, P. A. (2022). On the cognitive mechanisms supporting prosocial disobedience in a post-genocidal context. *Scientific Reports, 12*(1), 1–15. https://doi.org/10.1038/s41598-022-26460-z

Caspar, E. A., Ioumpa, K., Arnaldo, I., Angelis, L. Di, Gazzola, V., & Keysers, C. (2022). Commanding or Being a Simple Intermediary: How Does It Affect Moral Behavior and Related Brain Mechanisms? *ENeuro, 9*(5). https://doi.org/10.1523/ENEURO.0508-21.2022

Caspar, E. A., Ioumpa, K., Keysers, C., & Gazzola, V. (2020). Obeying orders reduces vicarious brain activation towards victims' pain. *NeuroImage, 222,* 117251. https://doi.org/10.1016/j.neuroimage.2020.117251

Caspar, E. A., Lo Bue, S., da Gama, P. A., Haggard, P., & Cleeremans, A. (2020). The effect of military training on the sense of agency and outcome processing. *Nature Communications, 11*(1), 4366. https://doi.org/10.1038/s41467-020-18152-x

Caspar, E. A., Nicolay, E., Banderembaho, F., & Pech, G. P. (2024). Volition as a modulator of the intergroup empathy bias. *Social Neuroscience*, 19(5-6), 326-339.

Caspar, E. A., Pech, G. P., & Ros, P. (2025). Long-term affective and non-affective brain alterations across three generations following the genocide in Cambodia. *Biological Psychology, 197,* 109028. https://doi.org/https://doi.org/10.1016/j.biopsycho.2025.109028

Caspar, E. A., & Pech, G. P. (2024). Obedience to authority reduces cognitive conflict before an action. *Social Neuroscience, 19*(2), 94–105. https://doi.org/10.1080/17470919.2024.2376049

Caspar, E. A., Rovai, A., Lo Bue, S., & Cleeremans, A. (2025). Neural correlates of the sense of agency in free and coerced moral decision-making among civilians and military personnel. *Cerebral Cortex, 35*(3), bhaf049. https://doi.org/10.1093/cercor/bhaf049

Campbell, D. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin, 54*(4): 297–312.

Chabin, T., Gabriel, D., Comte, A., & Pazart, L. (2022). Audience Interbrain Synchrony During Live Music Is Shaped by Both the Number of People Sharing Pleasure and the Strength of This Pleasure. *Frontiers in Human Neuroscience, 16,* 855778. https://doi.org/10.3389/fnhum.2022.855778

Chabin, T., Gabriel, D., Comte, A., Haffen, E., Moulin, T., & Pazart, L. (2022). Interbrain emotional connection during music performances is driven by physical proximity and individual traits. *Annals of the New York Academy of Sciences*, *1508(1)*, 178–195. https://doi.org/10.1111/nyas.14711

Cheetham, M., Pedroni, A., Antley, A., Slater, M., & Jäncke, L. (2009). Virtual Milgram: empathic concern or personal distress? Evidence from functional MRI and dispositional measures. *Frontiers in Human Neuroscience, 3*, 802. https://www.frontiersin.org/articles/10.3389/neuro.09.029.2009

Chen, C., Martínez, R. M., & Cheng, Y. (2020). The key to group fitness: The presence of another synchronizes moral attitudes and neural responses during moral decision-making. *NeuroImage, 213*(155). https://doi.org/10.1016/j.neuroimage.2020.116732

Chouinard, P. A., & Stewart, R. (2020). Having a live huntsman spider on a rubber hand does not modulate the rubber-hand illusion in a top-down manner. *Experimental Brain Research, 238*(9), 2053–2066. https://doi.org/10.1007/s00221-020-05869-w

Christensen, J. F., & Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: A principled review. *Neuroscience and Biobehavioral Reviews, 36*(4), 1249–1264. https://doi.org/10.1016/j.neubiorev.2012.02.008

Christensen, L. (1988). Deception in Psychological Research: When is its Use Justified? *Personality and Social Psychology Bulletin, 14*(4), 664–675. https://doi.org/10.1177/0146167288144002

Cialdini, R. B. (2001). The science of persuasion. *Scientific American, 284*(2), 76-81. https://doi.org/10.1038/scientificamerican0201-76

Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, *55*(1), 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Clemente, M., Rey, B., Rodríguez-Pujadas, A., Barros-Loscertales, A., Baños, R. M., Botella, C., Alcañiz, M., & Ávila, C. (2014). An fMRI Study to Analyze Neural Correlates of Presence during Virtual Reality Experiences. *Interacting with Computers, 26*(3), 269–284. https://doi.org/10.1093/iwc/iwt037

Corneille, O., & Gawronski, B. (2024). Self-reports are better measurement instruments than implicit measures. *Nature Reviews Psychology*, 3, 835–846. https://doi.org/10.1038/s44159-024-00376-z

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences, 111*(48), 17320-17325. https://doi.org/10.1073/pnas.1408988111

Cronin, K. A., Jacobson, S. L., Bonnie, K. E., & Hopper, L. M. (2017). Studying primate cognition in a social setting to improve validity and welfare: A literature review highlighting successful approaches. *PeerJ, 2017*(8), 1–35. https://doi.org/10.7717/peerj.3649

Cruz-Garza, J. G., Brantley, J. A., Nakagome, S., Kontson, K., Megjhani, M., Robleto, D., & Contreras-Vidal, J. L. (2017). Deployment of Mobile EEG Technology in an Art Museum Setting: Evaluation of Signal Quality and Usability. *Frontiers in Human Neuroscience,* 11. https://doi.org/10.3389/fnhum.2017.00527

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108*(2), 353–380. https://doi.org/10.1016/j.cognition.2008.03.006

Czeszumski, A., Eustergerling, S., Lang, A., Menrath, D., Gerstenberger, M., Schuberth, S., Schreiber, F., Rendon, Z. Z., & König, P. (2020). Hyperscanning: A Valid Method to Study Neural Inter-brain Underpinnings of Social Interaction. *Frontiers in Human Neuroscience*, 14, 1–17. https://doi.org/10.3389/fnhum.2020.00039

Dale, M. T., & Gawronski, B. (2023). Brains, trains, and ethical claims: Reassessing the normative implications of moral dilemma research. *Philosophical Psychology, 36*(1), 109–133. https://doi.org/10.1080/09515089.2022.2038783

Davidesco, I., Matuk, C., Bevilacqua, D., Poeppel, D., & Dikker, S. (2021). Neuroscience Research in the Classroom: Portable Brain Technologies in Education Research. *Educational Researcher, 50*(9), 649–656. https://doi.org/10.3102/0013189X211031563

Debener, S., Minow, F., Emkes, R., Gandras, K., & de Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology, 49*(11), 1617–1621. https://doi.org/10.1111/j.1469-8986.2012.01471.x

DeCaro, M. S. (2014). Enclothed cognition and controlled attention during insight problem-solving. *The Journal of Problem Solving, 7*(1), 8. https://doi.org/10.7771/1932-6246.1164

Decety, J., Chen, C., Harenski, C., & Kiehl, K. (2013). An fMRI study of affective perspective taking in individuals with psychopathy: imagining another in pain does not evoke empathy. *Frontiers in Human Neuroscience,* 7. https://www.frontiersin.org/articles/10.3389/fnhum.2013.00489

Demolliens, M., Isbaine, F., Takerkart, S., Huguet, P., & Boussaoud, D. (2017). Social and asocial prefrontal cortex neurons: A new look at social facilitation and the social brain. *Social Cognitive and Affective Neuroscience, 12*(8), 1241–1248. https://doi.org/10.1093/scan/nsx053

Deng, X., Chen, K., Chen, X., Zhang, L., Lin, M., Li, X., & Gao, Q. (2024). Parental involvement affects parent-adolescents brain-to-brain synchrony when experiencing different emotions together: an EEG-based hyperscanning study. *Behavioural brain research,* 458, 114734. https://doi.org/10.1016/j.bbr.2023.114734

di Fronso, S., Fiedler, P., Tamburro, G., Haueisen, J., Bertollo, M., & Comani, S. (2019). Dry EEG in Sports Sciences: A Fast and Reliable Tool to Assess Individual Alpha Peak Frequency Changes Induced by Physical Effort. *Frontiers in Neuroscience,* 13. https://doi.org/10.3389/fnins.2019.00982

Dikker, S., Michalareas, G., Oostrik, M., Serafimaki, A., Kahraman, H. M., Struiksma, M. E., & Poeppel, D. (2021). Crowdsourcing neuroscience: Inter-brain coupling during face-to-face interactions outside the laboratory. *NeuroImage*, 227. https://doi.org/10.1016/j.neuroimage.2020.117436

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. *Current Biology, 27*(9), 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002

Doell, K. C., Todorova, B., Vlasceanu, M., Bak Coleman, J. B., Pronizius, E., Schumann, P., Azevedo, F., Patel, Y., Berkebile-Wineberg, M. M., Brick, C., Lange, F., Grayson, S. J., Pei, Y., Chakroff, A., van den Broek, K. L., Lamm, C., Vlasceanu, D., Constantino, S. M., Rathje, S., … Van Bavel, J. J. (2024). The International Climate Psychology Collaboration: Climate change-related data collected from 63 countries. *Scientific Data, 11*(1), 1066. https://doi.org/10.1038/s41597-024-03865-1

Doyen, S., Klein, O., Pichon, C. L., & Cleeremans, A. (2012). Behavioral priming: it's all in the mind, but whose mind?. *PloS one, 7*(1), e29081. https://doi.org/10.1371/journal.pone.0029081

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES-a database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods, 42*(1), 351–362. https://doi.org/10.3758/BRM.42.1.351

Edwards, D. M., Franks, P., Friedgood, D., Lobban, G., & Mackay, H. C. G. (1969). An experiment on obedience. Unpublished student report, University of the Witwatersrand, Johannesburg, South Africa.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124. https://doi.org/10.1037/h0030377

Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019). The Psychology of Morality: A Review and Analysis of Empirical Studies Published From 1940 Through 2017. *Personality and Social Psychology Review, 23*(4), 332–366. https://doi.org/10.1177/1088868318811759

Everett, J. A. C., Pizarro, D. A., & Crockett, M. J. (2016). Inference of Trustworthiness From Intuitive Moral Judgments. *Journal of Experimental Psychology: General, 145*(6), 772–787. https://doi.org/10.1037/xge0000165

Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of a battery of computerized test systems by a troop of semi-free-ranging baboons (Papio papio). *Behavior Research Methods, 42*(2), 507–516. https://doi.org/10.3758/BRM.42.2.507

Fagot, J., Boë, L. J., Berthomier, F., Claidière, N., Malassis, R., Meguerditchian, A., Rey, A., & Montant, M. (2019). The baboon: A model for the study of language evolution. *Journal of Human Evolution, 126*, 39–50. https://doi.org/10.1016/j.jhevol.2018.10.006

Falk, E. B., Hyde, L. W., Mitchell, C., Faul, J., Gonzalez, R., Heitzeg, M. M., Keating, D. P., Langa, K. M., Martz, M. E., Maslowsky, J., Morrison, F. J., Noll, D. C., Patrick, M. E., Pfeffer, F. T., Reuter-Lorenz, P. A., Thomason, M. E., Davis-Kean, P., Monk, C. S., & Schulenberg, J. (2013). What is a representative brain? Neuroscience meets population science. *Proceedings of the National Academy of Sciences, 110*(44), 17615–17622. https://doi.org/10.1073/pnas.1310134110

Fan, S., Dal Monte, O., & Chang, S. W. C. (2021). Levels of naturalism in social neuroscience research. *IScience, 24*(7), 102702. https://doi.org/10.1016/j.isci.2021.102702

Fan, Y., & Han, S. (2008). Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study. *Neuropsychologia, 46*(1), 160-173. https://doi.org/10.1016/j.neuropsychologia.2007.07.023

Francis, K. B., Gummerum, M., Ganis, G., Howard, I. S., & Terbeck, S. (2018). Virtual morality in the helping professions: Simulated action and resilience. *British Journal of Psychology, 109*(3), 442–465. https://doi.org/10.1111/bjop.12276

Francis, K. B., Howard, C., Howard, I. S., Gummerum, M., Ganis, G., Anderson, G., & Terbeck, S. (2016). Virtual morality: Transitioning from moral judgment to moral action? *PLoS ONE, 11*(10), 1–22. https://doi.org/10.1371/journal.pone.0164374

Francis, K. B., Terbeck, S., Briazu, R. A., Haines, A., Gummerum, M., Ganis, G., & Howard, I. S. (2017). Simulating Moral Actions: An Investigation of Personal Force in Virtual Moral Dilemmas. *Scientific Reports, 7*(1), 1–11. https://doi.org/10.1038/s41598-017-13909-9

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review, 94*(4), 427. https://doi.org/10.1037/0033-295X.94.4.427

Gautheron, F., Quinton, J.-C., & Smeding, A. (2024). Conflict in moral and nonmoral decision making: an empirical study coupled with a computational model. *Cognitive Processing, 25*(2), 281–303. https://doi.org/10.1007/s10339-024-01178-0

Gautheron, F., Quinton, J.-C., Muller, D., & Smeding, A. (2023). Paradigm constraints on moral decision-making dynamics. *Journal of Behavioral Decision Making, 36*(4), e2324. https://doi.org/10.1002/bdm.2324

Gawronski, B., Armstrong, J., Conway, P., Friesdorf, R., & Hütter, M. (2017). Consequences, norms, and generalized inaction in moral dilemmas: The CNI Model of Moral Decision-Making. *Journal of Personality and Social Psychology, 113*(3), 343–376. https://doi.org/10.1037/pspa0000086

Gert, A. L., Ehinger, B. V., Timm, S., Kietzmann, T. C., & König, P. (2022). WildLab: A naturalistic free viewing experiment reveals previously unknown electroencephalography signatures of face processing. *The European journal of neuroscience, 56*(11), 6022–6038. https://doi.org/10.1111/ejn.15824

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and Conservatives Rely on Different Sets of Moral Foundations. Journal of Personality and Social Psychology, 96(5), 1029–1046. https://doi.org/10.1037/a0015141

Graham, J., Meindl, P., Beall, E., Johnson, K. M., & Zhang, L. (2016). Cultural differences in moral judgment and behavior, across and within societies. *Current Opinion in Psychology, 8*, 125–130. https://doi.org/10.1016/j.copsyc.2015.09.007

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the Moral Domain. *Journal of Personality and Social Psychology, 101*(2), 366–385. https://doi.org/10.1037/a0021847

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*(3), 1144–1154. https://doi.org/10.1016/j.cognition.2007.11.004

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron, 44*(2), 389–400. https://doi.org/10.1016/j.neuron.2004.09.027

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*(5537), 2105–2108. https://doi.org/10.1126/science.1062872

Greene, J., & Haidt, J. (2002). How (and where) does moral. *Trends in Cognitive Sciences, 6*(12), 517–523. http://tics.trends.com1364-6613/02/$-seefrontmatter

Gu, X., & Han, S. (2007). Attention and reality constraints on the neural processes of empathy for pain. *NeuroImage, 36*(1), 256–267. https://doi.org/10.1016/j.neuroimage.2007.02.025

Guerin, B. (2010). *Social facilitation.* Cambridge University Press.

Guglielmo, S. (2015). Moral judgment as information processing: An integrative review. *Frontiers in Psychology,* 6, 1–19. https://doi.org/10.3389/fpsyg.2015.01637

Gupta, I. (1983). Obedience to authority amongst university students: An experimental analysis. Unpublished doctoral dissertation, University of Delhi, Delhi, India.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814–834. https://doi.org/10.1037//0033-295x.108.4.814

Hamilton, L. S., & Huth, A. G. (2018). The revolution will not be controlled: natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience, 35*(5), 573–582. https://doi.org/10.1080/23273798.2018.1499946

Hammond, K. R., & Stewart, T. R. (2001). *The essential Brunswik: Beginnings, explications, applications*. New York: Oxford University Press.

Han, S., & Ma, Y. (2014). Cultural differences in human brain activity: A quantitative meta-analysis. *NeuroImage, 99*, 293–300. https://doi.org/10.1016/j.neuroimage.2014.05.062

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of Social Interaction in Human Brain Function. *Neuron, 88*(1), 181–193. https://doi.org/10.1016/j.neuron.2015.09.022

Harvey, E., Sandhaus, H., Jacobs, A. Z., Moss, E., & Sloane, M. (2024). The Cadaver in the Machine: The Social Practices of Measurement and Validation in Motion Capture Technology. Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, 1–23. https://doi.org/10.1145/3613904.3642004

Haslam, S. A., & Reicher, S. D. (2017). 50 years of "obedience to authority": From blind conformity to engaged followership. *Annual Review of Law and Social Science, 13*(1), 59-78. https://doi.org/10.1146/annurev-lawsocsci-110316-113710

Haslam, S. A., Reicher, S. D., Millard, K., & McDonald, R. (2015). 'Happy to have been of service': The Yale archive as a window into the engaged followership of participants in Milgram's 'obedience' experiments. *British Journal of Social Psychology, 54*(1), 55-83. https://doi.org/10.1111/bjso.12074

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Herrera C. D. (2001). Ethics, deception, and 'Those Milgram experiments'. *Journal of applied philosophy, 18*(3), 245–256. https://doi.org/10.1111/1468-5930.00192

Higgs S. (2015). Social norms and their influence on eating behaviours. *Appetite*, *86*, 38–44. https://doi.org/10.1016/j.appet.2014.10.021

Hobson, N. M., & Inzlicht, M. (2016). The mere presence of an outgroup member disrupts the brain's feedback-monitoring system. *Social Cognitive and Affective Neuroscience, 11*(11), 1698–1706. https://doi.org/10.1093/scan/nsw082

Hofstede, G. (1991). *Cultures and Organizations: Software of the Mind*. London, UK: McGraw-Hill.

Holland, C. A. C., Ebner, N. C., Lin, T., & Samanez-Larkin, G. R. (2019). Emotion identification across adulthood using the Dynamic FACES database of emotional expressions in younger, middle aged, and older adults. *Cognition and Emotion, 33*(2), 245–257. https://doi.org/10.1080/02699931.2018.1445981

Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2020). The 'Real-World Approach' and Its Problems: A Critique of the Term Ecological Validity. *Frontiers in Psychology, 11*, 1–12. https://doi.org/10.3389/fpsyg.2020.00721

Hortensius, R., & de Gelder, B. (2014). The neural basis of the bystander effect — The influence of group size on neural activity when witnessing an emergency. *NeuroImage, 93*, 53–58. https://doi.org/10.1016/j.neuroimage.2014.02.025

Huang, C., & Yu, R. (2018). Making mistakes in public: Being observed magnifies physiological responses to errors. *Neuropsychologia, 119*, 214–222. https://doi.org/10.1016/j.neuropsychologia.2018.08.015

Ioumpa, K., Gallo, S., Keysers, C., & Gazzola, V. (2024). Neural mechanisms of costly helping in the general population and mirror-pain synesthetes. *Scientific reports, 14*(1), 11617. https://doi.org/10.1038/s41598-024-62422-3

Jacobsen, N. S. J., Blum, S., Witt, K., & Debener, S. (2021). A walk in the park? Characterizing gait-related artifacts in mobile EEG recordings. *European Journal of Neuroscience, 54*(12), 8421-8440. https://doi.org/10.1111/ejn.14965

Jasińska, K. K., & Guei, S. (2018). Neuroimaging field methods using functional near infrared spectroscopy (nirs) neuroimaging to study global child development: rural sub-Saharan Africa. *Journal of visualized experiments: JoVE,* (132), 57165. https://doi.org/10.3791/57165

Jasińska, K., & Guei, S. (2022). Outcomes at the Bottom of the Pyramid. Learning, marginalization, and improving the quality of education in low-income countries. https://doi.org/10.11647/OBP.0256.13

Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review, 125*(2), 131–164. https://doi.org/10.1037/rev0000093

Kahneman, D. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica, 47*, 278. https://doi.org/10.2307/1914185

Kaspar, K., Krapp, V., & König, P. (2015). Hand washing induces a clean slate effect in moral judgments: a pupillometry and eye-tracking study. *Scientific reports,* 5, 10471. https://doi.org/10.1038/srep10471

Kim, H. S., & Sasaki, J. Y. (2014). Cultural neuroscience: Biology of the mind in cultural contexts. *Annual Review of Psychology*, 65, 487–514. https://doi.org/10.1146/annurev-psych-010213-115040

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization Theory and Research* (pp. 347-480). Chicago, IL: Rand McNally.

Körner, A., Deutsch, R., & Gawronski, B. (2020). Using the CNI model to investigate individual differences in moral dilemma judgments. *Personality and Social Psychology Bulletin, 46*(9), 1392–1407. https://doi.org/10.1177/0146167220907203

Körner, A., & Deutsch, R. (2023). Deontology and Utilitarianism in Real Life: A Set of Moral Dilemmas Based on Historic Events. *Personality & Social Psychology Bulletin, 49*(10), 1511–1528. https://doi.org/10.1177/01461672221103058

Kox E. & Caspar E. (2025). Behind bars: How prison shapes inmates' sense of agency and outcome processing. Manuscript submitted for publication.

LaBar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex, 13*(10), 1023-1033. https://doi.org/10.1093/cercor/13.10.1023

Ladouce, S., Donaldson, D. I., Dudchenko, P. A., & Ietswaart, M. (2017). Understanding Minds in Real-World Environments: Toward a Mobile Cognition Approach. *Frontiers in Human Neuroscience*, 10. https://doi.org/10.3389/fnhum.2016.00694

Ladouce, S., Mustile, M., Ietswaart, M., & Dehais, F. (2022). Capturing Cognitive Events Embedded in the Real World Using Mobile Electroencephalography and Eye-Tracking. *Journal of Cognitive Neuroscience, 34*(12), 2237–2255. https://doi.org/10.1162/jocn_a_01903

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Latané, B. (1981). The psychology of social impact. *American Psychologist, 36*(4), 343. https://doi.org/10.1037/0003-066X.36.4.343

Lau-Zhu, A., Lau, M. P. H., & McLoughlin, G. (2019). Mobile EEG in research on neurodevelopmental disorders: Opportunities and challenges. *Developmental Cognitive Neuroscience,* 36, 100635. https://doi.org/10.1016/j.dcn.2019.100635

Lee, M., Sul, S., & Kim, H. (2018). Social observation increases deontological judgments in moral dilemmas. *Evolution and Human Behavior, 39*(6), 611–621. https://doi.org/10.1016/j.evolhumbehav.2018.06.004

Lockwood, P. L., Klein-Flügge, M. C., Abdurahman, A., & Crockett, M. J. (2020). Model-free decision making is prioritized when learning to avoid harming others. *Proceedings of the National Academy of Sciences of the United States of America, 117*(44), 27719–27730. https://doi.org/10.1073/pnas.2010890117

López-Pérez, B., Ambrona, T., Wilson, E. L., & Khalil, M. (2016). The effect of enclothed cognition on empathic responses and helping behavior. *Social Psychology, 47*(4), 223–231. https://doi.org/10.1027/1864-9335/a000273

LoSchiavo, F. M., & Shatz, M. A. (2009). Reaching the neglected 95%. *American Psychologist, 64*(6), 565–566. https://doi.org/10.1037/a0016192

Lotto, L., Manfrinati, A., & Sarlo, M. (2014). A new set of moral dilemmas: Norms for moral acceptability, decision times, and emotional salience. *Journal of Behavioral Decision Making, 27*(1), 57-65. https://doi.org/10.1002/bdm.1782

Lucas, B. J., & Livingston, R. W. (2014). Feeling socially connected increases utilitarian choices in moral dilemmas. *Journal of Experimental Social Psychology*, 53, 1–4. https://doi.org/10.1016/j.jesp.2014.01.011

Malti, T., & Latzko, B. (2010). Children's moral emotions and moral cognition: towards an integrative perspective. *New directions for child and adolescent development, 2010*(129), 1–10. https://doi.org/10.1002/cd.272

Markus, H. R., & Kitayama, S. (2014). Culture and the self: Implications for cognition, emotion, and motivation. In: *College student development and academic life* (pp. 264-293). Routledge.

Martens, A., Kosloff, S., Greenberg, J., Landau, M. J., & Schmader, T. (2007). Killing begets killing: Evidence from a bug-killing paradigm that initial killing fuels subsequent killing. *Personality and Social Psychology Bulletin, 33*(9), 1251-1264. https://doi.org/10.1177/0146167207303020

Master, A., Markman, E. M., & Dweck, C. S. (2012). Thinking in categories or along a continuum: Consequences for children's social judgments. *Child Development, 83*(4), 1145–1163. https://doi.org/10.1111/j.1467-8624.2012.01774.x

Mathewson, K. E., Kuziek, J. P., Scanlon, J. E. M., & Robles, D. (2024). The moving wave: Applications of the mobile EEG approach to study human attention. *Psychophysiology*, *61*(9), e14603. https://doi.org/10.1111/psyp.14603

McGuire, J., Langdon, R., Coltheart, M., & Mackenzie, C. (2009). A reanalysis of the personal/impersonal distinction in moral psychology research. *Journal of Experimental Social Psychology, 45*(3), 577–580. https://doi.org/10.1016/j.jesp.2009.01.002

McKendrick, R., Parasuraman, R., Murtza, R., Formwalt, A., Baccus, W., Paczynski, M., & Ayaz, H. (2016). Into the Wild: Neuroergonomic Differentiation of Hand-Held and Augmented Reality Wearable Displays during Outdoor Navigation with Functional Near Infrared Spectroscopy. *Frontiers in Human Neuroscience,* 10. https://doi.org/10.3389/fnhum.2016.00216

Milgram, S. (1963). Behavioral Study of obedience. *The Journal of Abnormal and Social Psychology, 67(*4), 371–378. https://doi.org/10.1037/h0040525

Milgram, S. (1974). *Obedience to Authority: an Experiment View*. Harper & Row.

Milgram, S. (1984). Obedience. In R. J. Corsini (Ed.), *Encyclopedia of psychology* (Vol. 2, pp. 446–447). New York: Wiley.

Miller, G. A. (1969). Psychology as a means of promoting human welfare. *American psychologist, 24*(12), 1063. https://doi.org/10.1037/h0028988

Mironko, C. (2004). Igitero: means and motive in the Rwandan genocide. *Journal of Genocide Research, 6*(1), 47-60.

Mobbs, D., Yu, R., Rowe, J. B., Eich, H., FeldmanHall, O., & Dalgleish, T. (2010). Neural activity associated with monitoring the oscillating threat value of a tarantula. *Proceedings of the National Academy of Sciences, 107(*47), 20582-20586. https://doi.org/10.1073/pnas.1009076107

Moore, A. B., Lee, N. Y. L., Clark, B. A. M., & Conway, A. R. A. (2011). In defense of the personal / impersonal distinction in moral psychology research : Cross-cultural validation of the dual process model of moral judgment. *Judgment and Decision making,* 6(3), 186–195. https://doi.org/10.1017/S193029750000139X

Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., Grahe, J. E., McCarthy, R. J., Musser, E. D., Antfolk, J., Castille, C. M., Evans, T. R., Fiedler, S., Flake, J. K., Forero, D. A., Janssen, S. M. J., Keene, J. R., Protzko, J., Aczel, B., … Chartier, C. R. (2018). The Psychological Science Accelerator: Advancing Psychology Through a Distributed Collaborative Network. *Advances in Methods and Practices in Psychological Science, 1*(4), 501–515. https://doi.org/10.1177/2515245918797607

Nathan, K., & Contreras-Vidal, J. L. (2016). Negligible motion artifacts in scalp electroencephalography (EEG) during treadmill walking. *Frontiers in Human Neuroscience*, 9, 708. https://doi.org/10.3389/fnhum.2015.00708

Newman, A., Bavik, Y. L., Mount, M., & Shao, B. (2021). Data Collection via Online Platforms: Challenges and Recommendations for Future Research. *Applied Psychology, 70*(3), 1380–1402. https://doi.org/10.1111/apps.12302

Nick, C. (producer). (2010). Le Jeu de la Mort [The Game of Death]. France Television.

Nickel, P. J. (2009). Trust, staking, and expectations. Journal for the Theory of Social Behaviour, 39(3), 345-362. https://doi.org/10.1111/j.1468-5914.2009.00407.x

Noh, S. R., & Isaacowitz, D. M. (2013). Emotional faces in context: Age differences in recognition accuracy and scanning patterns. *Emotion, 13*(2), 238–249. https://doi.org/10.1037/a0030234

Northoff, G., & Heinzel, A. (2006). First-Person Neuroscience: A new methodological approach for linking mental and neuronal states. *Philosophy, Ethics, and Humanities in Medicine, 1*(1), 3. https://doi.org/10.1186/1747-5341-1-3

Norton, E. S., Manning, B. L., Harriott, E. M., Nikolaeva, J. I., Nyabingi, O. S., Fredian, K. M., Page, J. M., McWeeny, S., Krogh-Jespersen, S., MacNeill, L. A., Roberts, M. Y., & Wakschlag, L. S. (2022). Social EEG: A novel neurodevelopmental approach to studying brain-behavior links and brain-to-brain synchrony during naturalistic toddler-parent interactions. *Developmental psychobiology, 64*(3), e22240. https://doi.org/10.1002/dev.22240

Nozawa, T., Sakaki, K., Ikeda, S., Jeong, H., Yamazaki, S., Kawata, K. H. D. S., ... & Kawashima, R. (2019). Prior physical synchrony enhances rapport and inter-brain synchronization during subsequent educational communication. *Scientific reports, 9*(1), 12747. https://doi.org/10.1038/s41598-019-49257-z

Olson, J. A., Sandra, D. A., Colucci, É. S., Al Bikaii, A., Chmoulevitch, D., Nahas, J., Raz, A., & Veissière, S. P. L. (2022). Smartphone addiction is increasing across the world: A meta-analysis of 24 countries. *Computers in Human Behavior,* 129, 107138. https://doi.org/10.1016/j.chb.2021.107138

Parker, R. S., Haytko, D., & Hermans, C. (2009). Individualism and collectivism: Reconsidering old assumptions. *Journal of International Business Research, 8*, 127–139.

Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J., & Richardson, D. C. (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy* *of Sciences of the United States of America, 112*(13), 4170–4175. https://doi.org/10.1073/pnas.1415250112

Pascual, A., Felonneau, M. L., Guéguen, N., & Lafaille, E. (2014). Conformity, obedience to authority, and compliance without pressure to control cigarette butt pollution. *Social Influence, 9*(2), 83–98. https://doi.org/10.1080/15534510.2013.778214

Patil, I., Cogoni, C., Zangrando, N., Chittaro, L., & Silani, G. (2014). Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. *Social Neuroscience, 9*(1), 94–107. https://doi.org/10.1080/17470919.2013.870091

Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science, 36*(1), 163–177. https://doi.org/10.1111/j.1551-6709.2011.01210.x

Pech, G. & Caspar, E.A. (2023). Does the cowl make the monk? The effect of military and Red Cross uniforms on empathy for pain, sense of agency and moral behaviors. *Frontiers in Psychology*, 14. https://doi.org/10.3389/fpsyg.2023.1255835

Pech, P.P., & Caspar, E.A. (in press). A Cross-Cultural EEG Study of How Obedience and Conformity Influence Reconciliation Intentions. *Social, Cognitive and Affective Neuroscience*.

Pech, G. P., Gishoma, D., & Caspar, E. A. (2024). A novel electroencephalography-based paradigm to measure intergroup prosociality: An intergenerational study in the aftermath of the genocide against Tutsis in Rwanda. *Journal of Experimental Psychology. General, 153*(1), 241–254. https://doi.org/10.1037/xge0001480

Pechtel, P., & Pizzagalli, D. A. (2011). Effects of early life stress on cognitive and affective function: an integrated review of human literature. *Psychopharmacology, 214*(1), 55–70. https://doi.org/10.1007/s00213-010-2009-2

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), e13331. https://doi.org/10.1111/psyp.13331

Pinti, P., Merla, A., Aichelburg, C., Lind, F., Power, S., Swingler, E., Hamilton, A., Gilbert, S., Burgess, P. W., & Tachtsidis, I. (2017). A novel GLM-based method for the Automatic IDentification of functional Events (AIDE) in fNIRS data recorded in naturalistic environments. *NeuroImage, 155*, 291–304. https://doi.org/10.1016/j.neuroimage.2017.05.001

Piper, S. K., Krueger, A., Koch, S. P., Mehnert, J., Habermehl, C., Steinbrink, J., Obrig, H., & Schmitz, C. H. (2014). A wearable multi-channel fNIRS system for brain imaging in freely moving subjects. *Neuroimage,* 85, 64-71. https://doi.org/10.1016/j.neuroimage.2013.06.062

Poulsen, A. T., Kamronn, S., Dmochowski, J., Parra, L. C., & Hansen, L. K. (2017). EEG in the classroom: Synchronised neural recordings during video presentation. *Scientific Reports, 7*(1), 43916. https://doi.org/10.1038/srep43916

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. Nature Reviews Neuroscience, 20(8), 495–505. https://doi.org/10.1038/s41583-019-0179-4

Reicher, S. D., & Haslam, S. A. (2020). Engaged followership and engaged fellowship: Toward a unified analysis of harm-doing and helping. In L. S. Newman (Ed.*), Confronting humanity at its worst: Social psychological perspectives on genocide* (pp. 195–219). Oxford University Press. https://doi.org/10.1093/oso/9780190685942.003.0008

Reis, P., Hebenstreit, F., Gabsteiger, F., von Tscharner, V., & Lochmann, M. (2014). Methodological aspects of EEG and body dynamics measurements during motion. *Frontiers in Human Neuroscience*, 8. https://doi.org/10.3389/fnhum.2014.00156

Richer, N., Bradford, J. C., & Ferris, D. P. (2024). Mobile neuroimaging: What we have learned about the neural control of human walking, with an emphasis on EEG-based research. *Neuroscience & Biobehavioral Reviews, 162,* 105718. https://doi.org/10.1016/j.neubiorev.2024.105718

Roberts, G., Holmes, N., Alexander, N., Boto, E., Leggett, J., Hill, R. M., Shah, V., Rea, M., Vaughan, R., Maguire, E. A., Kessler, K., Beebe, S., Fromhold, M., Barnes, G. R., Bowtell, R., & Brookes, M. J. (2019). Towards OPM-MEG in a virtual reality environment. *NeuroImage*, 199, 408–417. https://doi.org/10.1016/j.neuroimage.2019.06.010

Ronca, V., Flumeri, G. Di, Giorgi, A., Vozzi, A., Capotorto, R., Germano, D., Sciaraffa, N., Borghini, G., Babiloni, F., & Aricò, P. (2024). o-CLEAN: a novel multi-stage algorithm for the ocular artifacts' correction from EEG data in out-of-the-lab applications. *Journal of Neural Engineering, 21*(5), 56023. https://doi.org/10.1088/1741-2552/ad7b78

Schein, C., & Gray, K. (2018). The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm. *Personality and Social Psychology Review, 22*(1), 32–70. https://doi.org/10.1177/1088868317698288

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences, 36*(4), 393–414. https://doi.org/10.1017/S0140525X12000660

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. Trends in cognitive sciences, 10(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Seyll, L., Sezibera, V., Masabo, F., & Caspar, E.A. (2025). *Neural processing of obedience and resistance among former genocide perpetrators and rescuers*. Manuscript submitted for publication.

Shanab, M. E., & Yahya, K. A. (1978). A cross-cultural study of obedience. *Bulletin of the Psychonomic Society, 11*(4), 267–269. https://doi.org/10.3758/BF03336827

Sharma, D., Booth, R., Brown, R., & Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the Stroop task. *Psychonomic Bulletin & Review, 17*(1), 52–58. https://doi.org/10.3758/PBR.17.1.52

Shaver, K. G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness*. New York, NY: Springer-Verlag.

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on Generality (COG): A Proposed Addition to All Empirical Papers. *Perspectives on Psychological Science, 12*(6), 1123–1128. https://doi.org/10.1177/1745691617708630

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N., & Sanchez-Vives, M. V. (2006). A virtual reprise of the Stanley Milgram obedience experiments. *PloS one, 1*(1), e39. https://doi.org/10.1371/journal.pone.0000039

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science, 303*(5661), 1157–1162. https://doi.org/10.1126/science.1093535

Skulmowski, A., Bunge, A., Kaspar, K., & Pipa, G. (2014). Forced-choice decision-making in modified trolley dilemma situations: a virtual reality and eye tracking study. *Frontiers in Behavioral Neuroscience*, 8, 426. https://doi.org/10.3389/fnbeh.2014.00426

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic Stimuli in Neuroscience: Critically Acclaimed. *Trends in Cognitive Sciences, 23*(8), 699–714. https://doi.org/10.1016/j.tics.2019.05.004

Spatola, N., Belletier, C., Normand, A., Chausse, P., Monceau, S., Augustinova, M., Barra, V., Huguet, P., & Ferrand, L. (2018). Not as bad as it seems: When the presence of a threatening humanoid robot improves human performance. *Science Robotics, 3*(21), 2–4. https://doi.org/10.1126/scirobotics.aat5843

Stangl, M., Maoz, S. L., & Suthana, N. (2023). Mobile cognition: imaging the human brain in the 'real world.' *Nature Reviews Neuroscience, 24*(6), 347–362. https://doi.org/10.1038/s41583-023-00692-y

Staub, E. (2014). Obeying, Joining, Following, Resisting, and Other Processes in the Milgram Studies, and in the Holocaust and Other Genocides: Situations, Personality, and Bystanders. *Journal of Social Issues, 70*(3), 501–514. https://doi.org/10.1111/josi.12074

Stringfellow, J. S., Liran, O., Lin, M.-H., & Baker, T. E. (2024). Recording Neural Reward Signals in a Naturalistic Operant Task Using Mobile-EEG and Augmented Reality. *ENeuro, 11*(8). https://doi.org/10.1523/ENEURO.0372-23.2024

Tassy, S., Oullier, O., Mancini, J., & Wicker, B. (2013). Discrepancies between judgment and choice of action in moral dilemmas. *Frontiers in Psychology*, 4, 1–8. https://doi.org/10.3389/fpsyg.2013.00250

Tebes, J. K. (2000). External validity and scientific psychology. *American Psychologist*, *55*(12), 1508–1509. https://doi.org/10.1037/0003-066X.55.12.1508

Testard, C., Tremblay, S., & Platt, M. (2021). From the field to the lab and back: neuroethology of primate social behavior. *Current Opinion in Neurobiology,* 68, 76–83. https://doi.org/10.1016/j.conb.2021.01.005

Thalmayer, A. G., Toscanelli, C., & Arnett, J. J. (2021). The neglected 95% revisited: Is American psychology becoming less American?. *American Psychologist, 76*(1), 116–129. https://doi.org/10.1037/amp0000622

Thébault Guiochon, A., Vives, E., Revillard, V., Trémolière, B., Djeriouat, H. (2025). *Sex Differences in Audiovisual Integration of Moral Appraisal in Third-Party Transgressions*. Manuscript submitted for publication.

Tian, T., Feng, X., Gu, R., Broster, L. S., Feng, C., Wang, L., Guan, Q., & Luo, Y. J. (2015). Modulation of the brain activity in outcome evaluation by the presence of an audience: An electrophysiological investigation. *Brain Research,* 1615, 139–147. https://doi.org/10.1016/j.brainres.2015.04.040

Toelch, U., & Dolan, R. J. (2015). Informational and Normative Influences in Conformity from a Neurocomputational Perspective. *Trends in cognitive sciences*, *19*(10), 579–589. https://doi.org/10.1016/j.tics.2015.07.007

Tricoche, L., & Caspar, E. A. (2024, December 3). The power of social influence on (non-)moral decision-making: A review of behavioral effects and neural mechanisms. https://doi.org/10.31219/osf.io/s6jtk

Tricoche, L., Pélisson, D., Longo, L., Koun, E., Poisson, A., Prado, J., & Meunier, M. (2023). Task-independent neural bases of peer presence effect on cognition in children and adults. *NeuroImage,* 277, 120247. https://doi.org/10.1016/j.neuroimage.2023.120247

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *The American journal of psychology, 9*(4), 507-533.

Tromp, J., Peeters, D., Meyer, A. S., & Hagoort, P. (2018). The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods, 50*(2), 862–869. https://doi.org/10.3758/s13428-017-0911-9

TRUST (2018) The TRUST Code – A Global Code of Conduct for Equitable Research Partnerships, DOI: https://doi.org/10.48508/GCC/2018.05

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.

Uziel, L. (2007). Individual differences in the social facilitation effect: A review and meta-analysis. *Journal of Research in Personality, 41*(3), 579–601. https://doi.org/10.1016/j.jrp.2006.06.008

Van Bavel, J. J., FeldmanHall, O., & Mende-Siedlecki, P. (2015). The neuroscience of moral cognition: From dual processes to dynamic systems. *Current Opinion in Psychology*, 6, 167–172. https://doi.org/10.1016/j.copsyc.2015.08.009

Vazire, S., Schiavone, S. R., & Bottesini, J. G. (2022). Credibility Beyond Replicability: Improving the Four Validities in Psychological Science. *Current Directions in Psychological Science, 31*(2), 162–168. https://doi.org/10.1177/09637214211067779

Vives, E., Thébault Guiochon, A., Tremolière, B., Falco, A., & Djeriouat, H. (2024) *Moral Anger and Disgust: Examining the Recipient vs. Initiator Focus in the Moral Evaluation of Transgressions*. Manuscript submitted for publication. https://doi.org/10.31234/osf.io/3dzp4

Wagner, S. L., Martin, C. A., & McFee, J. A. (2009). Investigating the “rescue personality”. *Traumatology*, *15*(3), 5–12. https://doi.org/10.1177/1534765609338499

Yarkoni T. (2020). The generalizability crisis. *The Behavioral and Brain Sciences*, 45, e1. https://doi.org/10.1017/S0140525X20001685

Young, L., & Saxe, R. (2009). An fMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience, 21*(7), 1396–1405. https://doi.org/10.1162/jocn.2009.21137

Zajonc, R. B. (1965). Social facilitation. *Science*, 149, 269–274. https://doi.org/10.1126/science.149.3681.269

Zapała, D., Augustynowicz, P., & Tokovarov, M. (2022). Recognition of Attentional States in VR Environment: An fNIRS Study. *Sensors, 22*(9), 3133. https://doi.org/10.3390/s22093133

Zheng, J., Hu, L., Li, L., Shen, Q., & Wang, L. (2021). Confidence modulates the conformity behavior of the investors and neural responses of social influence in crowdfunding. *Frontiers in Human Neuroscience,* 15, 766908.