Virtual Reality Paradigms for the Assessment of Human Anxiety-Related Behavior

Examining Reliability, Individual Differences, and the Introduction of a Novel Paradigm

Dissertation

zur Erlangung des Doktorgrades

Doctor of Philosophy (PhD)

der Medizinischen Fakultät der Universität Hamburg

Vorgelegt von

Lateefah Ayashe Roth

aus Frankfurt am Main

Hamburg, 2025

Betreuer:in / Gutachter:in der Dissertation: Prof. Dr. Johannes Fuß

Gutachter:in der Dissertation: PD. Dr. Sarah Biedermann

Vorsitz der Prüfungskommission: Prof. Dr. Johannes Fuß

Mitglied der Prüfungskommission: Prof. Dr. Peer Briken

Mitglied der Prüfungskommission: PD. Dr. Sarah Biedermann

Datum der mündlichen Prüfung: 05.06.2026

“We’re not, after all, separate from the animal kingdom. We’re part of it.”

— Jane Goodall

Table of Contents

Summary ... 1

Zusammenfassung ... 4

List of PhD-Related Publications ... 8

Statement on Use of Published Material ... 9

Chapter 1 Introduction ... 10

Fear and Anxiety: Emotional Responses to Threat ... 10

Animal Models of Fear and Anxiety ... 12

Bridging the Gap Between Animal Models and Human Anxiety Research Through Virtual Reality. 16

Challenges and Opportunities in Applying VR-Based Anxiety Assessment ... 19

Aims of the Dissertation ... 22

Chapter 2 Methods ... 25

General Methods ... 25

Setup and Virtual Environments ... 25

Questionnaires ... 28

Physiological and Endocrinological Measurements ... 30

Heatmaps ... 31

Ethical Considerations ... 32

Study-Specific Methods ... 33

Study 1: Reliability of Repeated Measures on the Virtual EPM ... 33

Study 2: The Impact of Psychopathic Traits on Anxiety-Related Behaviors on the Virtual EPM . 36

Study 3: The Effect of Gaming Experience on Behavioral Responses on the Virtual EPM ... 38

Study 4: Validation of the Dark-Light Box in Virtual Reality ... 40

Chapter 3 Results ... 49

Study 1: Reliability of Repeated Measures on the Virtual EPM ... 49

Anxiety-Related Behaviors Upon Repeated Exposure ... 49

The Influence of Repeated Exposure to the EPM on Autonomic Reactions ... 51

The Subjective Response to Repeated EPM Exposure ... 53

Study 2: The Impact of Psychopathic Traits on Anxiety-Related Behaviors on the Virtual EPM ... 54

Associations of Psychopathic Traits With Measures of Anxiety-Related Constructs ... 54

Predictors and Spatial Patterns of EPM Behavior ... 56

Study 3: The Effect of Gaming Experience on Behavioral Responses on the Virtual EPM ... 59

Effects of Gaming Experience and Gender on EPM Behavior 59

Study 4: Validation of the Dark-Light Box in Virtual Reality *62*

Healthy Sample 62

Clinical Sample 72

Chapter 4 Discussion 76

Summary of Main Findings 76

Reliability and Validity of VR-Based Anxiety Paradigms 77

Further Evidence From the DLB: Clinical and Cross-Paradigm Analyses 81

Trait- and Experience-Related Variability in VR Anxiety Behavior 84

Methodological Considerations and Limitations 89

Conclusions and Implications for Translational Anxiety Research 91

Future Directions 94

Chapter 5 References 98

Chapter 6 Acknowledgements 123

Chapter 7 Supplementary Material 125

Curriculum Vitae 136

Der Lebenslauf wurde aus datenschutzrechtlichen Gründen entfernt

Summary

Approach-avoidance conflict paradigms are central tools in animal anxiety research. Setups like the Elevated Plus Maze (EPM), Open Field Test (OFT), and Dark-Light Box (DLB) are considered gold-standard tasks and build on rodents' innate aversion to open, elevated, or brightly lit areas. The natural curiosity for exploration, together with the animal's preference for enclosed, darker spaces, creates such an approach-avoidance conflict. This allows for the measurement of avoidance behaviors, such as reduced time in aversive compartments or longer latencies before entering them, which are considered behavioral markers of anxiety. Translating these paradigms for human use allows for the investigation of authentic human anxiety-related behavioral traits, greatly advancing conventional computer and self-report-based assessments. Virtual reality (VR) provides a powerful framework for this translation, as it enables the creation of immersive, lifelike environments that preserve ecological validity while offering experimental control.

The overarching aim of this dissertation was to investigate VR-based paradigms for assessing human anxiety-related behavior, addressing three central objectives: (1) testing the test-retest reliability of the already validated virtual EPM, (2) examining the influence of individual differences, specifically psychopathic traits and prior gaming experience, on behavior measured in the EPM, and (3) introducing and evaluating the validity of a novel virtual DLB paradigm.

Study 1 examined whether the virtual EPM is suitable for repeated testing in a sample of healthy participants (*N* = 39), an essential requirement for clinical and pharmacological applications. To minimize habituation and sensitization, strategies adapted from rodent research were applied: a 28-day inter-trial interval and a variation of the test environment. For the different test environments in VR, two versions were used: a stylized, game-like EPM version and a naturalistic, real-world resembling EPM. Results indicated that behavioral parameters showed strong correlations and no significant difference between first and repeated exposure. Likewise, psychophysiological responses (salivary alpha-amylase, skin conductance, and respiratory rate) increased at first and second exposure. Although skin conductance and heart rate showed slightly elevated responses at repeated exposure, indicating anticipatory anxiety and a mild sensitization effect. All results were found to be independent of the EPM version. Importantly, both the stylized and naturalistic versions yielded comparable outcomes, suggesting that a more realistic environment is not superior in terms of eliciting lifelike physiological, emotional, and behavioral responses. Taken together, these

findings demonstrate that the virtual EPM can be reliably used in longitudinal human studies, provided that appropriate retest strategies are implemented.

Study 2 examined anxiety-related behavior in relation to psychopathic traits in a non-clinical sample of healthy participants (*N* = 160) using the virtual EPM. Psychopathic traits were assessed with the Brief Questionnaire of Psychopathic Personality Traits (FPP), and total scores were significantly associated with reduced anxiety-related behavior. Multivariate regression analyses further revealed that the subdimensions Fearlessness and Lack of Empathy were significant predictors. These results provide the first direct evidence linking psychopathic traits to observable anxiety-related behavior in an immersive paradigm, lending support to the low-anxiety hypothesis of psychopathy, which relates to the ongoing debate about anxiety deficits in psychopathy and their possible influence on individual behavior. Beyond theoretical implications, these findings highlight the potential of VR-based tasks to capture clinically and forensically relevant behavioral tendencies in controlled laboratory settings.

Study 3 investigated whether prior gaming experience, assessed as a binary yes/no measure of regular gaming (at least weekly), influenced behavior in the virtual EPM and whether it could account for observed differences in EPM behavior of men and women. In a large sample of healthy participants (*N* = 428), gaming experience was associated with reduced anxiety-related behavior in women, while effects were negligible in men. Overall, men displayed less anxiety-related behavior than women, consistent with previous EPM findings in humans. These results identify gaming experience as an important experiential factor that may confound or modulate behavioral responses in VR and emphasize the need to systematically assess and control for it in future studies using immersive VR paradigms.

Study 4 introduced the DLB as a novel VR paradigm for assessing human anxiety-related behavior. In total, 74 healthy participants were tested in the virtual DLB. Exposure to the DLB elicited robust increases in heart rate, respiratory rate, skin conductance, and salivary alpha-amylase, together with a gradual rise in subjective anxiety from bright to dark areas, supporting face validity. Content validity was supported by the consistent pattern of subjective, physiological, and partly endocrine responses, but weakened by the lack of robust effects in classical avoidance indices (e.g., time in dark, covered distance). The ethological marker “arms-before-body” emerged as the most reliable behavioral marker for anxiety and differed between high- and low-anxiety participants. Concurrent validity was likewise partly supported, with convergence of subjective and physiological responses, but again limited by weak support from classical avoidance measures. Predictive validity was indicated by multivariate multiple regression analysis showing overall effects of sensation seeking and subjective anxiety on

DLB behavior. Subsequent coefficient *t*-tests with robust standard errors revealed that sensation seeking predicted greater exploration (covered distance, re-examinations), while subjective anxiety predicted arms-before-body behavior. Construct validity is conceptually strong due to the close alignment with rodent models of approach-avoidance conflict, but empirically incomplete given the limited robustness of behavioral findings and the absence of pharmacological testing.

In a clinical sample (*n* = 40 patients and individually matched controls), including anxiety disorders (AD), post-traumatic stress disorder (PTSD), and borderline personality disorder (BPD), no robust transdiagnostic patient-control differences were found. Although based on small samples and uncorrected tests, exploratory subgroup analyses suggested diagnosis-specific patterns. AD/PTSD patients showed reduced exploration, whereas BPD patients displayed a mixed profile, characterized by greater time spent in dark areas but heterogeneous effects across other measures. Cross-paradigm analyses with the EPM revealed convergent moderate correlations for exploration and latency measures, but also showed complementarity of the setups. On the one hand, the absence of strong correlations indicated that the paradigms are not redundant. On the other hand, ethological markers coded in the DLB, but not yet assessed in the EPM, captured behavioral dimensions beyond conventional indices. Taken together, these findings position the DLB as a promising proof of concept: it demonstrates strong face validity and partial support across content, concurrent, predictive, and construct domains, while highlighting the need for refinement and further validation, particularly in clinical and pharmacological contexts.

Across studies, this dissertation demonstrates that VR-based paradigms can yield reliable and ecologically valid measures of human anxiety-related behavior while also being sensitive to stable personality traits and experiential factors. At the same time, the findings illustrate that careful consideration of protocol design, participant characteristics, and validation criteria is essential for their broader translational and clinical application. VR paradigms therefore hold considerable potential as a methodological bridge between animal and human anxiety research but require continued refinement to establish their robustness and clinical utility.

Zusammenfassung

Annäherungs-Vermeidungs-Konfliktparadigmen sind zentrale Werkzeuge der Angstforschung bei Tieren. Versuchsaufbauten wie das Elevated Plus Maze (EPM), der Open Field Test (OFT) und die Dark-Light Box (DLB) gelten als Goldstandard und basieren auf der angeborenen Aversion von Nagetieren gegenüber offenen, erhöhten oder hell beleuchteten Bereichen. Das Zusammenspiel aus natürlicher Neugierde der Tiere und der Präferenz für geschützte, dunklere Bereiche erzeugt dabei einen Annäherungs-Vermeidungs-Konflikt. Dies ermöglicht die Messung von Vermeidungsverhalten, wie etwa einer verkürzten Aufenthaltsdauer in aversiven Bereichen oder verlängerten Latenzen bis zu deren erstmaligem Betreten, die als Verhaltensmarker für Angst gelten. Die Übertragung dieser Paradigmen auf den Menschen eröffnet die Möglichkeit, authentische angstbezogene Verhaltensweisen zu erfassen und konventionelle computerbasierte sowie selbstberichtsbasierte Verfahren entscheidend zu erweitern. Virtuelle Realität (VR) bietet hierfür einen besonders geeigneten Rahmen, da sie immersive, lebensnahe Umgebungen mit ökologischer Validität schafft und gleichzeitig eine präzise experimentelle Kontrolle ermöglicht.

Das übergeordnete Ziel dieser Dissertation war es, VR-basierte Paradigmen zur Erfassung menschlichen Angstverhaltens zu untersuchen, wobei drei Schwerpunkte verfolgt wurden: (1) die Prüfung der Reliabilität des bereits validierten virtuellen EPM bei wiederholter Testung, (2) die Untersuchung individueller Unterschiede, namentlich psychopathischer Persönlichkeitsmerkmale und vorheriger Spielerfahrung und deren Einfluss auf das Verhalten im virtuellen EPM, sowie (3) die Einführung und Validitätsprüfung eines neu entwickelten virtuellen DLB-Paradigmas.

Studie 1 prüfte an gesunden Proband:innen (N = 39), ob das virtuelle EPM für wiederholte Testungen geeignet ist, eine wesentliche Voraussetzung für klinische und pharmakologische Anwendungen. Zur Minimierung von Habituation und Sensitivierung wurden Strategien aus der Tierforschung übernommen: ein 28-tägiges Intervall zwischen den Testungen sowie die Variation der Testumgebung (spielähnliche vs. naturalistische VR-Version). Die Ergebnisse zeigten starke Korrelationen der Verhaltensmaße und keine signifikanten Unterschiede zwischen erster und wiederholter Testung. Psychophysiologische Reaktionen (Speichel- Alpha-Amylase, Hautleitfähigkeit, Atemfrequenz) stiegen in beiden Sitzungen an. Hautleitfähigkeit und Herzrate waren bei der Wiederholung leicht erhöht, was auf antizipatorische Angst und eine milde Sensitisierung hindeutet. Die Ergebnisse waren unabhängig von der VR-Version, sodass eine realistischere Umgebung keinen Vorteil für physiologische, emotionale oder behaviorale Reaktionen bot. Insgesamt belegen die

Ergebnisse, dass das virtuelle EPM für longitudinale Studien verlässlich einsetzbar ist, sofern geeignete Wiederholungskonzepte berücksichtigt werden.

Studie 2 untersuchte das angstbezogene Verhalten in Bezug auf psychopathische Merkmale in einer nicht-klinischen Stichprobe (*N* = 160) mithilfe des virtuellen EPM. Erfasst wurden die Merkmale mit dem Brief Questionnaire of Psychopathic Personality Traits (kurzer Fragebogen zu psychopathischen Persönlichkeitsmerkmalen; FPP). Die Gesamt-Scores korrelierten signifikant mit reduziertem angstbezogenem Verhalten. Multivariate Regressionsanalysen zeigten insbesondere die Subdimensionen Fearlessness (Furchtlosigkeit) und Lack of Empathy (Mangel an Empathie) als signifikante Prädiktoren. Dies stellt den ersten direkten Nachweis einer Verbindung zwischen psychopathischen Persönlichkeitsmerkmalen und beobachtbarem angstbezogenem Verhalten in einem immersiven Paradigma dar und stützt die Low-Anxiety-Hypothese der Psychopathieforschung, die sich auf die anhaltende Debatte über Angstdefizite bei Psychopathie und deren möglichen Einfluss auf das individuelle Verhalten bezieht. Über theoretische Implikationen hinaus verdeutlicht dies das Potenzial von VR-basierten Paradigmen, klinisch und forensisch relevante Verhaltenstendenzen unter kontrollierten Laborbedingungen abzubilden.

Studie 3 untersuchte, ob vorherige Spielerfahrung (erfasst als dichotome Variable: regelmäßiges Spielen mindestens einmal wöchentlich) das Verhalten im virtuellen EPM beeinflusst und ob dadurch beobachtete Geschlechtsunterschiede im EPM-Verhalten bei Menschen erklärt werden können. In einer großen Stichprobe (*N* = 428) zeigte sich, dass Spielerfahrung bei Frauen mit reduziertem angstbezogenem Verhalten verbunden war, während die Effekte bei Männern vernachlässigbar waren. Männer zeigten insgesamt weniger angstbezogenes Verhalten als Frauen, was frühere Befunde zum virtuellen EPM bestätigt. Diese Ergebnisse identifizieren Spielerfahrung als wichtigen erfahrungsbedingten Faktor, der Verhalten in VR-Paradigmen modulieren oder verzerren kann. Dies unterstreicht die Notwendigkeit, Spieleerfahrung in künftigen Studien systematisch zu erfassen und zu kontrollieren.

Studie 4 führte die DLB als neuartiges VR-Paradigma zur Erfassung angstbezogenen Verhaltens an gesunden Proband:innen (*N* = 74) ein. Die Exposition in der DLB rief deutliche Anstiege von Herzrate, Atemfrequenz, Hautleitfähigkeit und Speichel-Alpha-Amylase hervor, begleitet von einem graduellen Anstieg subjektiver Angst von hellen zu dunklen Bereichen, was die Augenscheinvalidität unterstützt. Inhaltsvalidität wurde durch die konsistenten Muster subjektiver, physiologischer und teilweise endokrinologischer Reaktionen gestützt, jedoch durch fehlende stabile Effekte klassischer Vermeidungsmaße (z. B. Zeit im Dunkeln, zurückgelegte Distanz) abgeschwächt. Das ethologische Maß „Arme-vor-dem-Körper“ erwies sich als stabilster Verhaltensmarker für Angst und zeigte Unterschiede zwischen Proband:innen mit hoher und niedriger subjektiver Angst. Die Übereinstimmungsvalidität wurde teilweise gestützt durch die Konvergenz subjektiver und physiologischer Reaktionen, blieb aber durch schwache klassische Vermeidungsmaße eingeschränkt. Prädiktive Validität zeigte sich in einer multivariaten Regressionsanalyse, die Gesamteffekte von Sensation Seeking und subjektiver Angst auf das Verhalten bestätigte. Nachgelagerte *t*-Tests der Regressionskoeffizienten zeigten, dass Sensation Seeking stärkeres Explorationsverhalten (zurückgelegte Distanz, Re-Examinationen) vorhersagte, während subjektive Angst mit Arme-vor-dem-Körper-Verhalten assoziiert war. Die Konstruktvalidität ist konzeptionell stark durch die enge Anlehnung an Tiermodelle von Annäherungs-Vermeidungs-Konflikten, empirisch jedoch unvollständig aufgrund begrenzter stabiler Verhaltenseffekte und fehlender pharmakologischer Validierung.

In einer klinischen Stichprobe ($n = 40$ Patient:innen individuell gepaart mit Kontrollen), einschließlich Angststörungen (AD), Posttraumatischer Belastungsstörung (PTBS) und Borderline-Persönlichkeitsstörung (BPS), zeigten sich keine robusten transdiagnostischen Unterschiede. Explorative Subgruppenanalysen, basierend auf kleinen Stichproben und unkorrigierten Tests, deuten jedoch auf diagnosespezifische Muster hin: AD/PTBS-Patient:innen zeigten reduziertes Explorationsverhalten, während BPS-Patient:innen ein gemischtes Profil mit längerer Zeit im Dunkeln, jedoch heterogenen Effekten in anderen Maßen aufwiesen. Kreuzparadigmatische Analysen mit dem EPM ergaben konvergente moderate Korrelationen für Explorations- und Latenzmessungen, zeigten aber auch die Komplementarität der Setups. Einerseits deutete das Fehlen starker Korrelationen darauf hin, dass die Paradigmen nicht redundant sind. Andererseits erfassten ethologische Marker, die im DLB kodiert, aber im EPM noch nicht untersucht wurden, Verhaltensdimensionen, die über herkömmliche Indizes hinausgingen. Insgesamt positionieren diese Ergebnisse die DLB als vielversprechenden Konzeptnachweis: Das Paradigma zeigt starke Augenscheinvalidität und teilweise Unterstützung in den Bereichen Inhalts-, Übereinstimmungs-, prädiktiver und Konstruktvalidität. Die Studie weist jedoch auch auf die Notwendigkeit weiterer Verfeinerung und Validierung hin, insbesondere in klinischen und pharmakologischen Kontexten.

Zusammenfassend zeigt diese Dissertation, dass VR-basierte Paradigmen zuverlässige und ökologisch valide Maße menschlichen Angstverhaltens liefern können und dabei sowohl sensitiv für stabile Persönlichkeitsmerkmale als auch für erfahrungsbedingte Faktoren sein können. Gleichzeitig verdeutlichen die Befunde, dass eine sorgfältige Berücksichtigung von Testprotokollen, Teilnehmermerkmalen und Validitätskriterien entscheidend für ihre breitere translationale und klinische Anwendung ist. VR-Paradigmen besitzen damit erhebliches Potenzial als methodische Brücke zwischen Tier- und Humanforschung zur Angst, bedürfen jedoch weiterer Verfeinerung, um ihre Robustheit und klinische Nützlichkeit zu sichern.

List of PhD-Related Publications

Publications for the PhD thesis

Biedermann, S. V., **Roth, L.**, Biedermann, D., & Fuss, J. (2024). Reliability of repeated exposure to the human elevated plus-maze in virtual reality: Behavioral, emotional, and autonomic responses. *Behavior Research Methods, 56*, 187–198. https://doi.org/10.3758/s13428-022-02046-5
–The author contributed to data collection, data preparation, and writing and editing of the manuscript.

Voulgaris, A., Biedermann, S. V., Biedermann, D., Bründl, S., **Roth, L.**, Wiessner, C., Briken, P., & Fuss, J. (2024). The impact of psychopathic traits on anxiety-related behaviors in a mixed reality environment. *Scientific Reports, 14,* 11832. https://doi.org/10.1038/s41598-024-62438-9
–The author contributed to data collection and data preparation, and reviewing and editing of the final manuscript.

Manuscripts submitted or in preparation for the PhD thesis

Roth, L., Biedermann, S. V., Wiessner, C., Biedermann, D., & Fuss, J. (submitted). The effect of gaming experience on behavioral responses to the human elevated plus-maze in virtual reality.
–The author contributed to the design of the study, conducted data collection and analysis, and wrote the manuscript.

Roth, L., Biedermann, S. V., Wolf, A. E. M., Briken, P., & Fuss, J. (in preparation). The dark-light box in mixed reality: assessment of a novel behavioral assay to study human anxiety-related behaviors.
–The author contributed to the design of the study, conducted data collection and analysis, and wrote the manuscript.

Statement on Use of Published Material

This dissertation is based on material also used for publication of the research findings, and for manuscripts prepared for publication. Many parts, including text passages, figures, tables, and statistical results, are therefore identical or nearly identical to the published versions. Overlaps in content are therefore inherent. To avoid repetition and improve readability, these works are not cited at each relevant passage but listed here to ensure complete transparency regarding the use of previously published material and soon-to-be-published manuscripts.

The correspondence between the dissertation and the (forthcoming) publications is as follows:

Study 1 is based on Biedermann et al. (2024)

Study 2 is based on Voulgaris et al. (2024)

Study 3 is based on Roth et al. (submitted)

Study 4 is based on Roth et al. (in preparation)

The respective introduction and discussion sections partly draw on conceptual and background passages from these works.

Chapter 1 Introduction

Fear and Anxiety: Emotional Responses to Threat

Emotions are fundamental to life in humans and a variety of non-human animal species (hereafter animals). Even though humans have long been thought to have the sole capacity for emotions, a growing body of animal research has demonstrated that many animals show a range of affective experiences from basic to more complex emotions (Bateson et al., 2011; Bekoff, 2000; Carranza-Pinedo et al., 2025; Hoehfurtner et al., 2025; Kremer et al., 2020; Panksepp, 2010; Solvi et al., 2016). Supporting the notion that emotions are products of evolution (Darwin, 1897). Especially the simplest of emotional responses are found in most vertebrate animal species as they are ancestral tools for survival (Panksepp, 2010). Anxiety and fear are two emotional responses that are highly adaptive in the face of various dangers that living beings encounter. Therefore, anxiety and fear as emotions represent responses to threats, with underlying structures that are evolutionarily conserved across species(LeDoux, 2012; Zych & Gogolla, 2021). For example, neural mechanisms like brain circuits of homologous brain regions (Panksepp, 2011), or action patterns of muscles (Ekman et al., 1980, 1983; Ekman & Friesen, 1986; McNaughton & Zangrossi, 2008). These emotional responses to threats are expressed through observable patterns of defensive behavior, which together form an animal's defensive repertoire. Animal literature has described at least six key defensive behaviors: risk assessment, freezing, avoidance, defensive threat, defensive attack, and flight (D. C. Blanchard & Blanchard, 2008). Defensive behaviors have been found to be similar throughout mammals (Edmunds, 1974; Shuhama et al., 2007), including humans, as assessed via self-reports (D. C. Blanchard et al., 2001; Perkins & Corr, 2006; Shuhama et al., 2008). The features of threats, for instance discreteness or ambiguity, defensive distance, and situational features like the presence of a conspecific or possible hiding places, modulate defensive behaviors, which allows a clear ethological differentiation between anxiety and fear (D. C. Blanchard, 2001).

While fear is a response to a direct and imminent threat, for example, a predator or a painful stimulus, anxiety is a response of a potential danger, like the odor of a predator or an ambiguous threat stimulus (R. J. Blanchard et al., 2008; McNaughton & Zangrossi, 2008). They therefore have distinguishable defensive directions. Fear enables animals to leave dangerous situations, such as fleeing from a predator (active avoidance). Whereas anxiety allows animals to either enter uncertain situations, as in approaching for cautious risk assessment, or to withhold from the situation (passive avoidance) (Gray & McNaughton, 2000; McNaughton &

Corr, 2004). However, anxiety is by no means less adaptive than fear, as ambiguous situations require careful assessment of risk, rather than a reactive fearful response. This allows non-threatening stimuli to be identified, reducing the time spent on defensive behavior and allowing individuals to return to other important activities such as foraging, reproduction, and self-care (R. J. Blanchard et al., 1990). At the same time, in case of a true threat, being vigilant (a form of risk assessment) helps the detection of danger (Eilam et al., 2011). However, the adaptive value of anxiety diminishes if the defensive behaviors become disproportionately intense for a given threat, displayed to an inappropriate stimulus, or persistent in the absence of actual danger. Given that defensive behaviors are time and labor intensive, the incorrect performance is at best energy-costly and at worst deadly, which makes them maladaptive (R. J. Blanchard et al., 2008).

In humans, anxiety typically also extends into elaborate cognitive processes such as rumination, anticipation, and worry. Human features of anxiety can be viewed as being very similar to the risk assessment responses observed in animals. This is because they both involve focused cognitive efforts to understand and solve difficult situations that are ambiguous or ill-defined (D. C. Blanchard & Blanchard, 2008; Eilam et al., 2011). Maladaptive anxiety responses in humans also involve unwanted reactions that are extreme, or occur to inappropriate stimuli. These responses underlie a range of clinical disorders classified as anxiety disorders, which significantly burden individuals' lives (Nutt et al., 2008). Therefore, understanding how such responses emerge, how they become dysregulated, and how this manifests, has been a central aim of anxiety and fear research. Analyzing defensive responses provides a window into the underlying mechanisms of fear and anxiety, offering insight into these essential questions. Because many defensive behaviors and their neurobiological underpinnings are evolutionarily conserved across species, animal models offer a valuable tool for investigating the origins, functions, and dysfunctions of anxiety or fear-related processes.

A major contribution of animal research has been the development of behavioral paradigms that simulate distinct types of threat. Paradigms involving approach-avoidance conflict are especially relevant for modeling anxiety because they involve situations of uncertainty, where potential danger is possible but not immediate. In these situations the individual must resolve the competing motivations of approaching a goal (e.g., novelty, food, or shelter) or avoiding harm. The resulting sustained behavioral and cognitive tension is a hallmark of anxiety (Gray & McNaughton, 2000; McNaughton & Corr, 2004). In humans, clinical pharmacology further shows that anxiolytic drugs such as Benzodiazepines are effective in treating disorders involving prolonged threat and uncertainty (e.g., generalized anxiety), but not specific phobias that reflect acute, distinct fear (McNaughton & Zangrossi, 2008). This well-demonstrated pharmacological distinction is mirrored in animal models of anxiety, where passive avoidance behaviors are reduced by anxiolytics, while active fear responses typically are not (R. J. Blanchard & Blanchard, 1989; Pellow et al., 1985). This convergence of behavioral and pharmacological evidence supports the ethological distinction between fear and anxiety in distinct underlying mechanisms and distinct defensive states. It has helped shape modern anxiety research, with animal models and their associated paradigms remaining a key component.

Animal Models of Fear and Anxiety

Animal models are generally designed to mimic specific aspects of human diseases or disorders. Using animal models can reduce time and costs of research and is argued to be especially useful for situations where human research becomes ethically difficult or impossible. Growing technical advances, e.g. gene engineering, allowed for the development of increasing experimental devices and methods to analyze human conditions in different model organisms with improved precision. These model organisms can include a range of animals like fish, canines, felines, or primates. However, rodents, particularly rats and mice are the most commonly used species (Crawley & Paylor, 1997; Litvin et al., 2008; Rosenthal & Brown, 2007). This is, in part, because rodents provide high experimental accessibility: they are small and easy to house and handle, have relatively short reproductive cycles, and are available in well characterized genetic strains (Nestler et al., 2002; Rosenthal & Brown, 2007; van der Staay, 2006). More importantly, despite evolutionary distance, rodents share key neuroanatomical and neurochemical systems with humans, including the amygdala, hippocampus, and agranular regions of the prefrontal cortex – regions known to regulate fear and anxiety (La-Vu et al., 2020; McNaughton & Corr, 2004; Panksepp, 2011; Preuss & Wise, 2022). This phylogenic continuity, which posits that similar behaviors observed across species may arise from homologous neural mechanisms, is an implicit assumption of all animal models (Cryan & Holmes, 2005; Kalueff et al., 2007; McNaughton & Zangrossi, 2008; van der Staay, 2006). To ensure the translational value of these models in psychiatric research, their validity in reflecting human conditions must be carefully evaluated.

A model's utility in translational research depends on its validity across multiple dimensions, including behavioral relevance, pharmacological responsiveness, and theoretical grounding (McNaughton & Zangrossi, 2008). Face validity refers to the extent to which the behavior observed in the animal resembles the human condition it's supposed to model, such as reduced activity or impaired sustained attention and memory in models of depression. Predictive validity is the proven capacity of a model to predict outcomes in the condition is meant to represent (McNaughton & Zangrossi, 2008; van der Staay, 2006). It is commonly established when the model responds to pharmacological treatments in a way that mirrors human clinical outcomes, for instance, when benzodiazepines reduce avoidance behavior in rodents, as they reduce anxiety symptoms in humans (Belzung & Griebel, 2001; Cryan & Holmes, 2005). Construct validity, arguably the most critical, reflects whether the model is based on theoretical mechanisms known to underlie the respective disorders, such as alterations in neurobiological systems or neural pathways (Nestler & Hyman, 2010; Willner, 1984). Therefore, high construct validity, ideally grounded in a sound theoretical framework supported by a broad evidence base, can still provide valuable insights into the human condition when predictive validity fails, by deepening the understanding of underlying mechanisms (McNaughton & Zangrossi, 2008). Beyond this classical triad, additional dimensions have also been emphasized such as content validity, referring to whether a model captures the full range of features relevant to the disorder (e.g. including behavioral, physiological, and cognitive aspects rather than focusing on a single symptom), and concurrent validity, which is demonstrated when model outcomes are consistent with other measures of the same construct, such as when avoidance behavior parallels stress hormone elevations or autonomic arousal (Biedermann et al., 2017). Although no single model can capture the full complexity of human disorders, these validity criteria are meant to ensure that the observed behaviors in animals are meaningfully related to human conditions and can be interpreted within a translational framework.

Experimental animal models of fear and anxiety are frameworks that incorporate specific behavioral paradigms to screen novel pharmaceuticals, study behavioral phenomena, and understand underlying etiology (Litvin et al., 2008). One important distinction in animal models of fear and anxiety lies in whether the paradigms elicit behavioral responses through conditioned or unconditioned procedures. Conditioned models typically rely on associative learning, as famously demonstrated by Pavlov's dog, where a neutral stimulus (e.g., a bell) was repeatedly paired with food until it alone triggered salivation (Jarius & Wildemann, 2015; Pavlov, 1927). In fear research, this principle is applied by pairing a neutral cue with an aversive event, such as a tone followed by a foot shock, to induce a learned fear response.

Such paradigms are used to study fear responses to discrete, predictable threats (Davis, 1992; Fanselow, 1980; LeDoux, 2000; Maren, 2001). In contrast, unconditioned models evoke innate defensive behaviors without prior learning. Unconditioned paradigms exist on a continuum of threat intensity, from low-threat situations involving novelty or open spaces, to social tests involving social interactions or isolation, and high-threat contexts such as predator exposure or defensive test batteries. Depending on the intensity and nature of the threat, these paradigms may elicit behaviors associated with either fear or anxiety (D. C. Blanchard, 2001; Davis et al., 2010; Litvin et al., 2008; Perusini & Fanselow, 2015) (Figure 1.1).

Unconditioned paradigms involving novelty and exploration have proven particularly useful for anxiety research. Behavioral measures from such paradigms are regarded as anxiety-related behaviors sensitive to anxiolytic pharmacological agents (e.g., benzodiazepines) (Lister, 1987; Pellow et al., 1985). Because these tests do not require pretraining of fear learning, they are considered to have the advantage of more efficient running and analyzing, which allows for high-throughput screening of behavioral traits or drug effects (Cryan & Holmes, 2005; Lister, 1987). Moreover, unconditioned paradigms involving novelty/exploration typically are setup to have a safe and an aversive area and therefore implicitly involve approach-avoidance conflicts, characteristic of anxiety-related decision-making (Gray & McNaughton, 2000; McNaughton & Corr, 2004). These conflict situations are not only ethologically valid but also offer a powerful framework to study core mechanisms in modeling anxiety (Calhoon & Tye, 2015; Gray & McNaughton, 2000).

The internal conflict between the motivation to move toward rewarding stimuli (e.g. explore novel environments) and the drive to avoid aversive ones (e.g. potentially dangerous) is a dynamic, referred to as approach-avoidance conflict. Approach-avoidance conflict has long been recognized as a fundamental mechanism underlying anxiety across species (Gray & McNaughton, 2000; McNaughton & Corr, 2004; Miller, 1944). One of the earliest theoretical frameworks describing approach-avoidance conflict was proposed by Miller (1944), who conceptualized behavior as the outcome of competing motivational forces. He suggested that when an individual is simultaneously drawn toward and repelled from a stimulus, two gradients develop: an approach gradient that increases gradually with proximity, and an avoidance gradient that increases more steeply. This means that as the subject nears the goal, the avoidance tendency can overpower the approach drive, creating a zone of conflict. This point of hesitation, where competing motivations are equally strong, reflects a core feature of anxiety. Miller noted that such conflict does not arise from approach tendencies alone, but emerges when avoidance is introduced, especially when fear prevents escape, and yet competing drives keep the subject engaged. This dynamic is foundational in understanding anxiety as a state of unresolved tension between opposing motivational forces. Later research expanded on this framework, demonstrating that such conflicts can be systematically elicited and studied under controlled conditions (Aupperle & Paulus, 2010; Kalueff et al., 2007).

To systematically investigate approach-avoidance conflict and its underlying mechanisms, researchers have developed a variety of experimental paradigms, particularly in animal models. Among the most widely used unconditioned paradigms involving approach-avoidance conflict are the Elevated Plus Maze (EPM) (Handley & Mithani, 1984; Montgomery, 1955; Montgomery & Monkman, 1955) and the Dark-Light Box (DLB) (Crawley, 1981, 1985) (Figure 1.1). These paradigms capitalize on rodents' innate aversion to open or brightly lit spaces while leveraging their exploratory drive. Both paradigms place rodents in environments that simultaneously offer safe and aversive areas, prompting the animals to navigate conflicting motivations: exploring a novel, potentially dangerous space versus retreating to a familiar, secure one. In the EPM, for example, animals must choose between the safety of enclosed arms and the risky exploration of open arms. Similarly, the DLB tests an animal's willingness to leave a dark, safe enclosure and explore a brightly lit, aversive space. This generates a measurable behavioral tension, which reflects the core features of anxiety, where potential threat must be evaluated without the presence of immediate danger. These paradigms have proven effective in capturing anxiety-like behavior in form of risk assessment, hesitation, and passive avoidance. As mentioned before, these behaviors have been shown to be sensitive to anxiolytic compounds, such as benzodiazepines (Crawley, 1981; Lister, 1987; Pellow et al., 1985) and are distinctive of anxiety-related processes (Cryan & Holmes, 2005; Lister, 1987). Their ethological validity, ease of implementation, and interpretability have made them fundamental tools in preclinical anxiety research, allowing researchers to model the cognitive-motivational dynamics of anxiety.

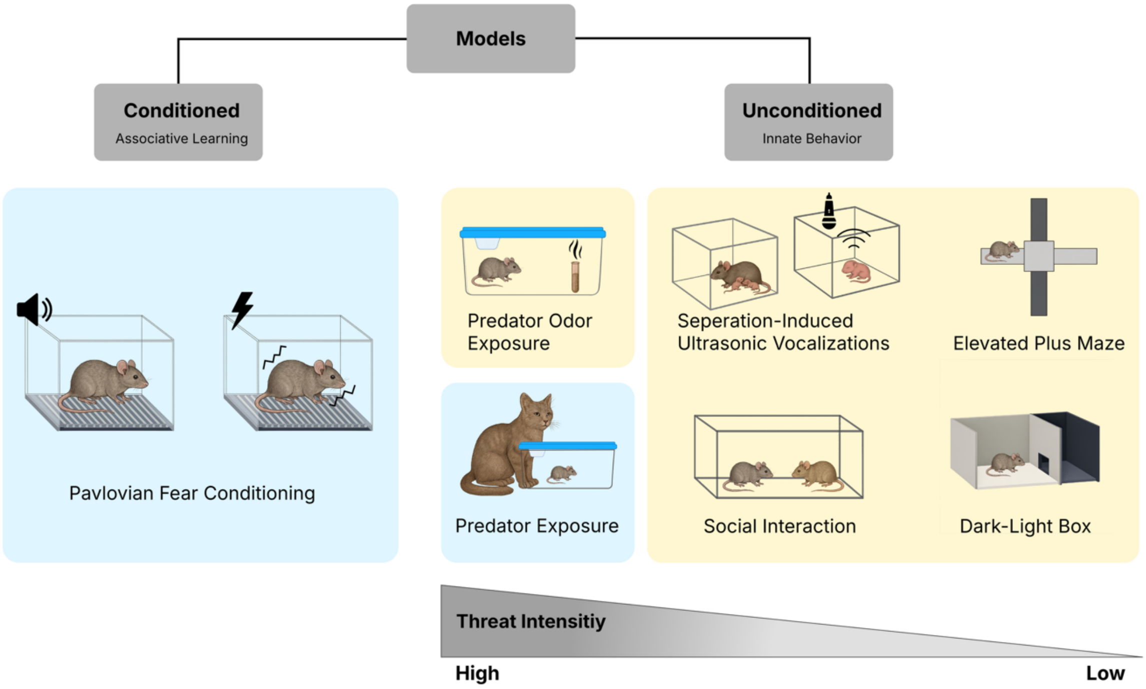


Figure 1.1: Examples of conditioned (left) and unconditioned (right) animal models of fear (blue) and anxiety (yellow) and their corresponding paradigms. Illustration and Copyright by Kari Massey, 2025. Created for this dissertation.

Bridging the Gap Between Animal Models and Human Anxiety Research Through Virtual Reality

Despite their value in isolating fundamental anxiety-related mechanisms, animal models cannot capture the full complexity of human psychopathology (Nestler & Hyman, 2010). While rodents and humans share evolutionarily conserved neural circuits situated in homologous brain regions, important structural and functional differences constrain direct translational generalizability. For example, rodents possess homologs of the agranular medial and orbital frontal cortex but lack homologs of the granular prefrontal cortex (Preuss & Wise, 2022). These granular areas, are considered a primate specialization, forming a cortical organization of interconnected higher-order association areas not present in non-primate mammals. Moreover, the granular prefrontal cortex is thought to be a region that is particularly expanded in humans and forms a substantial portion of the association cortex within the frontal lobe

(Preuss & Wise, 2022). Beyond neuroanatomy, defensive behavior of common laboratory animals has been documented to differ from wild conspecifics. Over generations, breeding and standardized housing (of especially rodents and primates) resulted in an artificial rift, a phenomenon referred to as "laboratorization" (Huck & Price, 1975; Litvin et al., 2008; Price, 1999; Würbel, 2002). For example, rats raised in laboratory settings were found to show more defensive behaviors (escape and confrontational defensive postures) compared to their wild counterparts (R. J. Blanchard et al., 1986). Similarly, EPM and open field testing of wild mice showed enhanced avoidance of open areas, which is attributed to higher levels of anxiety, compared to laboratory strains (Augustsson & Meyerson, 2004). However, the opposite has also been reported, with wild mice showing less anxiety-related behavior than laboratory mice (Holmes et al., 2000). The effects of laboratorization can undermine the validity of interpreted behavioral readouts in animal models, thus limiting their translational relevance to humans. Translational efforts are also challenged by differences in data sources and interpretative frameworks across species. While animal studies often rely on behavioral observations following pharmacological or lesion-based interventions, human research typically emphasizes subjective symptom reports or neuroimaging correlates. As Kirlic and colleagues (2017) pointed out, this can lead to a mismatch akin to "comparing apples to oranges." The incorporation of introspective data such as verbal descriptions and subjective ratings is crucial for understanding the lived experience of human anxiety. However, self-reports do present their own limitations.

While self-report remains a widely used method in clinical and experimental anxiety research, its reliability as a stand-alone measure is limited. It is known that subjective ratings often fail to correlate with actual behavior (Zinbarg, 1998). Moreover, it has been argued that avoidance and other defensive behaviors can be generated by non-conscious neural circuits that do not rely on subjective fear, indicating a possible disconnection between behavior and conscious experience (LeDoux, 2012). In anxiety-inducing contexts, individuals may exhibit anxiety-related behavior without consciously recognizing the fear stimulus or reporting increased anxiety (Bertini et al., 2013; Mineka & Öhman, 2002; Tamietto & De Gelder, 2010). This discrepancy suggests that subjective awareness does not always reflect the underlying affective state. Consequently, observable behaviors, such as movement patterns, postures, or avoidance tendencies, could add an objective and standardized assessment of anxiety-related processes. First successful translation of well-established a rodent paradigm into human research was done by Walz and colleagues (2016), by successfully implementing a human open field test on a soccer field, and reproducing core behavioral patterns observed in rodents. The open field test is traditionally used in rodents to index anxiety by contrasting exploration of open (aversive) versus peripheral (safe) zones. In the soccer field setting, human participants also avoided the center and preferred movement along the perimeter of the field, a behavior known in animal research as thigmotaxis (Walz et al., 2016).

However, while promising, such real-world setups pose significant logistical and methodological challenges, including limited standardizability, environmental confounds, and high resource demands. Virtual reality (VR) overcomes many of these limitations by enabling controlled, replicable environments that do not require extensive physical space (Felnhofer et al., 2015; Gromer et al., 2021; Schöne et al., 2023; Schöne, Sylvester, et al., 2021). By preserving the core task mechanics of established behavioral paradigms in virtual environments, VR enables standardized and high-throughput testing that would be difficult to achieve in real-world setups (Kisker, Lange, et al., 2021). Moreover, the high level of immersion achievable in VR enhances emotional engagement and facilitates the elicitation of lifelike behavioral responses (Baños et al., 2004; Schöne et al., 2023; Schöne, Sylvester, et al., 2021). Crucially, VR allows for the simultaneous collection of behavioral, physiological, and subjective data within standardized, ecologically valid environments (Biedermann et al., 2017; Kisker, Lange, et al., 2021). Compared to conventional laboratory tasks, which commonly reduce anxiety-related behavior to keystrokes or eye and joystick movements (e.g.: Aupperle et al., 2011; Schlund et al., 2016; for review see: D. C. Blanchard, 2017), this multimodal assessment enables researchers to explore the complex interplay between what participants report, what they feel, and how they behave, offering a richer, more dimensional understanding of human anxiety.

VR offers a promising medium for the backward translation of rodent anxiety paradigms into controlled, engaging human environments. Paradigms such as the EPM rely on threat features, like open or elevated spaces, that evoke evolutionarily conserved and therefore innate defensive responses across many mammalian species. These threat elements have been successfully adapted for human participants. For instance, Biedermann and colleagues (2017) translated the spatial structure and approach-avoidance conflict of the rodent EPM into a virtual maze. The task reproduced key behavioral patterns in humans, as participants avoided the open arms suspended over a virtual cliff and reported higher anxiety on them, supporting the cross-species validity of this behavior. Other VR environments have further broadened this approach by simulating additional ecologically meaningful contexts: dimly lit caves (Kisker, Lange, et al., 2021), dark village squares and alleys (Toet et al., 2009), and social situations (Lange & Pauli, 2019; Wieser et al., 2010). Across these studies, VR has proven effective in eliciting lifelike affective responses and measurable avoidance tendencies, supporting its utility for studying approach-avoidance conflicts in humans. Despite its methodological promise, VR-based research has yet to address potential limitations, such as temporal stability of responses, the influence of interindividual differences, and to explore its applicability in the study of human psychopathology.

Challenges and Opportunities in Applying VR-Based Anxiety Assessment

A central requirement for applying VR-based paradigms to clinical anxiety research lies in their ability to reliably track anxiety-related behavior over time, as stable test-retest profiles are necessary for longitudinal or repeated-measures designs. While virtual environments elicit lifelike behavioral and physiological responses (Bernardo et al., 2021; Felnhofer et al., 2015; Kisker, Gruber, et al., 2021; Kisker, Lange, et al., 2021; Schöne et al., 2023; Schöne, Kisker, et al., 2021), repeated exposure can lead to non-associative learning effects such as habituation or sensitization. These effects may compromise the temporal stability of VR-based measures. Habituation is defined as a "behavioral response decrement" that occurs with repeated exposure and is not attributable to sensory or motor fatigue (Rankin et al., 2009). In contrast, sensitization refers to a temporary increase in responsiveness before adaptation sets in (Groves & Thompson, 1970). These processes can affect behavioral, physiological, and subjective responses. While such effects are well documented in rodent models of anxiety (Carobrez & Bertoglio, 2005; Espejo, 1997; Holmes & Rodgers, 1998; Rodgers & Shepherd, 1993; Schneider et al., 2011; Treit et al., 1993), little is known about their role in VR-based human testing. The therapeutic effects of VR exposure therapy are based on habituation processes; therefore, therapeutic approaches rely on longer or more frequent exposures to allow habituation to occur (e.g. Gujjar et al., 2019). However, when using VR to assess approach-avoidance behavior in experimental paradigms, the occurrence and impact of habituation remain largely unknown. This represents a critical gap, as the use of within-subject designs in clinical trials depends on the temporal consistency of repeated assessments. Establishing test-retest reliability in VR is therefore essential to support its broader implementation in translational and clinical anxiety research.

Interindividual differences in anxiety-related behavior within virtual environments are shaped by both biological and experiential factors. A prominent example concerns sex differences, which have been shown to manifest inconsistently across species. In rodent models, female animals typically exhibit lower levels of anxiety-like behavior than males in paradigms such as the EPM and open field test (Domonkos et al., 2017; Imhof et al., 1993; Johnston & File, 1991; P. Knight et al., 2021; Ramos et al., 2002; Scholl et al., 2019). In contrast, validation studies of the virtual EPM in humans have demonstrated the opposite pattern, with women showing stronger avoidance behavior and higher anxiety indicators than men (Biedermann et al., 2017; Nouri et al., 2022). In human research, such variation in human anxiety behavior cannot be fully explained by biological factors alone. While hormonal and neuroanatomical differences likely contribute (Bangasser & Cuarenta, 2021; Donner & Lowry, 2013; McEwen, 2020), social factors must also be considered. One possible influence is video game experience, which remains more common among adolescent boys in many cultural contexts (Brunborg et al., 2013; Leonhardt & Overå, 2021; Tak & Catsambis, 2023). Prior gaming experience has been shown to affect cognitive and perceptual domains relevant to VR, including brain function, attentional capacity, and visuospatial skills (Palaus et al., 2017). It has been associated with enhanced attentional control (Bavelier, Achtman, et al., 2012; Bavelier, Green, et al., 2012; Dye et al., 2009; Green & Bavelier, 2003, 2006, 2007, 2012; Hubert-Wallander et al., 2011), improved visuomotor coordination (Palaus et al., 2017), and, in some cases, reduced susceptibility to cybersickness (M. M. Knight & Arns, 2006; Weech et al., 2020 but see Gamito et al., 2008, 2010). Gaming may also increase perceived presence, as was shown in a first-person shooter VR scenario (Gamito et al., 2010), though findings remain mixed (Alsina-Jurnet & Gutiérrez-Maldonado, 2010; Ling et al., 2013; Weech et al., 2020). Despite this relevance, its role in shaping anxiety-related behavior in VR remains underexplored. Some evidence suggests that participants with limited computer experience exhibit stronger physiological arousal during VR anxiety tasks (Gamito et al., 2008), and that experienced gamers may perceive fear and surprise less intensely in virtual fear contexts (Geslin et al., 2011). These findings highlight the importance of accounting for experiential factors when interpreting behavior.

In addition to biological and experience-based factors, interindividual differences in anxiety-related behavior may also be shaped by stable psychological traits. One example of particular relevance is psychopathy, a multidimensional personality construct with distinct affective-interpersonal and antisocial-lifestyle components (Hare & Neumann, 2010). Traditionally, psychopathy has been linked to diminished anxiety and fear responses (Cleckley, 1988; Fowles, 1980; Lykken, 1957, 1995). Early experimental research emphasized low physiological reactivity to aversive stimuli, leading to the “low-fear hypothesis” of psychopathy (Hare, 1978; Hare et al., 1978; Lykken, 1957). While some empirical studies support this view (Lilienfeld & Andrews, 1996; Neumann et al., 2013; Patrick et al., 1993), others challenge it

(Salekin et al., 2005; Schmitt & Newman, 1999). This debate reflects conceptual and methodological inconsistencies in how fear and anxiety are defined and measured in the literature on psychopathic research (N. E. Anderson et al., 2021; Derefinko, 2015). A meta-analysis cites findings showing that total psychopathy scores relate only weakly with anxiety and fear, whereas more differentiated patterns emerge at the subfactor level: Factor 1 traits, reflecting affective-interpersonal features such as superficial charm, grandiose sense of self-worth, lack of guilt, and callousness, was negatively correlated with anxiety and fear, while Factor 2 traits, linked to antisocial-lifestyle features such as need of stimulation, impulsivity, parasitic lifestyle, and irresponsibility, showed positive correlations with anxiety (Derefinko, 2015). Furthermore, individuals scoring high on affective-interpersonal items (Factor1) also had anomalous behavioral and physiological response patterns associated with low fearlessness and anxiety (Patrick, 1994; Thomson et al., 2019; Vaidyanathan et al., 2011). Neurobiological research supports these findings, highlighting abnormalities in threat processing pathways, including altered amygdala reactivity and impaired limbic-prefrontal circuitry (N. E. Anderson & Stanford, 2012; Birbaumer et al., 2005; Esteller et al., 2016; Pastor et al., 2003; Schultz et al., 2016). However, current research suggests that psychopathy does not reflect a general absence of anxiety or fear, but rather involves specific cognitive deficits in processing emotionally relevant information (N. E. Anderson et al., 2021; N. E. Anderson & Kiehl, 2012; Budhani et al., 2006; Larson et al., 2013) Yet, there are limited studies with a focus on anxiety- and threat-related behavior associated with psychopathy. In addition, the commonly used methods to assess anxiety or threat-related responses, such as self-report, autonomic responses, or the ability to identify fear in others, differ considerably (Centifanti et al., 2022; Derefinko, 2015; Hoppenbrouwers et al., 2016). Understanding how such trait-based impairments manifest in behavioral anxiety measures may advance both our knowledge of psychopathy and the development of VR-based assessment tools for clinical research across psychiatric populations.

While most human VR adaptations in anxiety research have focused on the EPM, relying on a single paradigm may constrain the scope of observable anxiety-related behaviors. In animal research, anxiety is widely understood as a multidimensional construct (Gray, 1982; Ramos & Mormède, 1997), and behavioral paradigms are often combined to capture complementary aspects of defensive responses (Canteras & Blanchard, 2008; Ramos, 2008; Ramos & Mormède, 1997). Cryan and Holmes (2005) argue that no single test offers a comprehensive model of anxiety, and that even presumably similar exploration-based tasks differ qualitatively in what they assess. Different tasks elicit exploration or avoidance in response to distinct aversive stimuli, such as open elevated spaces in the EPM or bright illumination in the DLB. They are therefore believed to index overlapping but partially distinct dimensions of anxiety-like behavior. By combining multiple tests, the risk of attributing findings to task-specific sensitivities (e.g., fear of heights) is reduced, allowing for stronger inferences about general anxiety tendencies (Ramos, 2008). This principle has clear relevance for translational models as well, particularly in VR research, where much of the existing human data are derived from the virtual EPM – a task in which anxiety-like behavior was shown to be shaped not only by anxiety, but also by specific traits such as acrophobia and sensation seeking (Biedermann et al., 2017; Madeira et al., 2021). Introducing complementary tasks that rely on different aversive dimensions are necessary as they extend the construct coverage of VR-based anxiety assessments and support a more differentiated understanding of human anxiety.

Aims of the Dissertation

This dissertation addresses key methodological and translational gaps identified in previous sections by systematically advancing the use of VR adaptations of classical animal paradigms for the assessment of human anxiety-like behavior. Building on prior work by Biedermann and colleagues (2017) that translated and validated the EPM for human use in VR, this research extends the paradigm's application by investigating its reliability for repeated measures and exploring individual factors that influence anxiety-related behavior. In addition, the dissertation introduces and investigates the validity of a VR-based DLB as a complementary approach-avoidance paradigm. Together, these studies aim to strengthen the methodological foundation for using VR-based anxiety assessments in translational and clinical research, ensuring their robustness for longitudinal designs and broader behavioral phenotyping.

To address methodological challenges regarding the test-retest reliability of VR-based anxiety measures and potential habituation or sensitization during repeated testing, **Study 1** investigated whether the human virtual EPM, previously validated for single-session assessments (Biedermann et al., 2017), provides stable test-retest measures of anxiety-related behavior in repeated applications. A central aim was to evaluate whether applying procedures known from rodent testing, specifically a 28-day inter-trial interval and change of testing room (Schneider et al., 2011), would help avoid effects such as habituation or sensitization in humans. In the case of the virtual EPM, a change of testing room would correspond to a change of the virtual test surroundings. An additional exploratory aim was to test whether environmental realism is necessary to evoke anxiety-related behavior by comparing naturalistic realism with a more stylized, game-like setting. By addressing these questions, this study sought to establish the temporal reliability of the virtual EPM and identify design factors that support its application in within-subject and longitudinal research contexts, advancing its use in translational and clinical anxiety studies.

Research on psychopathy has long suggested alterations in fear and anxiety processing (Hare, 1978; Hare et al., 1978; Lykken, 1957), yet findings have been inconsistent. Moreover, while the assessment methods of anxiety responses in most studies vary, only few studies have examined direct anxiety- and threat-related behavior in relation to psychopathy, leaving uncertainty about how these traits manifest in observable defensive behavior. **Study 2** addressed this gap by assessing whether psychopathic personality traits are linked to distinct approach-avoidance behaviors within the EPM. The study examined whether higher psychopathic trait scores, particularly on the Fearlessness dimension, predict reduced avoidance and increased exploratory behavior, reflected in measures such as time spent and entries into open arms. This approach extends existing research by introducing a translational behavioral measure of anxiety responses that mirrors established animal paradigms and allows examination of individual differences related to psychopathy. Through this investigation, the study aimed to improve understanding of how psychopathic traits influence anxiety-related responding in humans and to evaluate the potential of immersive VR paradigms for advancing research in experimental and, prospectively, clinical contexts.

Sex differences in anxiety-related behavior have been consistently observed in animal research(Domonkos et al., 2017; Imhof et al., 1993; Johnston & File, 1991; P. Knight et al., 2021; Ramos et al., 2002; Scholl et al., 2019). In humans, corresponding gender differences have also been reported. However, findings in VR settings show a reversed pattern compared to rodent models, with women displaying stronger avoidance behavior than men (Biedermann et al., 2017; Nouri et al., 2022). **Study 3** examined whether differences in gaming experience contribute to these differences in behavioral patterns between men and women in the virtual EPM. Given that gaming can enhance cognitive skills like improved visuospatial abilities and attentional capacities, which are beneficial in immersive VR settings (Palaus et al., 2017), it was hypothesized that gaming experience reduces anxiety-related behavior in the virtual EPM in participants with more gaming experience. As gaming is more prevalent among men than women (Brunborg et al., 2013; Leonhardt & Overå, 2021; Tak & Catsambis, 2023), the difference in gaming experience could also explain observed differences in anxiety-related behavior of men and women in the virtual EPM. The study analyzed behavioral measures such as time spent and entries into open arms, latency to first open-arm entry, and end exploration, investigating whether these are influenced by gender and prior gaming experience. By addressing this question, the study aimed to clarify a potential experiential factor underlying behavioral differences in VR-based anxiety assessments and to inform future experimental designs by highlighting gaming experience as a variable that may influence behavioral outcomes.

Current VR-based anxiety research has predominantly relied on the Elevated Plus Maze (EPM), which captures only a limited dimension of the construct fear and is partly influenced by fear of heights (Biedermann et al., 2017; Madeira et al., 2021). This constrains the construct coverage of translational anxiety assessments. To address this, **Study 4** introduced and explores the validity of a virtual DLB as a complementary paradigm to the virtual EPM. The study aimed to capture a distinct dimension of unconditioned anxiety by eliciting approach-avoidance behavior through contrasting dark (aversive) and bright (safe) areas. Validity was examined by assessing avoidance and exploratory behavior, physiological stress responses (skin conductance, heart rate, respiratory rate), and salivary biomarkers (cortisol, alpha-amylase). Associations with self-reported anxiety and personality traits were examined. In addition, the paradigm was applied to a clinical sample of psychiatric patients with anxiety disorders (AD), borderline personality disorder (BPD), and Post-traumatic-stress-disorder (PTSD), compared with matched healthy controls. By extending behavioral, physiological, and endocrinological assessments to a complementary paradigm that captures a distinct dimension of unconditioned anxiety, this study addresses a key limitation of current VR-based anxiety research and establishes a foundation for broader and more robust phenotyping of human defensive responding in both experimental and clinical contexts. Collectively, these studies aim to advance VR-based anxiety assessment by contributing to the evaluation of its reliability, capturing trait- and experience-based individual differences, and supporting broader construct coverage for translational and clinical research.

Chapter 2 Methods

General Methods

Setup and Virtual Environments

Elevated Plus Maze

The human virtual EPM used in all studies was developed and validated by Biedermann et al. (2017) as a mixed reality translation of the classical rodent assay. The apparatus consisted of a real physical wooden maze with four arms (30 cm wide, 20 cm high, 175 cm long) arranged in a plus shape (total 350 × 350 cm), installed in a laboratory space (550 × 550 cm). A virtual reality plus-maze, with the same shape, material and size as well as position as the existing physical maze, was presented in a virtual environment via head-mounted display (HTC Vive, Seattle, USA) and noise-canceling headphones (Bose QuietComfort 35, Framingham, USA). Spatial tracking was performed with two virtual reality tracking systems (HTC Vive Base Station, Seattle, USA) attached at 250 cm height at opposite walls in the laboratory room. To increase presence in the virtual environment, fans were used to simulate wind during experimental testing. Before each testing, the virtual and physical mazes were synchronized using a controller (HTC Vive controller, Seattle, USA), a purpose-designed mounting device, and a purpose-designed software (A+ cross), ensuring precise alignment of physical and virtual coordinates. The calibration software (A+ cross) aligned the axis of the 3D world space coordinate system with the arms of the physical maze. The world space coordinate system was created with its origin at the center of the maze, its x-axis aligned with the closed arms, and its z-axis aligned with the open arms of the plus-maze. Headset position and orientation during the experiment were sampled at 5 Hz for each sampling point i at time t to obtain a set of 3D positions $\vec{p}^i = (p_x^i, p_y^i, p_z^i)$ and 3D orientations $\vec{r}^i = (r_x^i, r_y^i, r_z^i)$. These measurements were then analyzed by the software (A+ cross) to evaluate participants' movement patterns on the maze. A sample point pi counts towards time spent on one arm if the absolute position on one of the axes, and thus distance from center, is larger than a pre-defined threshold value. The total time spent on one of the open arms t_o and the closed arms, t_c, is calculated as $t_{o,c} = \sum t^i [|p_{x,z}^i| > threshold]$. Sample points for which the inverse case, $p_{x,z}^i > threshold$ and $p_{x,z}^{i-1} < threshold$, was measured were counted towards the total number of entries $n_{[o,c]}$ to one of the arms.

Participants entered the room with their eyes closed and were guided by the experimenter towards the maze where they received the head-mounted display and headphones. They were then instructed to open their eyes, and after checking the vision in a baseline graphical environment, the virtual reality software was started (A+ cross, VirtualRealWorlds.com, Germany). A recorded voice instructed participants, who found themselves in a large virtual laboratory room, to step on the maze in front of them and walk towards the center of the maze. Here, they had to wait for 60 seconds to allow for baseline measurements. Participants received the instructions that they would be allowed to explore the environment on the maze once the scene had changed. The behavioral experiment started after 90 seconds with said scenery change from a large virtual laboratory room into a naturalistic scene. This new virtual environment consisted of the plus-maze placed on a virtual rocky cliff surrounded by water. Two opposite arms (closed arms) and the center of the maze were surrounded by rocks, while the other arms reached out over the water, which was roughly 55 m below (open arms) (Figure 2.1). Simultaneously with the change of the virtual environment, suitable ocean sounds and the two ventilators were started. The experimental testing lasted 300 seconds in which participants were allowed to explore the maze. For further detailed description and visualization please see Biedermann et al. (2017) and a video available under this link:

https://osf.io/9tphn/?view_only=f0b4ec62c5204122bacaf1bbae3d46d1.

Figure 2.1: Mixed-reality elevated plus maze (EPM) setup. Left panels show a participant walking on the physical cross-shaped apparatus, while the right panel depicts the corresponding virtual reality environment displayed in the head-mounted display.

Dark-Light Box

The virtual DLB consists of an experimental room (760 cm × 490 cm), which has been recreated in virtual reality at a 1:1 scale. Both the virtual and real room are divided into a bright (safe) area (7.82 m²) and a dark (aversive) area (28.44 m²). The test area was equipped with two virtual reality tracking systems (HTC Vive Base Station, Seattle, USA) mounted at a height of 250 cm on opposite walls. By tracking the head-mounted display, localization data was continuously recorded, allowing for the extraction of key behavioral parameters. Specifically, the system logged position data, timestamps, and start and end times, which were stored in a JSON file upon completion of the session. Each localization point was recorded along with a corresponding timestamp, enabling a detailed reconstruction of participant movements over time. For subsequent analysis, a custom-built analysis tool, developed within Unity, processed the JSON files. This tool parsed the recorded data and computed behavioral parameters, such as latencies, time spent in each area, velocity and distances covered within the DLB. Participants wore a wireless head-mounted display (HTC Vive Pro 2018, Seattle, USA) and noise-canceling headphones (Bose QuietComfort 35, Framingham, USA), and after the virtual reality software (Unity Version 2019.3.1f1) was launched, they initially found themselves in a virtual staging room. A recorded voice instructed them to stand still and wait for 60 seconds to allow for baseline psychophysiological measurements. They were further informed that they would be allowed to explore the environment once the scene changed. As soon as the scene transitioned, the behavioral testing began and lasted 300 seconds. Participants started in the bright area, facing the dark section, which gradually became darker as one moved towards the far end of the room. A small light source on the back wall of the dark area served as an incentive for exploration (Figure 2.2). If a participant stepped out of the bright area, the lighting in that area dimmed; upon returning, it brightened again. To reinforce the perception of safety and aversion, different auditory stimuli were used. The bright area featured birdsong, creating a calm and familiar atmosphere associated with safety. In contrast, a low-frequency hum played in the dark area, increasing in volume and becoming increasingly unpleasant as participants approached the sound sources located in the darkest corners. Additionally, a ventilator was activated at the start of the experiment to heighten the sense of presence. Positioned at the far end of the room, the airflow was only perceptible in the darkest areas. For a video demonstration of the setup, see: https://youtu.be/hASvpVahsC8.

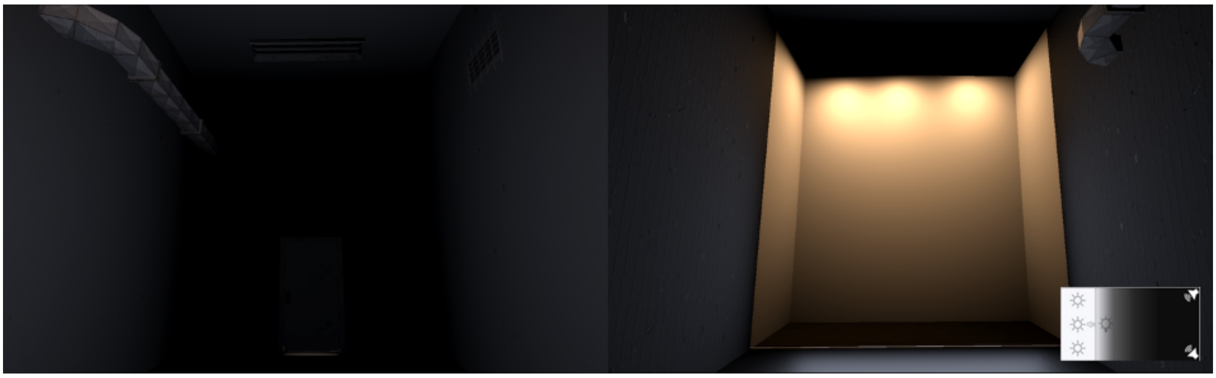


Figure 2.2: Virtual reality Dark-Light Box (DLB) paradigm. The left panel depicts the dark area (lightened here for illustration purposes), and the right panel shows the bright area of the environment. The inset on the bottom right illustrates the brightness and sound sources of the setup from an aerial perspective.

Questionnaires

Sensation Seeking Scale (SSSV)

The Sensation Seeking Scale Form V (SSSV; Zuckerman et al., 1978) was used to assess individual differences in the tendency to seek novel and stimulating experiences. The SSSV consists of 40 forced-choice items and assesses a general measure of sensation seeking as well as four subscales with 10 items each: (1) Thrill and Adventure Seeking (TAS; e.g., parachute jumping), (2) Experience Seeking (ES; e.g., exploring strange cities alone), (3) Disinhibition (DIS; e.g., desiring varied sexual experiences), and (4) Boredom Susceptibility (BS; e.g., preference for unpredictable friends). Internal consistency in the original validation ranged from $\alpha = 0.56$-0.82 across subscales, with total scores showing good reliability ($\alpha = 0.83$-0.86). The factorial structure was replicated across cultures and sexes (Zuckerman et al., 1978). Construct validity was supported by positive correlations with related personality traits, autonomy, field independence, substance use, and openness to experience (Zuckerman et al., 1972; Zuckerman & Link, 1968).

State-Trait Anxiety Inventory (STAI)

The State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970) was used to measure both situational and dispositional aspects of anxiety. The questionnaire comprises 40 self-report items with two subscales of 20 items each: the State Anxiety Scale (S-Anxiety; e.g., "I am tense"), which assesses anxiety "right now," and the Trait Anxiety Scale (T-Anxiety; e.g., "I worry too much over something that really doesn't matter"), which evaluates general anxiety proneness. Items are rated on a 4-point scale, with response options ranging from "not at all" to "very much so" for S-Anxiety and "almost never" to "almost always" for T-Anxiety. Scoring is reversed for 19 anxiety-absent items. Internal consistency coefficients were high (α = 0.86-0.95), while test-retest reliability ranged from 0.31-0.86, reflecting lower stability for the State scale due to its transient nature (Spielberger et al., 1970). Construct validity was supported by high correlations with other anxiety measures such as the Taylor Manifest Anxiety Scale (r = 0.73) and the Cattell and Scheier Anxiety Scale Questionnaire (r = 0.85) (Spielberger, 1970).

Simulator Sickness Questionnaire (SSQ)

The Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993) was applied to assess side effects of simulator and virtual reality exposure. The SSQ consists of 16 symptoms (e.g., nausea, eyestrain, dizziness), each rated on a 4-point Likert scale ranging from 0 "none" to 3 "severe". The three subscales (Nausea, Oculomotor, and Disorientation) can be combined into a weighted total score. Internal consistency was good, with Cronbach's α ranging from 0.84-0.95 across subscales and total score (Bouchard et al., 2007; Sevinc & Berkman, 2020). The SSQ demonstrated criterion validity by reliably distinguishing participants who discontinued exposure due to simulator sickness from completers (Balk et al., 2013) and showed sensitivity to VR stimuli of varying intensity (Sevinc & Berkman, 2020).

iGroup Presence Questionnaire (IPQ)

The iGroup Presence Questionnaire (IPQ; Schubert et al., 1999) was used to assess subjective presence in virtual environments. The IPQ consists of items rated on a 7-point Likert scale ranging from -3 "fully disagree" to +3 "fully agree" and comprises three subscales: (1) Spatial Presence (SP; e.g., "In the computer generated world I had a sense of being there"), (2) Involvement (INV; e.g., "I was completely captivated by the virtual world"), and (3) Realness (REAL; e.g., "How much did your experience in the virtual environment seem consistent with your real world experiences?"). A single general item ("Overall I had the impression of being in the virtual environment") is also included. Scores are calculated as mean values for the three subscales. Internal consistency was acceptable, with α = 0.68-0.87 across subscales (igroup Project Consortium, n.d.; Regenbrecht & Schubert, 2002). Confirmatory factor analyses supported the three-component structure across independent samples, and construct validity was evidenced by significant associations with immersion-related variables (Schubert et al., 1999, 2001).

Gender Reporting

All participants were asked to indicate their gender identity using the German phrasing “Welchem Geschlecht fühlen Sie sich zugehörig?” (Which gender do you identify with?), with response options “männlich,” “weiblich,” or “andere” (literal translation: male, female, other). In the published studies (Study 1 and 2) the literal translation “male/female” was used when reporting participant characteristics. However, because the item reflects a self-reported gender identity, the response options align with the gendered English translation “man”, “women”, “other”. Only participants identifying as male or female were included in the binary analyses of the studies. Therefore, in this dissertation all participants are referred to as women and men for inclusivity and sensitivity, in accordance with the APA guidelines on Bias-Free Language (*Publication Manual of the American Psychological Association (7th Ed.)*., 2020; Section 5.5, Gender and Noun Usage)

Physiological and Endocrinological Measurements

Psychophysiological and endocrine measures were collected only in Study 1 and Study 4. These included skin conductance level (SCL), respiratory rate (RR), and heart rate (electrocardiogram, ECG), recorded using BioNomadix wireless physiology devices connected to a BIOPAC MP150 data acquisition system (Biopac Systems, Goleta, CA, USA). Data were analyzed using AcqKnowledge software (version 4.4.1), with Ag/AgCl electrodes attached to the palmar surface of the non-dominant hand approximately 10 minutes before testing. Baseline levels were recorded 30 seconds before behavioral testing, and average levels were compared to average levels of 30-second intervals during behavioral testing, leading to 11 intervals (baseline, 0-30 s, 30-60 s, etc.). SCL data were log-transformed prior to analysis to approximate normality.

Saliva samples were collected at three time points: before testing (T0), immediately after testing (T1), and 15 minutes post-testing (T2). Samples were centrifuged and stored at -80 °C until analysis. In Study 1, only alpha-amylase activity was measured using a commercial liquid-phase enzymatic assay (RE80111, IBL International, Hamburg, Germany). In Study 4, both alpha-amylase and cortisol were analyzed, with alpha-amylase measured as in Study 1 and cortisol determined using an enzyme-linked immunosorbent assay (Cortisol Saliva ELISA, Tecan/IBL International, RE52611). Statistical analyses were conducted using IBM SPSS Statistics (Version 29) and included repeated-measures ANOVA, with relevant post hoc tests applied only in Study 1. Details of study-specific models, factors, and correction methods are provided in the individual study sections.

Heatmaps

Participant movement data on the EPM were visualized as heatmaps to illustrate spatial patterns of behavior. Position data from individual participants were combined and discretized into 10 × 10 cm spatial patches, resulting in 35 × 35 matrices that mapped the number of visits per maze location. Coordinates were transformed into integer grid indices and clipped to ensure that all points fell within the experimental boundaries. For every recorded coordinate $\binom{x}{y}$, the corresponding matrix entry was incremented according to:

$$M_{\{i,j\}} = M_{\{i,j\}} + 1$$

Each coordinate $\binom{x}{y}$ was assigned to a matrix entry (i, j) based on the condition:

$$i < \left\lfloor \frac{x}{10} \right\rfloor < i + 10 \text{ and } j < \left\lfloor \frac{y}{10} \right\rfloor < j + 10.$$

To account for differences in the number of participants or data points, values were normalized as:

$$M_{\{i,j\}}^{\{norm\}} = \frac{M_{\{i,j\}}}{N}$$

For group comparisons separate heatmaps were generated for different conditions (e.g., first vs. second session, gaming experience vs. no gaming experience, high vs. low psychopathy), and difference heatmaps were created by subtracting one group's normalized matrix from another:

$$M^{\{difference\}} = M_{\{norm\}}^{\{second\}} - M_{\{norm\}}^{\{first\}}$$

Heatmaps were visualized using the Python seaborn library. The color scale was adjusted to represent the full range of observed counts, ensuring that spatial occupancy patterns were accurately displayed. For difference heatmaps, color normalization was performed separately for each comparison to highlight location-specific behavioral differences.

Ethical Considerations

All studies included in this dissertation were conducted in accordance with the Declaration of Helsinki (2013) and Good Clinical Practice guidelines. Ethical approval was obtained from the local ethics committee of the Hamburg Medical Association (Ärztekammer Hamburg, Germany) for Study 1 and Study 4. For Study 2, approval was granted by the Psychological Ethics Committee of the University Medical Center Hamburg-Eppendorf. Written informed consent was obtained from all participants prior to testing, and participants were informed about the study procedures, potential risks, and their right to withdraw at any time without providing a reason. For analyses based on existing datasets (Study 3), only anonymized data were used, and researchers had no access to personally identifiable information.

Study-Specific Methods

Study 1: Reliability of Repeated Measures on the Virtual EPM

Participants

Recruitment for this study took place via a local biomedical research website, word of mouth, and university postings. Forty-four healthy participants (women = 22, men = 22) aged between 18 and 50 years got invited twice into the Human Behavior Laboratory at the University Medical Center Hamburg-Eppendorf. Following screening for somatic and psychiatric disorders and current drug use, participants provided written informed consent prior to testing. Five participants (11%) did not complete the second testing, resulting in a final sample of 39 participants (women = 19, men = 20; mean age = 25.67 ± 3.74 *SD* years). For the first testing session, the validated EPM_{Sea} was used (Biedermann et al., 2017). At the second session, 28 days after the first session, participants were randomly assigned to complete another naturalistic EPM_{Desert} (n = 18) or a stylized video game-like $EPM_{VideoGame}$ (n = 21) for repeated testing. Anxious temperament was assessed using the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970), and sensation seeking with the Sensation Seeking Scale (SSSV; Zuckerman et al., 1978). Gaming experience was measured via a purpose-designed questionnaire addressing prior gaming exposure (“ever played”) and “how much regular gaming time” participants had for computer and virtual reality games. After each EPM exposure, participants rated their experienced anxiety on a 0-9 scale (0 = “no anxiety”, 9 = “very strong anxiety”). Immersive presence during VR exposure was evaluated with the iGroup Presence Questionnaire (IPQ; Schubert et al., 1999), and side effects were monitored using the Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993). Detailed descriptions of the IPQ and SSQ questionnaires can be found in the General Methods section. All testing sessions were scheduled between 2 pm and 9 pm to minimize circadian influences on endocrine measures.

Additional VR Environments

To examine the effects of repeated exposure and the role of environmental realism, two additional virtual environments were created for the second session: a naturalistic desert scene (EPM_{Desert}) and a stylized video game-like scene with cubic patterns ($EPM_{VideoGame}$) (Figure 2.3). As in the validated EPM_{Sea} used during the first session, both environments included mixed-reality enhancements such as congruent sounds and airflow generated by a ventilator.

After a 28-day interval, participants were randomly assigned to one of the alternative environments, allowing investigation of whether differences in real-world resemblance influence approach-avoidance behavior during repeated testing. For videos please see https://osf.io/sbxpw/?view_only=870b2 d3ca4da47ff803d64621e4b0f50.

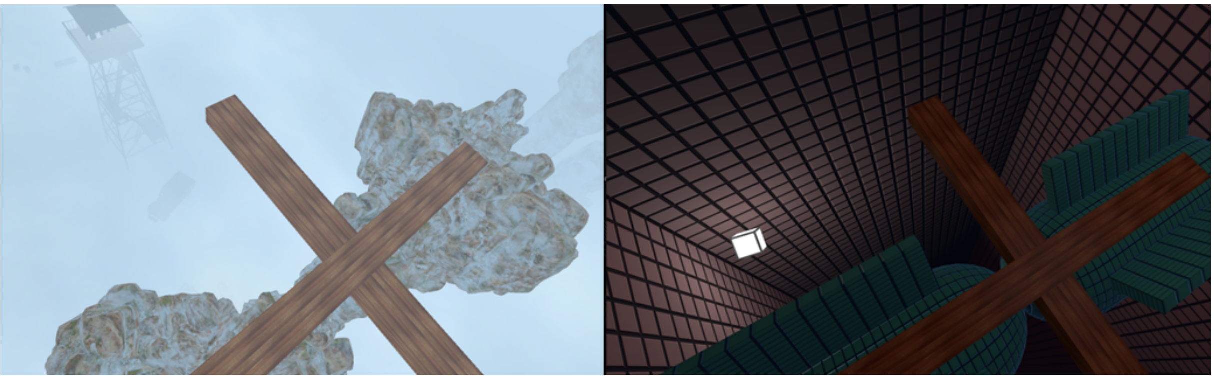


Figure 2.3: Two novel virtual reality environments for repeated exposure of the EPM: EPM Desert (left) resembles a real-world situation, while EPM Video Game (right) shows abstract shapes reminiscent of a video game. Biedermann et al., Behav Res., 56, 187-198 (2024).

Procedure

Upon arrival, participants were screened and provided informed consent at both sessions. They first completed psychometric questionnaires, after which physiological electrodes were attached and an initial saliva sample for alpha-amylase analysis was collected. Participants were then guided blindfolded into the testing room, fitted with the head-mounted display and headphones, and underwent a brief baseline phase in a neutral VR environment. For the first session, all participants explored the validated EPM_{Sea} environment, as described under EPM in the General Methods. After a 28-day interval, they returned for a second session and were randomly assigned to either the naturalistic EPM_{Desert} or the stylized $EPM_{VideoGame}$ environment. The VR setup, electrode placement, timing of baseline and exploration phases, and saliva sampling before, immediately after, and 15 minutes post-testing followed the standardized procedure described in the General Methods. Behavioral, physiological, and endocrine data were collected during each exposure. The following behavioral parameters were used for automated analyses: total time spent on open arms (time on open arms), number of entries of open arms, latency for the first entry of an open arms (latency open arm exploration), and time until participants reached the end of an open arms (latency end exploration). Moreover, time on closed arms and entries of closed arms were assessed as markers for locomotor activity.

For group-level behavioral visualizations, heatmaps were generated as described in the General Methods.

Statistical Analysis

During the first testing session, four SCL datasets (10%) were excluded due to poor quality. At the second testing, nine SCL datasets (23%), two RR datasets (5%), and two ECG datasets (5%) were excluded for the same reason. All statistical analyses were conducted using R (version 4.0.2; http://www.r-project.org) and IBM SPSS Statistics (version 22.0; IBM Corp., Armonk, NY, USA). An a priori power analysis performed with G*Power 3 (Faul et al., 2007) estimated that 38 participants would be needed to achieve a power of 0.85 for detecting a medium effect size ($d = 0.50$) in a paired t-test assessing changes in time spent on open arms ($\alpha = 0.05$). To check for potential baseline differences in anxiety-related factors between participants assigned to EPM_{Desert} and $EPM_{VideoGame}$, sociodemographic variables and trait questionnaire scores were compared using independent samples *t*-tests and chi-squared tests.

Test-retest reliability was evaluated using Pearson correlations and intraclass correlations (ICC, two-way mixed model, consistency type). Psychophysiological measures were analyzed using repeated-measures ANOVA with time (11 intervals: baseline, 0-30 s, 30-60 s, etc.) and exposure (first vs. second testing) as within-subject factors, and EPM version as the between-subject factor. Post hoc comparisons between baseline and the first 30-second interval were performed with paired t-tests. Alpha-amylase concentrations were analyzed with repeated-measures ANOVA including time (T0-T2) and exposure as within-subject factors and EPM version as a between-subject factor. Subjective ratings and behavioral measures of anxiety were analyzed similarly, with exposure as the within-subject factor and EPM version as the between-subject factor. Statistical significance was defined as $p < 0.05$. Effect sizes for ANOVAs are reported as *partial* η^2. Corrections for multiple comparisons were intentionally not applied to avoid reducing the likelihood of detecting differences between the EPM versions.

Study 2: The Impact of Psychopathic Traits on Anxiety-Related Behaviors on the Virtual EPM

Participants

Healthy volunteers (*N* = 170) were recruited via electronic and physical bulletin boards in public areas and on university campus. All participants self-reported the absence of neurological or psychiatric disorders and no regular use of recreational drugs or central nervous system medication. Ten individuals were excluded from analyses due to technical issues (*n* = 6) or failure to follow instructions (*n* = 4), resulting in a final sample of 160 participants. The majority were women (*n* = 107, 67.7%), aged 18-50 years (*M* = 25.62 years), and born in Germany (*n* = 154, 86.5%). Most reported being students (*n* = 101, 56.7%) or employed (*n* = 47, 26.4%); 9 participants (5.1%) were married and 1 participant (0.6%) divorced. Among women, 42% (*n* = 47) reported oral contraceptive use. This variable was not included in the analyses because a subsequent study from our lab showed no effect of estrogen doses on anxiety-related behavior on the human EPM (Nouri et al., 2022).

Procedure

On arrival, participants provided written informed consent and were screened for neurological or psychiatric disorders and medication or recreational drug use. They were then guided to a waiting area where they completed the questionnaires described in the Questionnaires section below. Afterward, participants entered the behavioral laboratory to perform the virtual EPM task (see General Methods for a full description of the setup and procedure). Immediately following the task, they rated their subjective anxiety for each maze location (center, closed arms, open arms) on a 0-9 scale. For group-level behavioral visualizations, heatmaps were generated as described in the General Methods.

Questionnaires

Prior to behavioral testing, participants completed a series of questionnaires to assess individual traits potentially influencing anxiety-related behavior. The Brief Questionnaire of Psychopathic Personality Traits (FPP; Etzler & Rohrmann, 2017) is a 30-item self-report measure adapted from the Psychopathic Personality Inventory (PPI; Lilienfeld & Andrews, 1996). It captures six dimensions of psychopathic traits: Lack of Empathy, Fearlessness, Narcissistic Egocentrism, Impulsivity, Social Manipulation, and Power. Items are rated on a 6-

point Likert scale ranging from 0 "absolutely not true" to 5 "absolutely true". Subscale scores are determined by summing five items each, while the total score is obtained by summing all 30 items. The subscale Fearlessness reflects reduced ability to anticipate or respond to potential negative consequences of one's actions, with sample items such as "When doing something, I rarely think that it might go wrong" and "There are few things that frighten me." Internal consistency was satisfactory, with McDonald's ω ranging from 0.63-0.80 across subscales and $\omega = 0.90$ for the total score. The factorial structure was stable across civilian and inmate samples, and construct validity was supported by high correlations with the PPI-R ($r = 0.68$) as well as significant associations with violent convictions and disciplinary sanctions (Etzler & Rohrmann, 2017).

The Sensation Seeking Scale Form V (SSSV; Zuckerman et al., 1978) consists of 40 forced-choice items designed to measure the propensity to seek varied, novel, and intense experiences. It yields an overall sensation seeking score as well as four distinct subscales: Thrill and Adventure Seeking (e.g., desire for physically risky activities like parachuting), Experience Seeking (e.g., openness to new sensory or mental experiences like solo exploring new cities), Disinhibition (e.g., preference for social and sexual variety), and Boredom Susceptibility (e.g., aversion to predictable routines). A detailed description of the SSSV can be found in the General Methods section.

The Acrophobia Questionnaire (AQ; Cohen, 1977) is a 40-item self-report questionnaire measuring fear of heights across two subscales: Anxiety (AQ-Anxiety) and Avoidance (AQ-Avoidance). Each subscale comprises 20 common height-relevant situations (e.g., "Walking on a foot-bridge over a highway"). The Anxiety scale is rated on a 7-point Likert scale from 0 "not anxious at all" to 6 "extremely anxious", while the Avoidance scale is rated on a 3-point scale from 0 "would not avoid" to 2 "would not do it under any circumstances". Subscale scores are obtained by summing across the 20 items, yielding a maximum of 120 for Anxiety and 40 for Avoidance. Internal consistency ranged from $\alpha = 0.72$-0.92 (Steinman & Teachman, 2014), and test-retest reliability was high (median $r = 0.82$; Cohen, 1977). Known-groups validity was supported by significantly lower scores in non-acrophobic compared to acrophobic samples, and convergent validity was evidenced by negative correlations with behavioral approach tests. The AQ was also sensitive to therapeutic change, and construct validity was further supported by correlations between AQ change scores and self-rated improvement (Cohen, 1977).

After completing the EPM task, participants also provided a post-task subjective anxiety rating, indicating their perceived anxiety level in each maze location (center, closed arms, open arms) on a single-item scale ranging from 0 “no anxiety” to 9 “very strong anxiety”.

Statistical Analysis

Statistical analyses were conducted using IBM SPSS Statistics 23.0 (IBM Corp., Armonk, NY, USA) and R software (version 4.0.2). Pearson correlation analyses were used to examine associations between psychopathic traits and anxiety-related measures. Additionally, participants were divided into three groups (low, medium, and high) based on their total scores on the Factorial Psychopathy Profile (FPP) using a quartile split. Descriptive statistics were calculated separately for these three FPP groups.

To further explore the influence of specific FPP dimensions, a multivariate linear regression model was computed with the FPP subscale scores as independent variables and four behavioral outcomes as dependent variables. Wilks' Lambda was used to evaluate the multivariate test. In addition to the multivariate approach, separate multiple regression models were fitted for each of the following dependent variables: time spent on the open arms, latency to first open arm entry, number of open arm entries, latency to exploration of the end of the open arm (endexploration), and subjective anxiety. All values are reported as mean ± standard error of the mean (SEM). Statistical significance was defined as $p < 0.05$.

Study 3: The Effect of Gaming Experience on Behavioral Responses on the Virtual EPM

Participants

The present study was based on existing datasets comprising a total of 446 healthy individuals aged 18 to 50 years, all of whom participated in previous behavioral studies using the validated EPM paradigm (see General Methods). Participants were recruited via word of mouth, online advertisements on local research recruitment platforms, and university announcements. All participants were naïve to the EPM and were screened for somatic and psychiatric disorders as well as current drug use. Only healthy individuals without current drug use were included in this study. Informed written consent was obtained from all participants prior to testing. The existing datasets were accessed for research purposes starting on October 15, 2023. The authors had no access to personally identifying information at any stage. Only complete cases were included in the statistical analyses. Sixteen participants did not complete the gaming experience questionnaire, and two lacked behavioral outcome data, resulting in a final sample of 428 individuals (women = 260, men = 168) with a mean age of 26.14 years (*SD* = 5.01).

Procedure

This study involved secondary analysis of existing datasets collected as part of previously conducted behavioral experiments using the virtual EPM. All participants had completed a standardized testing protocol that included exploration of the virtual EPM in a single session (as described under General Methods), followed by completion of self-report questionnaires. The single-item measure of gaming experience (described below) was part of the original assessment battery. For the present analysis, data access began on October 15, 2023. Only anonymized, complete datasets were included. No new data were collected for the present analysis. For group-level behavioral visualizations, heatmaps were generated as described in the General Methods.

Questionnaires

Gaming experience was defined as having regularly played video and computer games at least once per week. It was assessed using a purpose-designed, single-item questionnaire: "Have you ever played a computer game regularly (at least weekly)?" The item was dichotomous, with response options "yes" or "no." Based on this variable and the self-reported gender, the final dataset comprised four subgroups: men with gaming experience (n = 141), men without gaming experience (n = 27), women with gaming experience (n = 96), and women without gaming experience (n = 164).

Statistical Analysis

The statistical analysis was performed using Stata Statistical Software (StataCorp. 2023. Release 18. College Station, TX: StataCorp LLC.). Due to the distribution of the data, robust linear regressions were applied to examine the relationships between key behavioral parameters (time on open arms, number of open arm entries, latency to open arm exploration, and latency to end exploration) as outcome variables and the predictors gender and gaming experience. Also, the interaction between gender and gaming experience was included in the models. Significance levels were assessed at 0.05 level. For the regression models, the regression coefficients are reported with the 95% confidence interval. The global model fit is assessed with the F-Test and the explained variance is calculated via the coefficient of determination (R^2). To visualize the interaction effect between gender and gaming experience, plots of the estimated marginal means with 95% confidence intervals were generated. Additionally, to facilitate interpretation and comparison of effect sizes, the same regression models were conducted separately for each gender. For these stratified models, standardized regression coefficients for the effect of gaming experience on behavioral outcomes are reported.

Study 4: Validation of the Dark-Light Box in Virtual Reality

Participants

The study comprised two samples of participants. The healthy sample, used for general validation of the DLB, included a total of 74 individuals (men = 29, 39.2%; women = 45, 60.8%) aged between 18 and 61 years (*M* = 30.69, *SD* = 9.01), recruited through public advertising (flyers, online announcements, social media, and word of mouth). The clinical sample consisted of 40 psychiatric patients (men = 11, 27.5%; women = 29, 72.5%) aged between 19 and 59 years (*M* = 32.48, *SD* = 11.77), recruited from the Department of Psychiatry and Psychotherapy at the University Medical Center Hamburg-Eppendorf. All patients had one of the following primary diagnoses: anxiety disorder (AD, *n* = 15, 37.5%; including agoraphobia with/without panic disorder, panic disorder, generalized anxiety disorder, social phobia, or specific phobia), borderline personality disorder (BPD, *n* = 20, 50%), or post-traumatic stress disorder (PTSD, *n* = 5, 12.5%). To form a control group for the clinical sample, age- and gender-matched participants were selected from the healthy sample. The mean absolute age difference between patients and their matched controls was 1.2 years (*SD* = 2.4), and all controls were gender-matched (100%). As 14 additional matched controls were required, 14 healthy individuals were subsequently recruited, resulting in a final total of 74 healthy participants (see supplementary Figures S1 and S2 for recruitment flyers of both samples).

Participants in both samples were required to be at least 18 years old and to have sufficient fluency in spoken and written German. For the healthy sample, inclusion required the absence of current physical or mental illness. Individuals were excluded if they were pregnant or breastfeeding; taking medication (with the exception of over-the-counter drugs taken more than one week prior to participation and oral contraceptives); or had a history of internal medical conditions (e.g., cardiac arrhythmias or cardiovascular disease), chronic pain, neurological disorders (e.g., Parkinson’s disease), severe psychiatric disorders (e.g., psychosis, major depression, or severe anxiety disorders), acute infections, or severe orthopedic conditions affecting the lower limbs. For the clinical sample, participants were recruited from the Department of Psychiatry and Psychotherapy if they had at least one of the eligible psychiatric diagnoses (see above) and categorized by their primary diagnosis. Comorbid diagnoses were allowed except for the disorders mentioned in the exclusion criteria. The most frequent self-reported comorbidities in patients with AD and PTSD were depression (n = 8 and n = 3, respectively), and for patients with BPD depression (n = 4) and PTSD (n = 5). Patients were excluded if they were pregnant or breastfeeding; currently taking benzodiazepines; or had internal medical conditions, neurological disorders, acute infections, or severe orthopedic impairments of the lower limbs. Further exclusion criteria were current alcohol or drug abuse, the presence of a psychotic disorder or bipolar disorder, and acute suicidality.

Procedure

Participants were contacted via email and scheduled for testing at the Human Behavior Laboratory at the University Medical Center Hamburg-Eppendorf. Upon arrival, they first completed pre-DLB questionnaires (see section "Questionnaires" below). Healthy participants were then fitted with physiological recording equipment. Physiological measures and saliva sampling (for endocrinological measurements) were only conducted on healthy participants, as detailed in the General Methods section. Participants in the clinical sample proceeded directly to behavioral testing after completing the pre-DLB questionnaires. All participants began with the DLB virtual reality task. Afterward, the physiological equipment was removed from healthy participants, and all participants completed a second set of post-DLB questionnaires (see section "Questionnaires" below). Lastly, all participants underwent the behavioral testing in the EPM. No physiological data or saliva samples were collected during this second behavioral task for healthy participants. The exact VR testing protocol is described in the General Methods section.

Questionnaires

Pre-DLB, participants completed a set of questionnaires assessing sociodemographic characteristics, individual differences, and psychological traits relevant to the VR tasks. These included the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970) to measure trait and state anxiety levels, and the Sensation Seeking Scale (SSSV; Zuckerman et al., 1978) to assess sensation-seeking tendencies. Detailed descriptions of the STAI and the SSSV can be found in the General Methods section. Participants also reported their responses to the following standardized questionnaires:

The Agoraphobic Cognitions Questionnaire (ACQ; Chambless et al., 1984) assesses thoughts concerning negative consequences of experiencing anxiety. The ACQ consists of 14 items, of which six items reflect behavioral or social consequences of losing control (e.g., "I will go crazy") and eight items address catastrophic physical concerns (e.g., "I will have a heart attack"). Each item is rated on a 5-point scale ranging from 1 "thought never occurs" to 5 "thought always occurs". A total score is obtained by summing across all 14 items; subscale scores for social/behavioral versus physical concerns can additionally be derived. Internal consistency was satisfactory ($\alpha = 0.80$), and test-retest reliability was high ($r = 0.86$ across 8 days). Exploratory factor analysis supported a two-factor structure accounting for 46% of the variance. Construct validity was supported by moderate correlations with measures of avoidance, anxiety, depression, and neuroticism ($r = 0.21$-0.67), while discriminant validity was indicated by non-significant associations with psychoticism. The ACQ was shown to be responsive to treatment and reliably distinguished agoraphobic patients from controls (Chambless et al., 1984).

The Intolerance of Uncertainty Scale (IUS; Freeston et al., 1994) was used to assess negative beliefs, emotions, and behavioral responses associated with uncertainty. The scale consists of 27 items (e.g., "A small unforeseen event can spoil everything, even with the best of planning") rated on a 5-point Likert scale from 1 "not at all characteristic of me" to 5 "entirely characteristic of me", with higher scores reflecting greater intolerance of uncertainty. A total score is calculated by summing across all items (range 27-135). The IUS demonstrated excellent internal consistency ($\alpha = 0.91$). Known-groups validity was supported by significantly higher scores among individuals meeting criteria for generalized anxiety disorder (GAD) compared to non-GAD controls. Construct validity was evidenced by moderate to strong correlations with measures of worry, anxiety, and depression ($r = 0.52$-0.63), and partial correlations controlling for anxiety and depression remained significant, indicating discriminant validity (Freeston et al., 1994).

The Claustrophobia Questionnaire (CLQ; Radomsky et al., 2001) consists of 26 items to assess claustrophobic fear. Items are rated on a 5-point Likert scale ranging from 0 "not at all anxious" to 4 "extremely anxious." Factor analysis supported a two-factor structure: (1) Fear of Restriction (RS; e.g., "Being locked in a small dark room") and (2) Fear of Suffocation (SS; e.g., "Swimming while wearing a nose plug"). Scores are obtained by summing all items (range 0-104) or separately for each subscale. The CLQ demonstrated excellent internal consistency ($\alpha = 0.85$-0.96 across subscales), and test-retest reliability over a 2-week interval was high (r

= 0.77-0.89). Known-groups validity was supported by significantly higher scores in claustrophobic participants compared to controls. Discriminant validity was indicated by weak correlations with general measures of anxiety and depression (r = 0.12-0.33), while predictive validity was demonstrated by strong correlations with fear responses during exposure to enclosed spaces (r = 0.52-0.64) (Radomsky et al., 2001).

To assess fear of darkness, a 20-item Fear of Darkness Questionnaire (FOD) was used. The items capture subjective fears and physiological symptoms in situations without light (e.g., "I feel more nervous and anxious than usual," "I can feel my heart pounding rapidly"). Responses are given on a 4-point Likert scale ranging from 1 "sometimes or a small part of the time" to 4 "most or all of the time". The questionnaire was provided as an internal document in 2019; to the best of my knowledge, no published version or psychometric validation is available. In the present sample, the total scale showed high internal consistency (Cronbach's α = 0.89), and reliability estimates remained stable when individual items were deleted (range = 0.88-0.89). Items were coded such that higher scores indicate greater fear of darkness (positively worded items reverse-scored), and a sum score was computed. The complete set of items is presented in Supplementary Table S2.

Post-DLB, participants rated their subjective experience in VR using custom-designed items assessing perceptions and emotional reactions, each rated on a 9-point Likert scale (1 = “not at all”, 9 = “very strongly”). For example, subjective anxiety was assessed with the item: “During VR, I felt anxious.” Participants also completed the iGroup Presence Questionnaire (IPQ; Schubert et al., 1999) to measure their sense of presence in VR and the Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993) to assess symptoms of simulator-induced discomfort. Detailed descriptions of the IPQ and SSQ can be found in the General Methods section. In addition, participants were asked to provide spatially specific anxiety ratings by marking their perceived anxiety levels for different positions within the virtual room on an illustration of the environment, using a custom scale (0 = “no anxiety”, 9 = “very strong anxiety”).

Video Analysis

Behavioral testing in the DLB was video recorded for all participants (*N* = 114) and analyzed using focal sampling over the full 5-minute exposure period. The software BORIS (v7.13.9;Friard & Gamba, 2016) was used for behavioral coding. To develop the ethogram, initial exploratory analysis was conducted on videos of healthy participants. Prominent and recurring behaviors were identified and iteratively categorized. The ethogram was refined through multiple coding rounds, resulting in a finalized coding scheme. This finalized ethogram was then applied to the entire dataset, and all videos were coded accordingly (Table 2.1). Example videos of selected behaviors can be found under this link: https://osf.io/e7n3z/?view_only=2ddaae1993d14c88bfdcc7e6dbbf8881.

Table 2.1: Ethogram of observed behaviors in the DLB.

Behavior	Type	Description
Crossed arms	State event	Placing arms crossed into each other at chest height with no other movement or intention
Crossed arms disguised	State event	Crossing both arms in front of the body also holding on wrists, holding elbows, holding belly, fiddling on fingers, fiddling on sleeves or jewelry
One arm crossed	State event	Crossing one arm in front of the body using the arm to hold on the other wrist, elbow, belly, fiddle on sleeves or jewelry
Fiddling	State event	The act of touching, moving, or adjusting something in a small, often repetitive manner. Example: fiddling on sleeve or jewelry, fingers etc.
Crossed legs	State event	Standing with one leg crossed over the other
Foot edging/foot shifting	State event	Slightly lifting one foot inward or balancing on the toes, revealing the bottom of the shoe
Bouncing	State event	Bouncing/shaking of feet and legs
Shuffling feet	State event	Rubbing feet or one foot on ground
Exploration	State event	Examination of the DLB with hands and or feet. As movement: exploration of the DLB with one foot or arms extended moving forward with the rest of the body following
Re-examination	State event	Re-examination of already explored or visited parts of the DLB with hands and feet
One leg	State event	Standing shifting the weight on one leg
Looking up	State event	Looking up exposing throat
Arms bent	State event	Walking or standing with both forearms bend toward the upper arms. "L" or "V" shape at torso level

Out of sight	State event	Participant moves out of sight
Immobile	State event	Almost complete motionlessness without a tense body. Typically, long termed

To assess the reliability of behavioral coding, a second independent rater coded a randomly selected subset of 10 videos using the finalized ethogram. Coding was conducted in BORIS (V7.13.9; Friard & Gamba, 2016), and intraclass correlation coefficients (ICCs) were calculated using IBM SPSS Statistics (Version 29), applying the absolute agreement model (Hallgren 2012). Reliability was assessed for 15 behavioral categories, each coded for both duration and frequency, resulting in 30 coding units. Of the 30 behavioral units coded, 22 demonstrated good to excellent reliability (ICC > 0.60), while 8 showed only moderate to poor agreement (ICC < 0.60). Interpretation of ICC values followed the guidelines proposed by Cicchetti (1994), with thresholds for poor (< 0.40), moderate (0.40-0.59), good (0.60-0.74), and excellent reliability (≥ 0.75). Single-measures ICCs reflect how reliably an individual rater can code a given behavior, whereas average-measures ICCs represent the overall reliability of a behavioral category when assessed by multiple raters. Given that this study introduced a new ethogram, both single-measures and average-measures ICC values are reported in the Supplementary Material (see supplementary Table S3), providing a comprehensive evaluation of coder agreement and the robustness of the behavioral categories.

For the preparation of the statistical analysis, the set of coded behaviors was refined based on frequency of occurrence and inter-rater reliability (see Supplementary Tables S3, S4). Behaviors occurring in fewer than 5% of observations were considered rare and excluded. However, conceptually related rare behaviors were collapsed into broader categories to retain information while ensuring interpretability. Specifically, all behaviors involving the movement or placement of the arms in front of the torso (crossed arms, crossed arms disguised, one arm crossed, arms bent) were combined into a single category termed “arms before body behaviors.” Similarly, all leg- or foot-related behaviors (crossed legs, bouncing, foot edging/foot shifting, shuffling feet, one leg) were combined under the category “lower body behaviors.” These groupings, along with parameters recorded via the head-mounted display, resulted in the final selection of behavioral variables used for analysis (Table 2.2).

Table 2.2: Final selection of behavioral variables used for statistical analysis.

Behavioral variables	Definition	Measurement Method
Time in Dark (s)	Time in seconds that participants spent in the aversive (dark) parts of the DLB	Via head-mounted display
Average velocity (m/s)	The average velocity in meters per second at which participants moved	Via head-mounted display
Latency end exploration (s)	Time in seconds until the participant reaches the rear wall in the aversive (dark) part of the DLB	Via head-mounted display
Latency first visit (s)	Time in seconds until the participant stepped from the safe (bright) over to the aversive (dark) part of the DLB	Via head-mounted display
Total distance (m)	The total distance in meters that participants covered	Via head-mounted display
Exploration (duration: s, frequency: n)	Time in seconds and frequency counts of exploration behavior (see Ethogram)	Via video analysis
Re-examination (duration: s, frequency: n)	Time in seconds and frequency counts of re examination behavior (see Ethogram for details)	Via video analysis
Looking up (duration: s, frequency: n)	Time in seconds and frequency counts of looking up behavior (see Ethogram for details)	Via video analysis
Arms before body behaviors (duration: s, frequency: n)	Time in seconds and frequency counts of behavior that involves having arms in front of the torso (see paragraph above and Ethogram for details)	Via video analysis
Lower body behaviors (duration: s, frequency: n)	Time in seconds and frequency counts of behavior that involve feet and leg movements (see paragraph above and Ethogram for details)	Via video analysis

Statistical Analysis

Statical analysis was conducted using IBM SPSS Statistics (version 29.0; IBM Corp., Armonk, NY, USA) and R (version 4.3.1; http://www.r-project.org). Data from the healthy sample were used to analyze physiological and endocrinological responses in the DLB. To assess changes in autonomic and endocrine stress markers, physiological signals (SCL, RR, and ECG) were compared between a 30 second pre-test baseline interval and eleven subsequent 30 second intervals recorded during DLB exploration. Repeated-measures ANOVAs were applied for all physiological parameters. Salivary cortisol and alpha-amylase were analyzed at three time points: T0 (baseline), T1 (immediately post-test), and T2 (15 minutes post-test). Both markers were log-transformed and subjected to repeated-measures ANOVA with time as a within-subject factor. Bonferroni corrections were applied to all post hoc comparisons.

For analyzing subjective experience in VR, descriptive statistics (means and standard deviations) were computed for self-report measures of spatial presence, involvement, and realism (IPQ), as well as simulator sickness symptoms (SSQ; nausea, oculomotor disturbances, disorientation) across the entire healthy sample (N = 74). To assess spatial variation in anxiety responses across the DLB, subjective fear ratings were analyzed using repeated-measures ANOVA with Greenhouse-Geisser correction, followed by Bonferroni-corrected post hoc t-tests. Only healthy participants who visited all defined DLB areas ($n = 40$ complete cases) were included in this analysis.

Behavioral data were analyzed to investigate relationships with independent variables such as questionnaire scores and general anxiety ratings. To meet assumptions of normality, several parameters (latencies, re-examination, arms-before-body, lower-body behaviors) were log-transformed. Pearson correlations were first conducted in SPSS to identify associations between independent variables and behavioral outcomes. The correlation matrix was further used to identify the main behavioral variables with clear linear relationships to the independent variables and to avoid including redundant behavioral variables in further analyses.

Multivariate multiple regression was used to assess the predictive value of these independent variables. Multicollinearity was assessed using the Variance Inflation Factor (VIF), which ranged from 1.046 to 2.079, indicating no collinearity concerns. The corresponding tolerance values (0.481 to 0.956) further supported this. To account for moderate heteroscedasticity in three outcome variables (re-examination, arms before body, and lower body behaviors) observed in regression residuals, overall multivariate test statistics were assessed using Pillai's Trace in SPSS, which is regarded as the most robust statistic. Individual regression coefficients were evaluated in R using t-tests of coefficients with the *coeftest* function and robust standard errors (HC1 estimator, *sandwich* package). *P*-values were adjusted using the Benjamini-Hochberg correction. Effect sizes were reported as *semi-partial* R^2 (sr^2), reflecting the unique contribution of each predictor to the outcome variables.

To complement the multivariate regression analyses, group-level comparisons were conducted to examine whether behavioral measures in the virtual DLB differed between participants reporting high versus low levels of subjective state anxiety during the task. Participants were dichotomized into low- and high-anxiety groups using a median split on in-task anxiety ratings (scale 1-9), resulting in 41 participants in the low-anxiety group (ratings 1-3; LA) and 33 participants in the high-anxiety group (ratings 4-9; HA). The same seven behavioral outcome variables were selected for group comparisons: looking-up frequency, time spent in dark areas (s), total distance covered (m), log latency to first visit (s), log re-examination frequency, log arms-before-body duration, and log lower-body behaviors duration. For each variable, Welch's *t*-tests were used to compare LA and HA groups, as this test does not assume equal variances and is robust against unequal group sizes. To account for multiple testing across the seven comparisons, *p*-values were adjusted using the Benjamini-Hochberg procedure. Additionally, an exploratory analysis of looking-up duration was done post hoc and thus not corrected for multiple testing.

To examine convergent validity between DLB and EPM parameters, Pearson correlations were calculated using matched behavioral data for each participant. Of the initial 74 healthy participants, 71 completed both paradigms (due to early termination of the EPM) and were included in this analysis.

Data from the clinical sample were used for comparisons between the clinical group (n = 40) and individually matched healthy controls (n = 40). Paired-samples t-tests were used to analyze each behavioral variable. Standardized effect sizes were calculated as Cohen's d_z, derived from the t-statistic and sample size. Multiple comparisons were corrected using the Benjamini-Hochberg procedure. To further explore the diagnostic specificity of DLB responses, exploratory subgroup analyses were performed. Participants with Anxiety Disorder (AD, n = 15) and Post-Traumatic Stress Disorder (PTSD, n = 5) were combined due to overlapping symptomatology and limited sample size. This resulted in two equally sized diagnostic subgroups (n = 20 each): AD/PTSD and Borderline Personality Disorder (BPD). Each subgroup was compared to its matched controls using paired-samples *t*-tests. Given the small sample sizes and exploratory nature of this analysis, no correction for multiple comparisons was applied. These comparisons serve as a preliminary investigation of subgroup-specific behavioral patterns that may guide future diagnosis-focused research.

Chapter 3 Results

Study 1: Reliability of Repeated Measures on the Virtual EPM

As participants were randomly assigned to the two groups, sociodemographic factors, including video game and virtual reality experience, as well as trait anxiety and sensation seeking (SSSV total score), were assumed to be equally distributed between the groups. As a result, these potential confounders were not considered in the subsequent analyses.

Anxiety-Related Behaviors Upon Repeated Exposure

Repeated exposure did not have a significant effect on core behavioral measures in the EPM. Specifically, no significant differences were found in time on open arms ($F_{1,37} = 1.80$, $p = 0.188$, $\eta2_{partial} = 0.046$), number of entries of open arms ($F_{1,37} = 1.21$, $p = 0.278$, $\eta2_{partial} = 0.032$), latency of open arm exploration ($F_{1,37} = 0.66$, $p = 0.422$, $\eta2_{partial} = 0.018$), or latency of open arm end exploration ($F_{1,37} = 3.21$, $p = 0.081$, $\eta2_{partial} = 0.080$; Table 3.1). Similarly, no significant differences were observed for locomotor activity, such as the total distance covered on the EPM ($F_{1,37} = 1.75$, $p = 0.194$, $\eta2_{partial} = 0.045$) and the average velocity on closed arms ($F_{1,37} = 0.31$, $p = 0.581$, $\eta2_{partial} = 0.008$) also showed not significant difference between the baseline and second testing sessions. Additionally, the versions of the EPM ($EPM_{VideoGame}$ vs. EPM_{Desert}) did not significantly affect behavior (all $p > 0.3$).

Table 3.1: Mean and standard deviation (*SD*) of the core behavioral markers on the EPM. Biedermann et al., Behav Res., 56, 187-198 (2024).

Behavioral measures	**1st exposure Mean (*SD*)**	**2nd exposure Mean (*SD*)**	**$EPM_{VideoGame}$ Mean (*SD*)**	**EPM_{Desert} Mean (*SD*)**
entries open arms (n)	4.1 (2.6)	3.6 (2.3)	3.5 (2.3)	3.8 (2.2)
latency open arms exploration (s)	73.5 (98.6)	85.8 (88.1)	87.0 (96.9)	84.5 (79.5)
latency end exploration (s)	180.6 (119.5)	211.6 (110.8)	191.9 (124.2)	234.6 (91.0)
time on open arms (s)	91.0 (56.1)	81.2 (50.2)	79.6 (58.5)	83.2 (40.1)

total distance covered (m)	13.4 (5.9)	12.6 (4.8)	13.0 (4.8)	12.2 (5.0)

A high correlation was found between baseline and repeated exposure for anxiety-related behavior on the EPM (Figure 3.1; all $p < 0.001$). To assess test-retest reliability for the behavioral measures, intra-class correlations (ICC) were calculated for all behavioral measures. An ICC > 0.7 is typically considered "sufficient reliability" in reliability studies (Hripcsak & Heitjan, 2002). All ICC values for the behavioral measures in this study exceeded this threshold: entries = 0.72 [95% CI = 0.46-0.85], latency = 0.71 [95% CI = 0.45-0.85], latency-end = 0.72 [95% CI = 0.46-0.85], time = 0.80 [95% CI = 0.62-0.89], total distance = 0.84 [95% CI = 0.69-0.92], velocity = 0.79 [95% CI = 0.60-0.89]. Visual inspection of the heatmaps revealed no noticeable differences between the two time points (Figure 3.2).

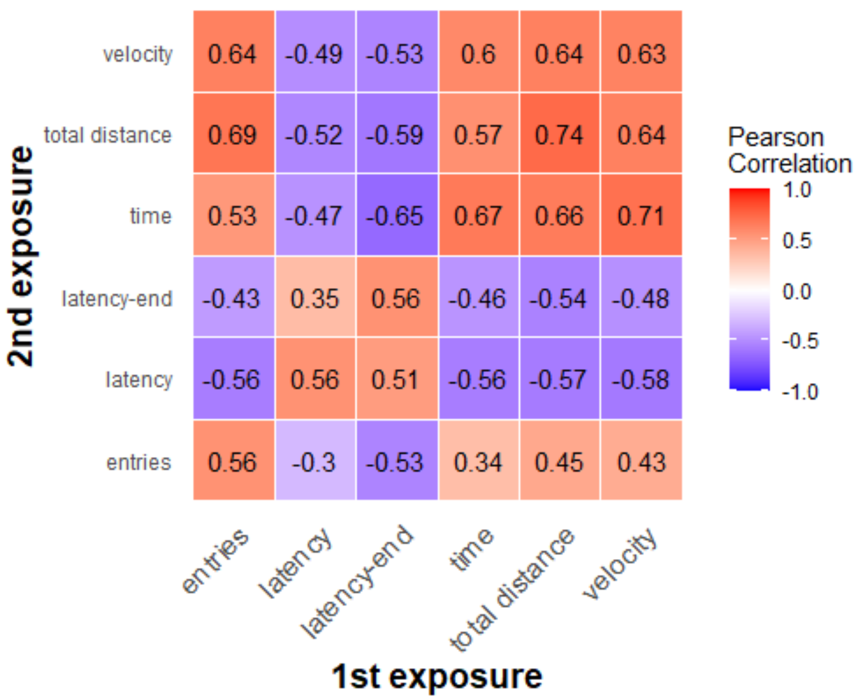


Figure 3.1: Correlation matrix of behavioral measures of the EPM at 1st and 2nd exposure. Numbers represent correlation coefficients. Four measures of anxiety-related behavior: entries = total entries open arms, latency = latency open arms exploration, latency-end = latency end exploration, time = time on open arms; as well as two measures of locomotion: total distance = total distance covered, velocity = average velocity on closed arms. Biedermann et al., Behav Res., 56, 187-198 (2024).

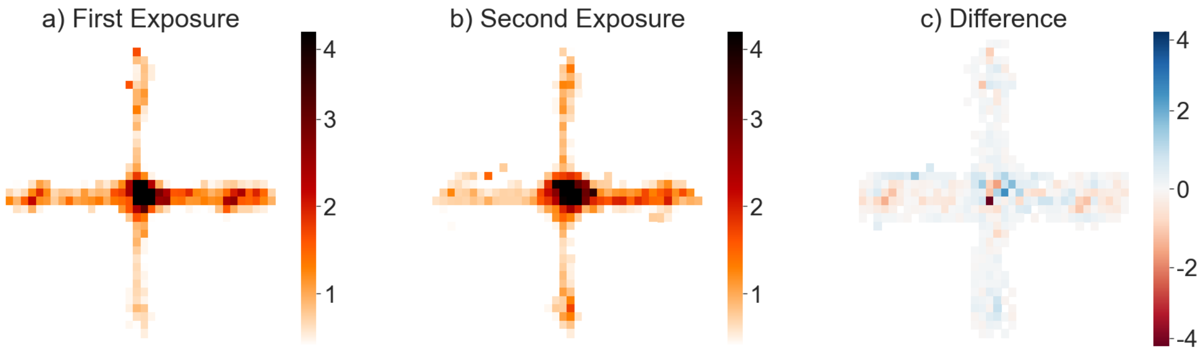


Figure 3.2: Heatmaps indicate length of stay of at least 0.5 s for the first (a) and second (b) exposure on 10 cm² areas. Darker values indicate a longer average length of stay in that area. Figure c depicts differences in mean length of stay between the first and the second measurement. Red hues indicate that participants spent less time in an area, and blue hues indicate that participants spent more time in an area on the second exposure. Biedermann et al., Behav Res., 56, 187-198 (2024).

The Influence of Repeated Exposure to the EPM on Autonomic Reactions

Exposure to the EPM elicited activation of the autonomous nervous system. However, repeated exposure to the EPM altered the pattern of autonomic response for some of the markers. Salivary alpha-amylase levels, expectedly, increased immediately following EPM exposure, as reflected by a significant effect of time ($F_{2, 74} = 5.278$, $p = 0.007$, $\eta2_{partial} = 0.125$). Repeated exposure as well as EPM version did not significantly affect the salivary alpha-amylase response (all $p > 0.3$). Post hoc comparison revealed a significant increase of salivary alpha-amylase levels between T0 and T1 during first (t =-3.62, df = 38, p = 0.001) and with repeated exposure (t = -2.25, df = 38, p = 0.03).

In terms of physiological responses, both respiratory rate (main effect of time: $F_{6.53, 216.64} = 16.88$, $p < 0.001$, $\eta2_{partial} = 0.345$) and skin conductance levels (main effect of time: $F_{2.52, 83.28} = 17.41$, $p < 0.001$, $\eta2_{partial} = 0.431$) showed increases over time. Repeated exposure had no effect on respiratory rate ($p > 0.3$), but it did lead to increased skin conductance levels ($F_{1, 83.28} = 4.98$, $p = 0.036$, $\eta2_{partial} = 0.178$). Heart rate also showed a significant effect of time ($F_{5.26, 202.39} = 3.68$, $p < 0.001$, $\eta2_{partial} = 0.103$) in the repeated-measures ANOVA. However, heart rate pattern at repeated exposure was different then at first exposure, indicated by a significant interaction between repeated exposure and time ($F_{6.33, 202.39} = 3.52$, $p = 0.002$, $\eta2_{partial} = 0.099$). Post hoc testing revealed significant increase in heart rate (t = -4.52, df = 38, $p < 0.001$), respiratory rate (t = -8.27, df = 38, $p > 0.001$), and skin conductance levels (t = -6.61, df = 36, $p < 0.001$) between baseline and onset of the experiment (0-30s) during first exposure. At repeated exposure, an increase was observed only for respiratory rate (t = -7.29, df = 36, p > 0.001) and skin conductance levels (t = -8.51, df = 29, p < 0.001). No significant change in heart rate (t = 0.36, df = 36, p = 0.72) was found between baseline and onset of behavioral testing during repeated exposure. Interestingly, at repeated exposure baseline heart rate levels were descriptively higher than those observed during first exposure, with no response following the onset of EPM testing (Figure 3.3). No significant effect of EPM version was found on physiological measures (all p > 0.3).

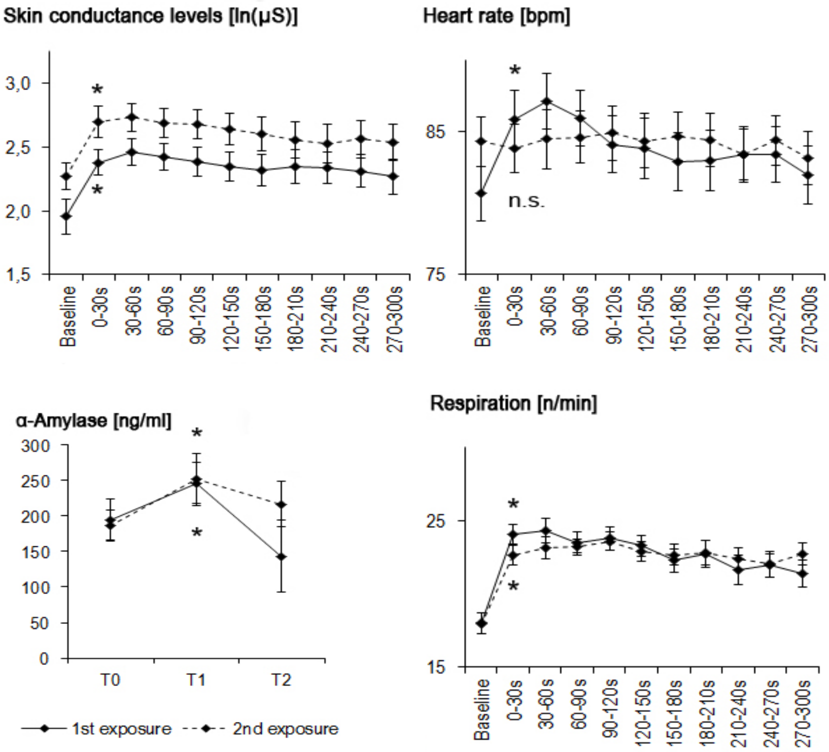


Figure 3.3: Repeated exposure to the EPM and autonomic responses. T0 = before, T1 = directly after and T2 = 15 minutes after EPM. Data points represent Means ± SEM. * represents a significant difference between T0 and T1 or baseline and 0-30s in post hoc tests. N.s. is not significant for the post hoc comparison between baseline and 0-30s at 2nd exposure. Biedermann et al., Behav Res., 56, 187-198 (2024).

The Subjective Response to Repeated EPM Exposure

Between the first and second exposure to the EPM no significant differences were observed in Spatial presence (p = 0.67), emotional involvement (p = 0.13), and perceived realism (p = 0.92), as measured by the iGroup Presence Questionnaire, or total side effects (p = 0.11) as assessed by the Simulator Sickness Questionnaire. These measures were also comparable between the $EPM_{videoGame}$ and EPM_{Desert} versions (spatial presence p = 0.90, emotional involvement p = 0.63, perceived realism p = 0.94, total side effects p = 0.22). However, participants reported higher levels of anxiety at second (repeated) exposure to the EPM ($F_{1,37}$ = 6.19, p = 0.018, $\eta 2_{partial}$ = 0.143; Figure 3.4). No significant difference in subjective anxiety between $EPM_{VideoGame}$ and EPM_{Desert} was found (p = 0.71). Self-reported anxiety levels at baseline and second testing highly correlated with each other (r = 0.61, $p < 0.001$).

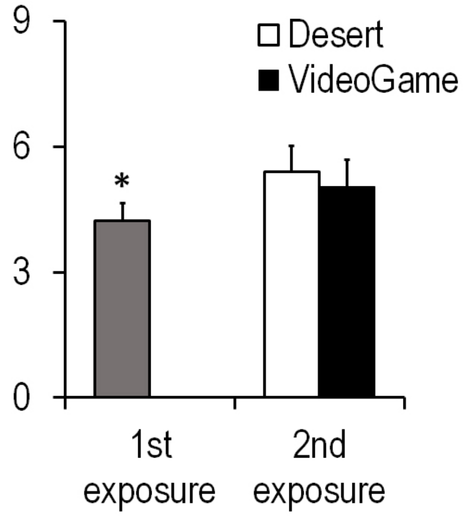
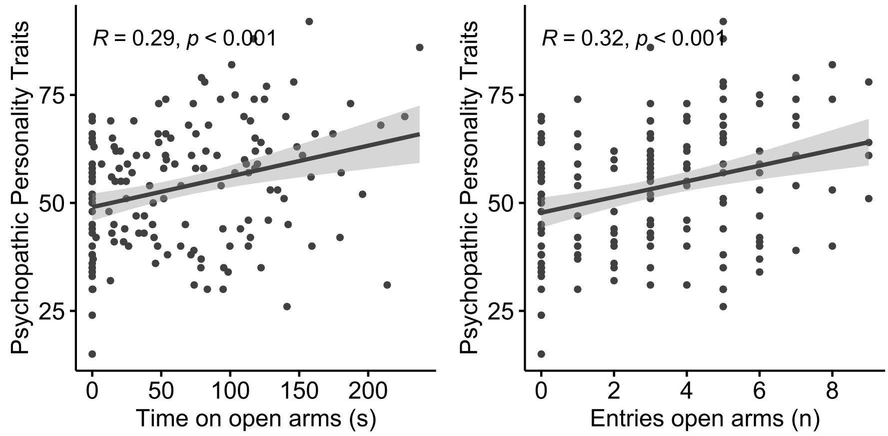


Figure 3.4: Subjective anxiety levels at 1st and 2nd exposure to the EPM.
Bars represent Means + *SEM*, * $p < 0.05$. Biedermann et al., Behav Res., 56, 187-198 (2024).

Study 2: The Impact of Psychopathic Traits on Anxiety-Related Behaviors on the Virtual EPM

Associations of Psychopathic Traits With Measures of Anxiety-Related Constructs

Higher psychopathic traits were associated with increased approach behavior and reduced avoidance on the EPM. The total FPP sum score correlated significantly with all measures of anxiety-like behavior on the EPM: time spent on open arms (r = 0.30, p < 0.001), number of entries into open arms (r = 0.32, p < 0.001), latency open arms exploration (r = -0.29, p < 0.001), and latency open arms end exploration (r = -0.30, p < 0.001; Figure 3.5). These correlations remained significant when controlling for age and gender in partial correlations (time on open arms: r = 0.26, p = 0.002; number of entries to open arms: r = 0.26, p = 0.002; latency open arm exploration: r = -0.23, p = 0.006; latency open arm end exploration: r = -0.21, p = 0.012). On the subscale level, not only the Fearlessness subscale of the FPP but also the other subscales were related to behavior on the EPM (Table 3.2). Psychopathic traits were also negatively correlated with subjective levels of anxiety on the EPM (r = -0.23, p = 0.004). Sensation seeking (SSSV) was positively associated with psychopathic traits (r = 0.33, p < 0.001), whereas general levels of acrophobia (AQ) were not significantly correlated (r = -0.13, p = 0.11). Additional correlations at the subscale level, as well as associations with acrophobia and sensation seeking, are provided in Supplementary Table S1.



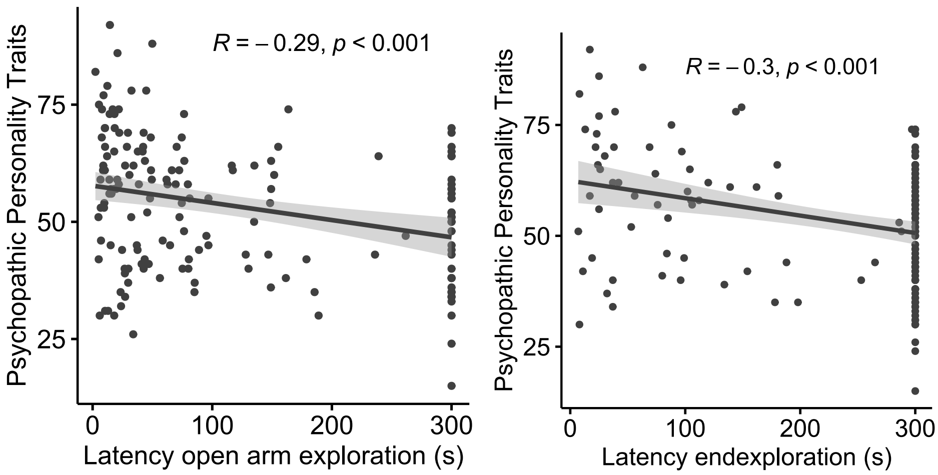


Figure 3.5: The relation between the core features of anxiety-related behavior on the EPM and sum-scores of the FPP. Voulgaris et al., Sci Rep 14, 11832 (2024).

Table 3.2: Pearson Correlations between psychopathic traits and behavioral measures on the EPM as well as subjective anxiety. Values represent Pearson correlation coefficients with associated *p*-values. Voulgaris et al., Sci Rep 14, 11832 (2024).

	Lack of Empathy	**Fearlessness**	**Narcissistic Egocentricity**	**Impulsivity**	**Social Manipulation**	**Power**	**PPT Sum-score**
Latency open arm exploration (s)	-0.198 (*p* = 0.012*)	-0.359 (*p* < 0.001**)	-0.138 (*p* = 0.083)	-0.106 (*p* = 0.182)	-0.164 (*p* = 0.039*)	-0.142 (*p* = 0.073)	-0.288 (*p* < 0.001**)
Latency end exploration (s)	-0.124 (*p* = 0.118)	-0.221 (*p* = 0.005**)	-0.157 (*p* = 0.048*)	-0.194 (*p* = 0.014*)	-0.175 (*p* = 0.027*)	-0.205 (*p* = 0.009**)	-0.302 (*p* < 0.001**)
Time on open arms (s)	0.175 (*p* = 0.027*)	0.336 (*p* < 0.001**)	0.169 (*p* = 0.033*)	0.155 (*p* = 0.050)	0.151 (*p* = 0.056)	0.184 (*p* = 0.020*)	0.295 (*p* < 0.001**)
Number of entries open arm	0.297 (*p* < 0.001**)	0.299 (*p* < 0.001**)	0.182 (*p* = 0.021*)	0.104 (*p* = 0.189)	0.141 (*p* = 0.076)	0.174 (*p* = 0.028*)	0.315 (*p* < 0.001**)

Subjective anxiety	-0.125 (p = 0.115)	-0.237 (p = 0.003**)	-0.082 (p = 0.303)	-0.068 (p = 0.390)	-0.160 (p = 0.044*)	-0.087 (p = 0.274)	-0.226 (p = 0.004**)

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Predictors and Spatial Patterns of EPM Behavior

Linear regression analyses revealed that the factor Fearlessness of the FPP predicted time on open arms ($p = 0.001$). Latency to explore the open arms was predicted by the factor Lack of empathy of the FPP ($p = 0.048$) and Fearlessness ($p < 0.001$), while latency to reach the end of the open arms was predicted by the factor Impulsivity of the FPP ($p = 0.049$). The number of entries into the open arms was predicted by Lack of empathy ($p < 0.001$) and Fearlessness ($p = 0.012$). Multivariate analyses showed that Fearlessness and Lack of empathy jointly contributed to all models (Table 3.3). Heatmaps depicting the time participants spent on the EPM in relation to the FPP sum score illustrate the behavioral patterns associated with psychopathic traits (Figure 3.6).

Table 3.3: Multivariable linear regression models for the association between psychopathic personality traits and elevated plus-maze/subjective anxiety. Table shows standardized regression coefficients with the associated *p*-values. Voulgaris et al., Sci Rep 14, 11832 (2024).

Predictor	Multivariate model (*p*-values)	Time on open arms (s)	Latency open arm exploration (s)	Latency open arm end exploration (s)	Number of entries open arm	Subjective anxiety (not included in multivariate model)
Lack of empathy	0.029	0.14 (*p* = 0.104)	-0.16 (*p* = 0.048)	-0.12 (*p* = 0.153)	0.27 (*p* = 0.001)	0.02 (*p* = 0.839)
Fearlessness	0.004	0.28 (*p* = 0.001)	-0.31 (*p* < 0.001)	-0.14 (*p* = 0.110)	0.21 (*p* = 0.012)	-0.21 (*p* = 0.016)
Narcissistic egocentrism	0.987	-0.02 (*p* = 0.863)	0.04 (*p* = 0.648)	0.04 (*p* = 0.729)	-0.01 (*p* = 0.892)	0.09 (*p* = 0.379)
Impulsivity	0.374	0.14 (*p* = 0.110)	-0.11 (*p* = 0.197)	-0.17 (*p* = 0.049)	0.11 (*p* = 0.185)	0.08 (*p* = 0.362)
Social manipulation	0.831	0.04 (*p* = 0.696)	-0.09 (*p* = 0.352)	-0.09 (*p* = 0.364)	0.06 (*p* = 0.520)	-0.09 (*p* = 0.340)
Power	0.912	0.00 (*p* = 0.929)	0.05 (*p* = 0.643)	-0.05 (*p* = 0.649)	0.00 (*p* = 0.999)	-0.05 (*p* = 0.636)
Model fit: R^2		0.14	0.16	0.10	0.16	0.07
Sample size	160	160	160	160	160	160

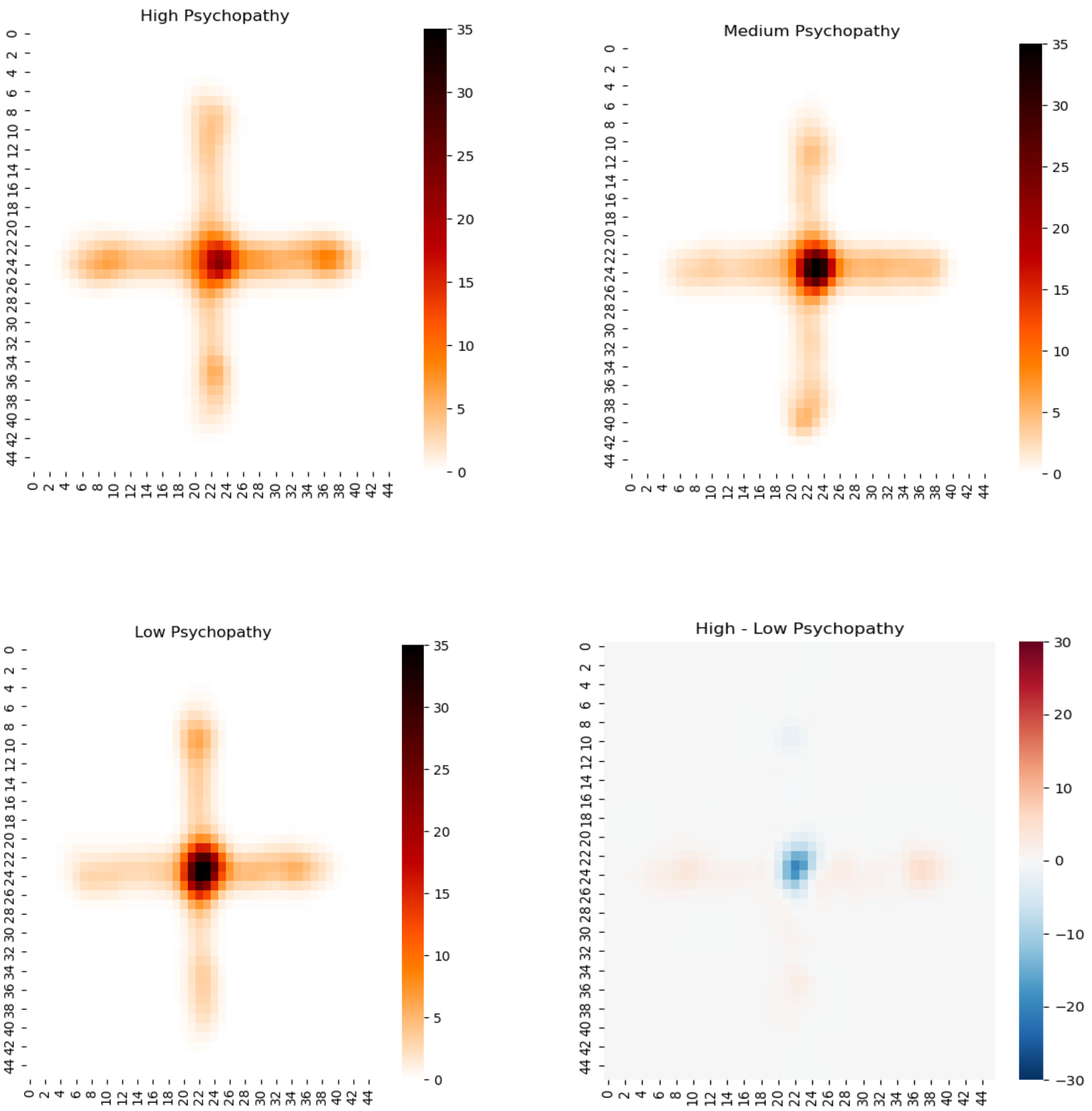


Figure 3.6: Heatmaps reveal time spent on the EPM in relation to the FPP sum score. The three groups of low, medium and high psychopathy were constructed using a quartile split into low, medium and high FPP. Sociodemographic characteristics of the groups were Low FPP: 74.5% women, 25.5% men, age: Mean= 26.1 years, *SD* = 5.1 years; Medium FPP: 73.3% women, 26.7 men, age: Mean = 25.4 years, *SD* = 5.8 years; High FPP: 60.0% women, 40.0 men, age: Mean = 24.6 years, *SD* = 5.6 years. Voulgaris et al., Sci Rep 14, 11832 (2024).

Study 3: The Effect of Gaming Experience on Behavioral Responses on the Virtual EPM

Effects of Gaming Experience and Gender on EPM Behavior

Robust linear regression models including gender, gaming experience, and their interaction accounted for 15% of the variance in latency to open arms exploration ($F(3, 423) = 22.74$, $p < 0.001$), 17.1% in open arm entries ($F(3, 423) = 28.99$, $p < 0.001$), 18% in time spent on open arms ($F(3, 423) = 30.25$, $p < 0.001$), and 18.6% in latency to end exploration ($F(3, 423) = 32.65$, $p < 0.001$) (Table 3.4). The interaction of gender and gaming experience did not reach significance for three of the four behavioral parameters (all $p > 0.05$), reaching significance only for latency to open arm exploration ($p = 0.012$).

Table 3.4: Linear regression analyses with robust standard errors for behavioral parameters. Predictors were gender (man vs. woman), gaming experience (yes vs. no), and their interaction. Values represent regression coefficients (*B*), 95% confidence intervals (CI), and *p*-values in the format *B* [95% CI], *p*. Reference categories: women (gender), no (gaming experience). Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; . $p \leq 0.1$.

Predictor	Time on open arms (s)	Entries on open arms (n)	Latency open arm exploration (s)	Latency end exploration (s)
Gender	61.57 [34.40, 88.74], $p < 0.001$***	2.35 [1.35, 3.35], $p < 0.001$***	-89.50 [-130.31, -48.70], $p < 0.001$***	-115.84 [-161.43, -70.24], $p < 0.001$***
Gaming experience	25.91 [12.51, 39.30], $p < 0.001$***	1.19 [0.58, 1.81], $p < 0.001$***	-68.83 [-95.16, -42.49], $p < 0.001$***	-42.59 [-67.57, -17.62], $p < 0.001$***
Gender × Gaming experience	-27.97 [-58.92, 2.98], $p = 0.076$.	-1.01 [-2.19, 0.18], $p = 0.095$.	60.51 [13.25, 107.77], $p = 0.012$*	47.96 [-6.15, 102.06], $p = 0.082$.
R^2	0.180	0.171	0.150	0.186

Across all behavioral measures, Gender showed a significant main effect: man spent more time on the open arms, made more open arm entries, and had shorter latencies for both open arms exploration and end exploration compared to women (all $p < 0.001$). Gaming experience also showed a significant main effect ($p < 0.001$), with post hoc inspection of marginal means indicating that this effect was primarily attributable to women, who exhibited greater exploration when they had gaming experience (Figure 3.7). In contrast, stratified analysis of gaming experience showed effect sizes for gaming experience in men showed no substantial effect (Table 3.5).

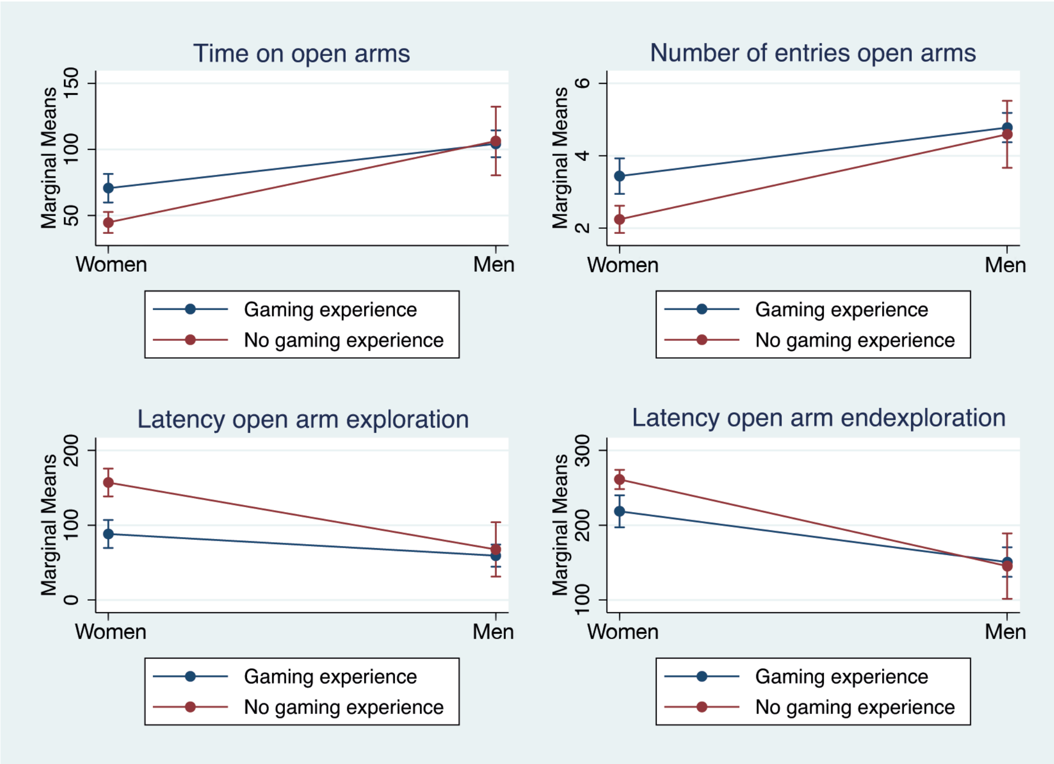


Figure 3.7: Behavioral parameters for man and women with (blue) and without (red) gaming experience (ever played at least weekly). Values are given as marginal means with 95% confidence intervals.

Table 3.5: Effect sizes in stratified analyses of gaming experience.

Behavioral measures	Subgroup	Standardized regression coefficients
Time on open arms (s)	women	0.23
	men	-0.01
Number of entries on open arms	women	0.24
	men	0.03

Latency open arm exploration (s)	women	-0.29
	men	-0.03
Latency end exploration (s)	women	-0.22
	men	0.02

Heatmap visualizations (Figure 3.8) further illustrate that gaming experience was associated with greater exploration of the open arms in women but not in men. These behavioral patterns are consistent with the main effect of gaming experience being largely driven by the women subgroup.

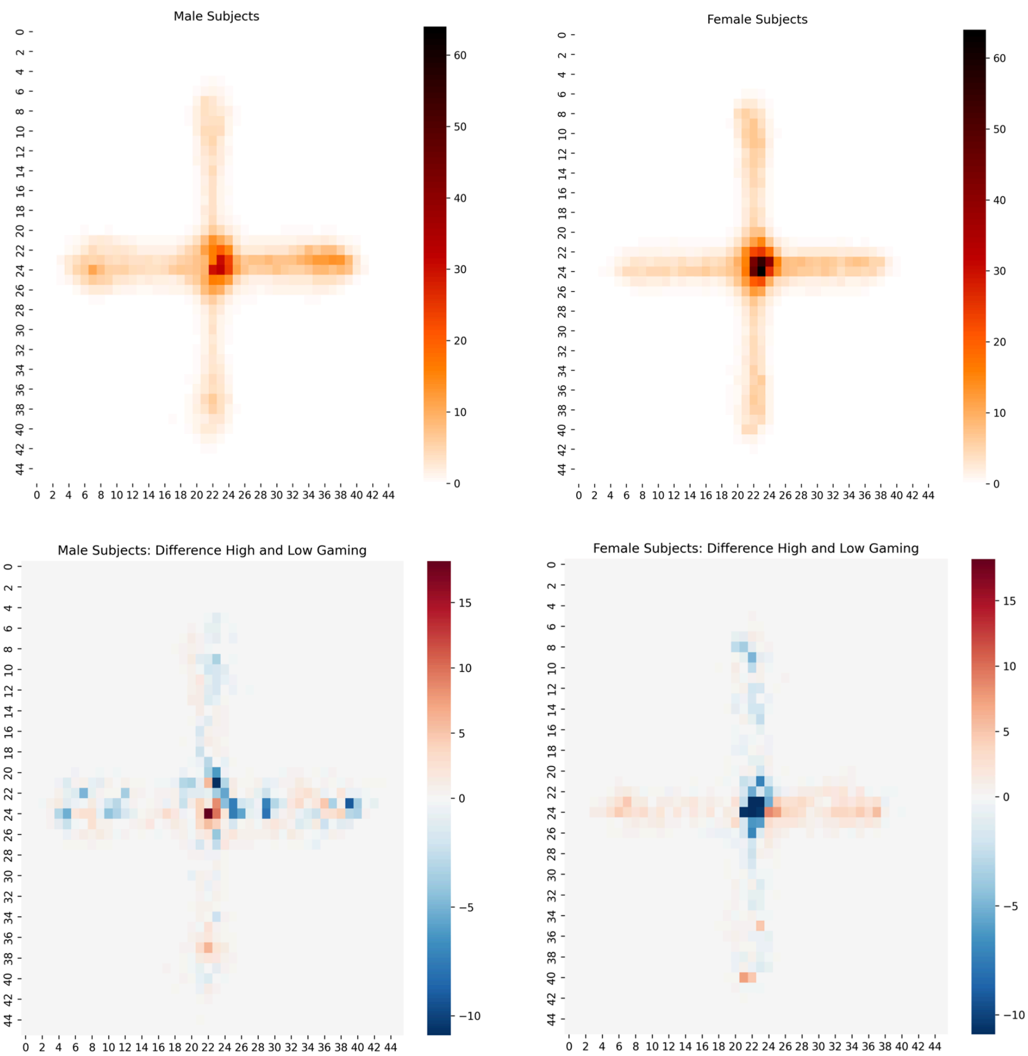


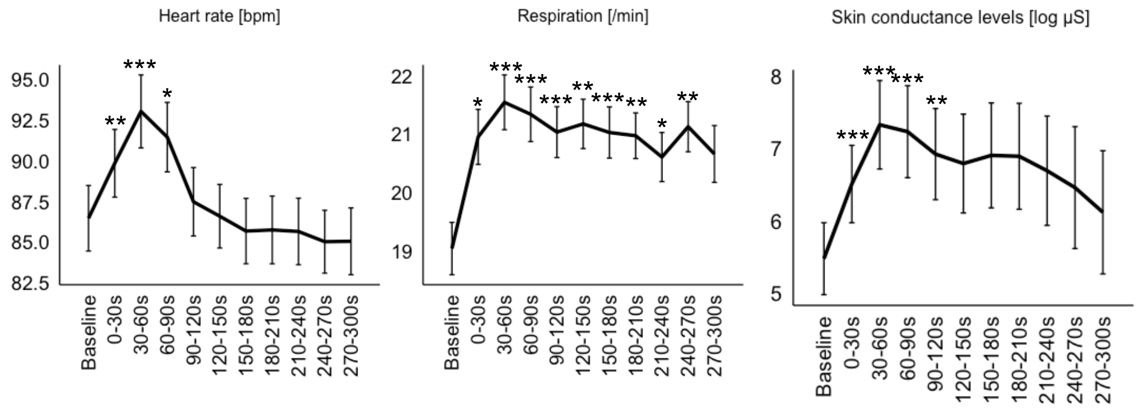
Figure 3.8: Heatmaps illustrate that women spent more time in the center of the EPM (upper panels), while gaming experience was associated with more exploration of the EPM arms in women but not men (lower panels).

Study 4: Validation of the Dark-Light Box in Virtual Reality

Healthy Sample

Physiological and Endocrinological Responses in the DLB

Behavioral testing in the DLB induced a physiological activation with a 7.6% increase in heart rate, a 13.2% increase in respiratory rate, and a 33.8% increase in skin conductance levels when comparing averages of baseline levels with those from the first minute (30-60s interval) of exposure to the DLB. A repeated-measures ANOVA with a Greenhouse-Geisser correction confirmed significant increases in heart rate ($F(4.77, 276.41) = 14.01$, $p < 0.001$, $\eta^2 partial = 0.20$), respiratory rate ($F(6.39, 370.82) = 6.50$, $p < 0.001$, $\eta^2 partial = 0.10$), and skin conductance levels ($F(2.39, 128.81) = 5.67$, $p < 0.01$, $\eta^2 partial = 0.10$). Post hoc tests showed significant increases between baseline levels and follow-up time points during the first minute of DLB exposure (Figure 3.9).



DLB exposure activated the sympathetic stress system, reflected in a main effect of time on saliva alpha-amylase levels ($F(2, 114) = 19.22$, $p < 0.001$, $\eta^2 partial = 0.25$). Alpha-amylase activity showed a significant increase of 11.45% directly after VR exposure (T1) compared to baseline (T0) ($p < 0.01$). This was followed by a decrease of 23.76% between levels directly after exposure (T1) and 15 minutes post-exposure (T2) ($p < 0.001$) (Table 3.6, Figure 3.10). In

Figure 3.9: Physiological reactions to the DLB show significant increase in heart rate ($n = 59$) respiratory rate ($n = 59$) and skin conductance levels ($n = 55$) compared to baseline. Values are given as means ± standard error. Significance levels are indicated for comparison with baseline: $p \leq 0.05$= *, $p \leq 0.01$=**, $p \leq 0.001$= ***.

contrast, DLB exposure did not activate the HPA axis activity: cortisol only showed a descriptive increase of 5.44% between baseline and 15 min after exposure (T2) that did not reach significance.

Table 3.6: Descriptive statistics of endocrinological measurements.
T0: baseline, T1: directly after VR exposure, T2: 15 min after VR exposure.

Measure	**Time point**	**Mean**	***SD***	***N***
Log Cortisol	T0	0.64	0.29	60
	T1	0.66	0.30	60
	T2	0.67	0.29	60
Log α Amylase	T0	2.09	0.28	58
	T1	2.16	0.25	58
	T2	2.03	0.27	58

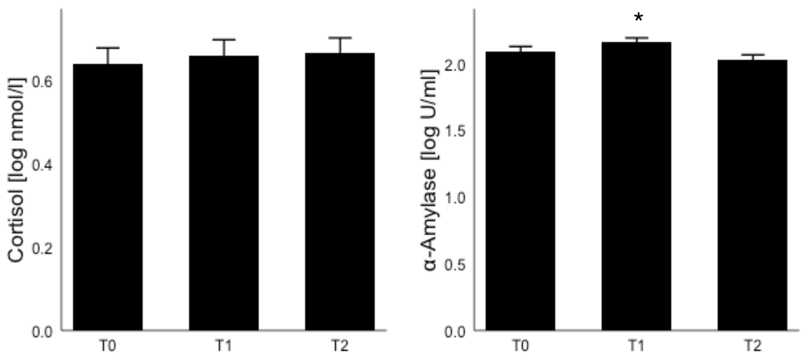


Figure 3.10: means for saliva cortisol ($n = 60$) and alpha-amylase levels ($n = 58$) before (T0) and directly after (T1) as well as at 15 min (T2) after exposure to the DLB. Values are given as means ± standard error. Significance levels are indicated for comparison with baseline: $p \leq 0.05$= *, $p \leq 0.01$=**, $p \leq 0.001$= ***.

Subjective Experiences in the DLB

Participants (n = 74) reported a strong sense of presence in the virtual environment, as indicated by average scores on the iGroup Presence Questionnaire (spatial presence: 14.6 ± 3.7; involvement: 5.9 ± 4.3; experienced realism: 8.7 ± 3.5). Symptoms of simulator sickness were also observed with the Simulator Sickness Questionnaire (total score: 27.6 ± 21.1; nausea: 17.2 ± 18.1; oculomotor symptoms: 24.0 ± 18.7; disorientation: 33.7 ± 29.0).

A repeated-measures ANOVA indicated that the position in the DLB setup had a significant effect on participants' specific anxiety ratings ($F(2.52, 98) = 41$, $p < 0.001$, $\eta^2 partial = 0.51$). The mean subjective anxiety ratings for each position are presented in Figure 3.11. Post hoc comparisons revealed a gradual increase in subjective anxiety from brighter to darker areas. There was a significant increase in subjective anxiety between the starting area and all three positions in the second area (minimum $p < 0.01$). Furthermore, in the second area, subjective anxiety ratings were significantly higher at the two outer positions (left and right wall) compared to the first area (minimum $p < 0.05$), while no significant difference was observed for the middle positions of both areas. The darker areas (3 and 4) had significantly higher subjective anxiety ratings compared to all brighter areas ($p < 0.001$). Additionally, the darkest area (4) showed significantly higher ratings than the third area across all three positions (minimum $p < 0.05$).

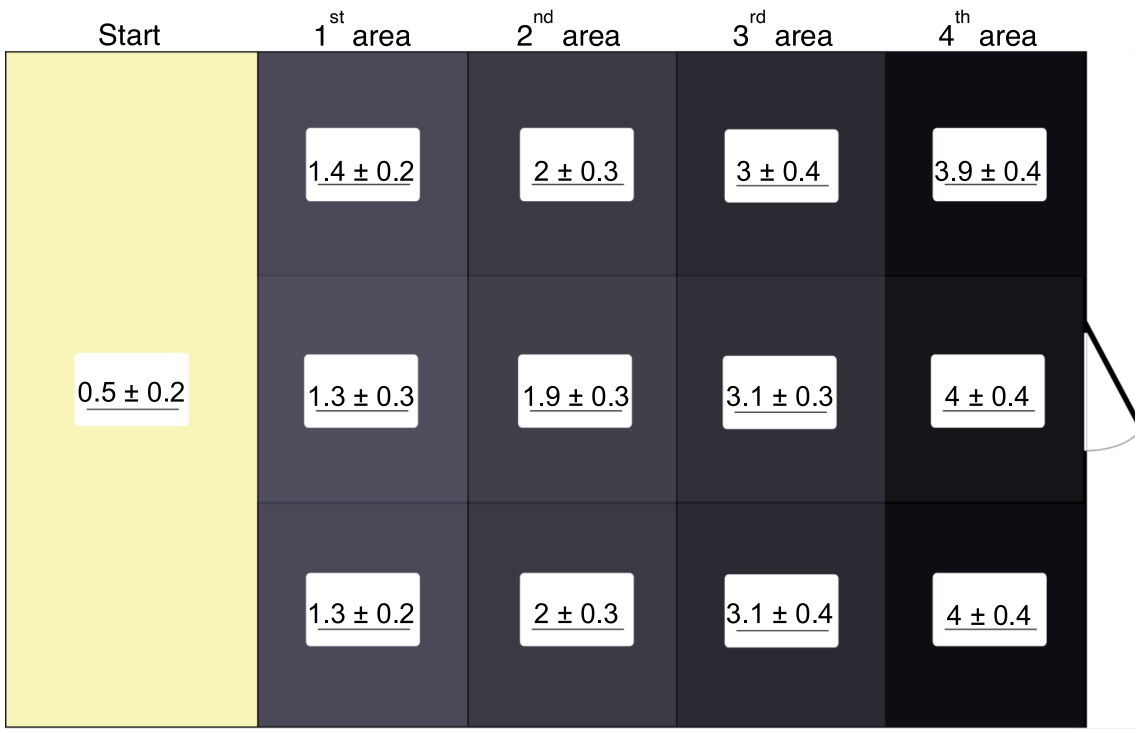


Figure 3.11: Subjective anxiety ratings for specific positions in the DLB ($N = 40$).The lighting conditions are represented in grayscale, while the lightest area is on the left (start) and the darkest on the right (4th area) of the DLB. Values are given as means ± standard error.

Behavioral Measures in the DLB

A Pearson correlation matrix showed minimal ($r = 0.10$-0.20) to moderate ($r = 0.30$-0.40) associations between DLB behavioral variables and questionnaire scores and subjective anxiety ratings (Figure 3.12). Behavioral parameters such as looking up, time in dark, average velocity, total distance, and lower body behaviors showed mainly negative relationships with most of the questionnaires and subjective anxiety ratings, whereas both latencies and arms before body behaviors were mostly positive correlated with questionnaires and subjective anxiety ratings. To avoid redundancy and focus on variables with clearer linear relations, the following were selected for further analysis: Looking up frequency, time in dark (s), total distance (m), log latency first visit (s), log re-examination frequency, log arms before body duration, log lower body behaviors duration.

	ACQ	STAI-T	SSSV	IUS	FOD	CLQ	STAI-S	Subjective Anxiety	N
exploration frequency (n)	-0.017	-0.003	0.161	0.159	0.096	-0.025	-0.022	-0.022	74
exploration duration (s)	-0.061	-0.047	0.188	0.101	0.049	0.015	-0.081	0.050	74
looking up frequency (n)	-0.108	-0.246	0.231	-0.125	-0.265	-0.114	-0.126	-0.188	74
looking up duration (s)	-0.007	0.047	0.163	-0.016	-0.191	0.000	-0.012	-0.249	74
time in dark (s)	-0.178	-0.224	0.258	-0.129	-0.242	-0.153	-0.193	-0.266	74
average velocity (m/s)	-0.169	-0.185	0.297	-0.155	-0.012	-0.050	0.049	-0.227	74
total distance (m)	-0.168	-0.184	0.297	-0.153	-0.011	-0.049	0.050	-0.227	74
log latency of first visit (s)	0.184	0.321	-0.228	0.154	0.208	-0.004	0.096	0.080	74
log latency of endexploration (s)	0.048	0.140	-0.256	0.062	0.145	0.060	0.031	0.134	74
log reexamination (n)	-0.018	-0.065	0.391	-0.033	0.024	-0.177	-0.001	-0.127	74
log reexamination duration (s)	-0.005	-0.057	0.363	-0.032	-0.013	-0.199	-0.027	-0.159	74
log arms before body frequency (n)	0.020	-0.049	0.075	0.123	0.118	0.177	0.019	0.326	74
log arms before body duration (s)	0.017	-0.045	0.072	0.132	0.169	0.206	0.169	0.432	74
log lower body behaviors frequencey (n)	-0.090	-0.038	0.317	-0.160	-0.132	-0.001	-0.045	-0.066	74
log lower body behaviors duration (s)	-0.121	-0.050	0.288	-0.167	-0.190	-0.095	-0.128	-0.107	74

Figure 3.12: Pearson correlations between behavioral variables and questionnaire sum scores. The intensity of the correlation is indicated through a color gradient where a correlation coefficient of 1 is depicted in deep red, 0 in white, and -1 in deep purple. Abbreviations: ACQ (Agoraphobic Cognitions Questionnaire), STAI-T (State-Trait Anxiety Inventory: trait), SSSV (Sensation Seeking Scale Form V), IUS (Intolerance of Uncertainty Scale), FOD (Fear of Darkness Questionnaire), CLQ (Claustrophobia Questionnaire), STAI-S (State-Trait Anxiety Inventory: state).

Multivariate multiple regression showed significant overall effects of sensation seeking on all behavioral parameters measured in the DLB (Pillai's Trace = 0.391, $F(7, 59) = 5.41$, $p < 0.001$, $\eta^2 partial = 0.39$) and subjective anxiety (Pillai's Trace = 0.232, $F(7, 59) = 2.55$, $p = 0.023$, $\eta^2 partial = 0.23$). Other predictors were not significant (Table 3.7).

Table 3.7: Multivariate tests of the effects of questionnaire sum scores and subjective anxiety on behavioral variables. Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; . $p \leq 0.1$. ACQ (Agoraphobic Cognitions Questionnaire), STAI-T (State-Trait Anxiety Inventory: trait), SSSV (Sensation Seeking Scale Form V), IUS (Intolerance of Uncertainty Scale), FOD (Fear of Darkness Questionnaire), CLQ (Claustrophobia Questionnaire), STAI-S (State-Trait Anxiety Inventory: state).

Effect	Pillai's Trace	*F*	Hypothesis *df*	Error *df*	*p*	Partial Eta-Squared
Constant	0.433	6.443[b]	7	59	<0.001	0.433
ACQ	0.092	0.851[b]	7	59	0.551	0.092
STAI-T	0.200	2.103[b]	7	59	0.057.	0.200
SSSV	0.391	5.409[b]	7	59	<0.001***	0.391
IUS	0.057	0.512[b]	7	59	0.822	0.057
FOD	0.119	1.141[b]	7	59	0.351	0.119
CLQ	0.166	1.680[b]	7	59	0.131	0.166
STAI-S	0.083	0.767[b]	7	59	0.617	0.083
Subjective Anxiety	0.232	2.553[b]	7	59	0.023*	0.232

a. Design: Constant + ACQ + STAI-T + SSSV + IUS + FOD + CLQ + STAI-S + subjective anxiety

b. Exact Statistic

The *t*-tests for the coefficients with robust standard errors revealed that sensation seeking was positively associated with total distance (Estimate = 1.21, *SE* = 0.38, $t = 3.18$, corrected $p = 0.046$, $sr^2 = 0.13$) and was positively associated with log-transformed re-examination frequency (Estimate = 0.02, *SE* = 0.01, $t = 3.69$, corrected $p = 0.026$, $sr^2 = 0.17$). Subjective Anxiety had a positive association with log-transformed arms-before-body behavior (Estimate = 0.18, *SE* = 0.06, $t = 3.15$, corrected $p = 0.046$, $sr^2 = 0.13$). None of the remaining behavioral variables were significantly affected by the questionnaire sum scores or subjective anxiety (Table 3.8, for full results see supplementary Table S5).

Table 3.8: Summary of *t*-tests of coefficients for the effects of questionnaire sum scores and subjective anxiety on behavioral variables. Values represent unstandardized estimates with adjusted *p*-values (Benjamini-Hochberg correction). Significance codes: *** p <0.001; ** p < 0.01; * p < 0.05; . $p \leq 0.1$. Abbreviations: ACQ (Agoraphobic Cognitions Questionnaire), STAI-T (State-Trait Anxiety Inventory: trait), SSSV (Sensation Seeking Scale Form V), IUS (Intolerance of Uncertainty Scale), FOD (Fear of Darkness Questionnaire), CLQ (Claustrophobia Questionnaire), STAI-S (State-Trait Anxiety Inventory: state). Additional statistics (SE, *t*-values, sr^2) are reported in the Supplementary Material Table S5.

Predictor	Looking up (n)	Time in dark (s)	Total distance (m)	Log First Visit (s)	Log Reexamination (n)	Log arms before body (s)	Log lower body behaviors (s)
ACQ	-0.11 (*p* = 0.765)	-2.25 (*p* = 0.503)	-0.61 (*p* = 0.470)	0.02 (*p* = 0.238)	0.01 (*p* = 0.549)	-0.02 (*p* = 0.585)	-0.01 (*p* = 0.549)
STAI-T	-0.25 (*p* = 0.238)	-2.01 (*p* = 0.336)	-0.71 (*p* = 0.201)	0.02 (*p* = 0.077).	-0.01 (*p* = 0.408)	-0.03 (*p* = 0.238)	0.01 (*p* = 0.549)
SSSV	0.36 (*p* = 0.238)	3.96 (*p* = 0.178)	1.21 (*p* = 0.046*)	-0.02 (*p* = 0.169)	0.02 (*p* = 0.026*)	0.03 (*p* = 0.396)	0.04 (*p* = 0.134)
IUS	0.07 (*p* = 0.549)	0.80 (*p* = 0.470)	0.02 (*p* = 0.945)	-0.00 (*p* = 0.582)	0.00 (*p* = 0.549)	0.01 (*p* = 0.549)	-0.01 (*p* = 0.549)
FOD	-0.32 (*p* = 0.277)	-2.30 (*p* = 0.408)	0.04 (*p* = 0.968)	0.02 (*p* = 0.238)	0.01 (*p* = 0.549)	0.02 (*p* = 0.652)	-0.02 (*p* = 0.391)
CLQ	0.09 (*p* = 0.468)	0.95 (*p* = 0.408)	0.33 (*p* = 0.336)	-0.01 (*p* = 0.077).	-0.01 (*p* = 0.336)	0.00 (*p* = 0.945)	0.01 (*p* = 0.468)
STAI-S	0.10 (*p* = 0.652)	-0.04 (*p* = 0.978)	0.65 (*p* = 0.255)	-0.01 (*p* = 0.549)	0.00 (*p* = 0.790)	0.01 (*p* = 0.765)	-0.00 (*p* = 0.789)
Subjective Anxiety	-0.67 (*p* = 0.336)	-8.93 (*p* = 0.238)	-2.83 (*p* = 0.077.)	0.02 (*p* = 0.549)	-0.00 (*p* = 0.725)	0.18 (*p* = 0.046*)	-0.01 (*p* = 0.931)

Welch's *t*-tests on the seven preselected behavioral variables with Benjamini-Hochberg correction across comparisons revealed that HA participants showed significantly longer arms-before-body duration than LA participants (adjusted $p = 0.042$, $d = -0.68$, 95% CI [-1.16, -0.19]).

None of the other six behavioral variables showed significant group differences after correction (all adjusted $p > .08$). (Table 3.9, Figure 3.13). An additional exploratory analysis was performed using looking-up duration instead of frequency. This revealed that HA participants spent less time looking up than LA participants (M_{HA} = 37.1, *SD* = 26.5 vs. M_{LA} = 51.9, *SD* = 29.8), $t(71.2) = 2.25$, $p = 0.027$, $d = 0.52$, 95% CI [0.06, 0.99]. However, as this variable was not corrected for multiple testing; if included, the effect would not survive Benjamini-Hochberg adjustment (see supplementary Figure S3).

Table 3.9: Welch's t-tests comparing behavioral measures in the virtual DLB between low-anxiety (LA; *n* = 41) and high-anxiety (HA; *n* = 33) participants, classified using a median split of subjective anxiety ratings (1-9, median = 3). Values are reported as group means with standard deviations (SD). Adjusted *p*-values are Benjamini-Hochberg corrected. Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; . $p \leq 0.1$.

Variable	Mean (*SD*) LA	Mean (*SD*) HA	*t*	*df*	*p*	Adjusted *p*	*d* (95% CI)
Log arms before body (s)	0.64 (0.84)	1.31 (1.13)	-2.86	57.75	0.006	0.042*	-0.68 (-1.16, -0.19)
Time spent in dark (s)	203.17 (68.32)	159.36 (89.79)	2.32	58.6	0.024	0.085	0.55 (0.07, 1.02)
Total distance (m)	61.04 (21.24)	51.70 (20.34)	1.92	69.80	0.058	0.136	0.45 (-0.02, 0.91)
Log reexami-nation (n)	0.38 (0.3)	0.3 (0.31)	1.18	67.72	0.24	0.357	0.28 (-0.19, 0.74
Looking up (n)	15.73 (8.13)	13.58 (7.95)	1.15	69.26	0.255	0.357	0.27 (-0.19, 0.73)
Log latency first visit (s)	1.55 (0.4)	1.62 (0.48)	-0.67	62.13	0.507	0.554	-0.16 (-0.62, 0,31)

Log lower body behaviors (s)	0.9 (0.66)	0.8 (0.72)	0.6	66.02	0.554	0.554	0.14 (-0.32, 0.6)

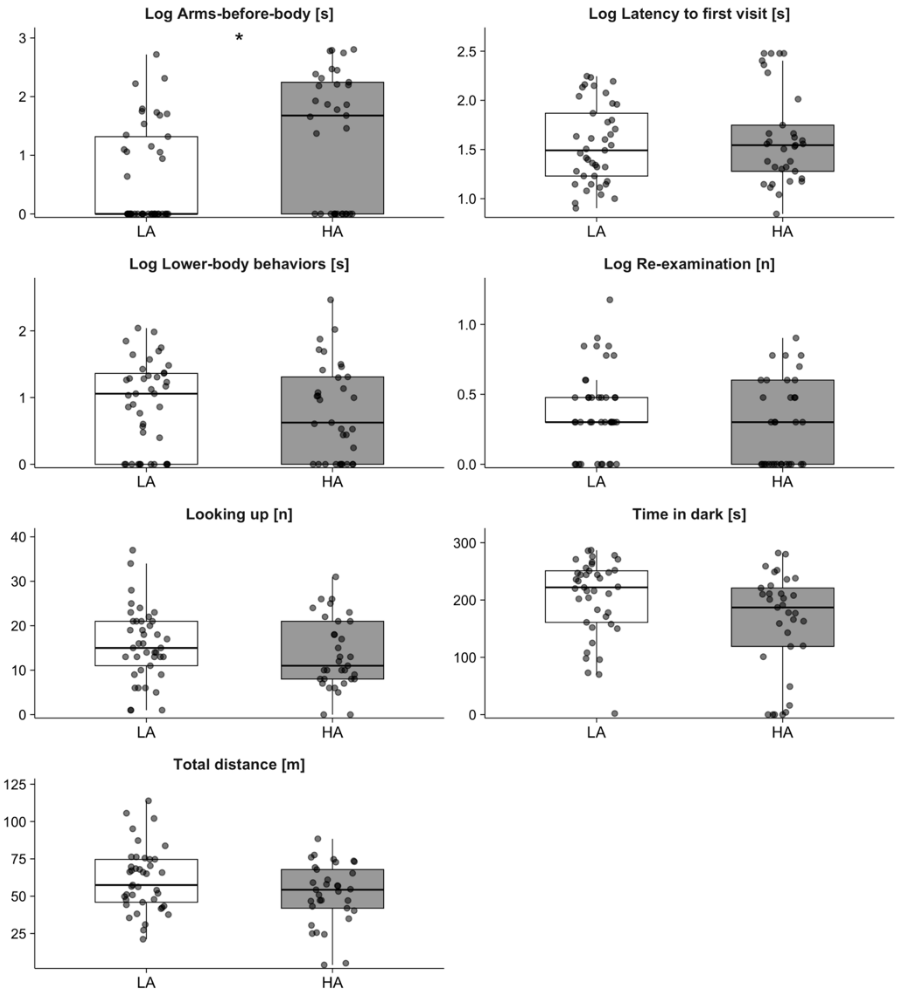


Figure 3.13: Group comparisons of behavioral measures in the virtual DLB between low-anxiety (LA; *n* = 41) and high-anxiety (HA; *n* = 33) participants, classified using a median split of subjective anxiety ratings (1-9, median = 3). Boxplots show medians, interquartile ranges, and individual data points. Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; . $p \leq 0.1$.

Pearson correlations were computed to examine the relationships between core behavioral variables of the EPM and the DLB (Figure 3.14). Most correlations were minimal (r = 0.10-0.20) to moderate (r = 0.30-0.40), with no strong correlations (r > 0.50) observed. DLB parameters such as time in dark, looking up, total distance, and re-examination were positively correlated with EPM measures including time on open arms, total distance, and number of open arm entries. Latencies in both tests were negatively correlated with most other behavioral variables but positively correlated with each other. Log-transformed arms-before-body duration in the DLB showed a similar but weaker pattern to the latencies, whereas log-transformed lower body behaviors duration had the weakest correlations with the EPM parameters (r = 0.00-0.10).

EPM / DLB	Time on open arm (s)	Total distance (EPM)	Number of entries open arm	Latency open arm first visit (s)	Latency open arm end explored (s)	N
looking up freqency (n)	0.279	0.314	0.138	-0.270	-0.317	71
time in dark (s)	0.285	0.364	0.294	-0.432	-0.189	71
total distance (m)	0.320	0.447	0.389	-0.440	-0.152	71
log latency first visit (s)	-0.263	-0.306	-0.225	0.318	0.153	71
log reexamination frequency (n)	0.426	0.301	0.381	-0.345	-0.300	71
log arms before body duration (s)	-0.114	-0.181	-0.177	0.141	0.152	71
log lower body behaviors duration (s)	0.063	0.027	0.054	0.043	-0.107	71

Figure 3.14: Pearson correlations between core behavioral variables of the EPM (horizontal) and selected behavioral variables of the DLB (vertical). The intensity of the correlation is indicated through a color gradient where a correlation coefficient of 1 is depicted in deep red, 0 in white, and -1 in deep

Clinical Sample

Transdignostic Behavioral Differences Between Patients and Controls

Paired samples *t*-tests were conducted to compare patients (n = 40) and individually matched controls (n = 40) across the behavioral variables. No statistically significant group differences were found either before or after *p*-values were adjusted for multiple comparisons using the Benjamini-Hochberg correction (Table 3.10). The largest uncorrected effect was observed for log re-examination frequency ($t(39) = -1.78$, $p = 0.082$, Cohen's $dz = -0.282$), which did not reach statistical significance. All other variables showed small effect sizes (Cohen's dz = -0.070 to 0.246) and non-significant *p*-values. A post hoc power analysis based on the largest observed effect size (*dz* = 0.282) indicated that the current sample size of 40 matched pairs provided approximately 41% power to detect an effect of this magnitude at $\alpha = 0.05$. Achieving the conventional 80% power threshold would have required approximately 100 matched pairs.

Table 3.10: Results of paired *t*-tests of patients ($n = 40$) and matched controls ($n = 40$) across behavioral variables and post hoc power estimates. Adjusted *p*-values are Benjamini-Hochberg corrected.

Variable	*t*	*df*	*p*	95% CI (Lower, Upper)	Mean (Control)	Mean (Patient)	Adjusted *p*	*dz*	Post hoc power
looking up (n)	-1.57	39	0.125	(-6.47, 0.82)	14.20	11.38	0.299	-0.248	0.334
time in dark (s)	0.84	39	0.406	(-23.32, 56.47)	181.65	198.23	0.568	0.133	0.130
total distance (m)	-1.07	39	0.292	(-16.13, 4.99)	58.93	53.35	0.512	-0.169	0.181
log first visit (s)	-0.67	39	0.506	(-0.27, 0.13)	1.64	1.57	0.590	-0.106	0.100
log reexamination (n)	-1.78	39	0.082	(-0.23, 0.01)	0.30	0.19	0.299	-0.282	0.413
log arms before body (s)	1.56	39	0.128	(-0.10, 0.78)	0.94	1.28	0.299	0.246	0.329
log lower body behaviors (s)	-0.44	39	0.662	(-0.33, 0.21)	0.78	0.72	0.662	-0.070	0.072

Exploratory Subgroup Analyses of AD/PTSD and BPD Patients

In exploratory subgroup analyses, behavioral differences were examined between patients in the AD/PTSD and BPD diagnostic groups and their individually matched controls. No correction for multiple comparisons was applied. *P*-values are therefore uncorrected and should be interpreted with caution. In the AD/PTSD group, patients covered less total distance than controls ($t(19) = -2.14$, $p = 0.045$, $dz = -0.74$), with additional non-significant trends toward reduced looking-up frequency ($t(19) = -1.79$, $p = 0.090$, $dz = -0.46$) and re-examination frequency ($t(19) = -1.66$, $p = 0.114$, $dz = -0.54$). Other variables, including time in dark areas, latency to first visit, and lower-body movement durations, showed minimal differences ($dz < 0.25$). In the BPD group, patients spent more time in dark areas than controls ($t(19) = 2.10$, $p = 0.050$, $dz = 0.73$). Small to moderate, non-significant differences were observed for greater arms-before-body duration ($t(19) = 1.18$, $p = 0.251$, $dz = 0.42$) and greater total distance ($t(19) = 1.07$, $p = 0.297$, $dz = 0.36$). Lower-body behaviors were slightly reduced ($t(19) = -1.15$, $p = 0.266$, $dz = -0.39$), with a non-significant difference toward increased latency to first visit ($t(19) = -1.09$, $p = 0.289$, $dz = -0.36$) and lower looking-up and re-examination frequencies (dz = -0.26 and -0.20, respectively) (Table 3.11).

Table 3.11: Results of exploratory paired samples t-tests comparing patients with AD/PTSD and BPD diagnoses to their matched controls across behavioral variables.

Variable	Diagnostic Group	*t*	*df*	*p*	95% CI (Lower, Upper)	*dz*
log first visit (s)	AD/PTSD	-0.01	19	0.996	[-0.32, 0.31]	-0.0
time in dark (s)	AD/PTSD	-0.58	19	0.571	[-78.01, 44.31]	-0.19
total distance (m)	AD/PTSD	-2.14	19	0.045	[-35.78, -0.41]	-0.74
log arms before body (s)	AD/PTSD	0.67	19	0.512	[-0.5, 0.96]	0.22
log lower body behaviors (s)	AD/PTSD	0.71	19	0.488	[-0.28, 0.57]	0.21
log reexamination (n)	AD/PTSD	-1.66	19	0.114	[-0.34, 0.04]	-0.54
looking up (n)	AD/PTSD	-1.79	19	0.09	[-7.38, 0.58]	-0.46
log first visit (s)	BPD	-1.09	19	0.289	[-0.39, 0.12]	-0.36
time in dark (s)	BPD	2.1	19	0.05	[0.09, 99.91]	0.73
total distance (m)	BPD	1.07	19	0.297	[-6.61, 20.51]	0.36
log arms before body (s)	BPD	1.18	19	0.251	[-0.34, 1.24]	0.42
log lower body behaviors (s)	BPD	-1.15	19	0.266	[-0.74, 0.22]	-0.39

log reexamination (n)	BPD	-0.62	19	0.542	[-0.28, 0.15]	-0.2
looking up (n)	BPD	-0.87	19	0.393	[-7.63, 3.13]	-0.26

Chapter 4 Discussion

The discussion is organized around overarching themes that reflect central aspects of the research questions. Within each thematic section, findings from the individual studies are considered and integrated to provide both a focused examination of each study's contributions and a broader synthesis across the dissertation.

Summary of Main Findings

This dissertation sets out to evaluate and extend the translational potential of virtual reality (VR) adaptations of classical animal paradigms of anxiety, namely the Elevated Plus Maze (EPM) and the Dark-Light Box (DLB), for the use in human participants. The virtual EPM demonstrated test-retest reliability of behavioral measures, indicating temporal stability of approach-avoidance behavior over repeated exposures. Psychopathic traits were associated with increased approach tendencies and reduced avoidance on the EPM, suggesting that the paradigm is also sensitive to anxiety-related processes of other psychological conditions than anxiety disorders. Gaming experience influenced behavioral responses in women, pointing to the role of experiential factors in behavioral differences in VR-based anxiety assessment. Finally, the virtual DLB elicited clear subjective and physiological anxiety-related responses, alongside a general effect of subjective anxiety and sensation seeking on behavior. Within the behavioral domain, clear effects were found in the ethological marker arms-before-body, whereas classic avoidance measures were less robust. Thus, while its validity on the behavioral level was only partly supported, the DLB nevertheless reached proof-of-concept status and may provide complementary information to the EPM by capturing partially distinct facets of anxiety-related behavior. Taken together, these findings support the use of VR as a standardized and ecologically valid tool for assessing anxiety-related behavior in humans, while also emphasizing the impact of stable traits and prior experience on observed behavioral patterns.

Reliability and Validity of VR-Based Anxiety Paradigms

The present set of studies provides strong evidence that VR-based adaptations of classical anxiety paradigms can yield reliable and valid measures of human anxiety-related behavior. **Study 1** demonstrated high correlations between the same behavioral measures and across different measures comparing first and second testing showing temporal stability across a 28-day interval. This stability was further supported by consistent increases in respiratory frequency and salivary α-amylase, reflecting autonomic activation across exposures. Importantly, no indication of habituation in core behavioral measures was observed, which supports the interpretation of these responses as stable rather than subject to simple response decrement. This is in line with the assumption that strong aversive stimuli typically do not show substantial response decrement after a single repetition with interval between exposures (Rankin et al., 2009). The absence of habitation effects also distinguishes the virtual EPM from other paradigms eliciting stress responses, such as the Trier Social Stress Test, where subjective ratings often habituate across repeated testing (Boesch et al., 2014; S. Cohen et al., 2000; Jönsson et al., 2010), while autonomic markers do not consistently do so (Gerra et al., 2001; Jönsson et al., 2010).

At the same time, autonomic and subjective markers such as heart rate, skin conductance, and self-reported anxiety showed altered patterns at retest, likely reflecting anticipatory anxiety or sensitization processes. At second testing, heart rate was already elevated during the baseline period and no longer exhibited a distinct rise at task onset, whereas skin conductance levels and subjective anxiety ratings were generally higher across the second exposure. The anticipation of aversive situations involving novelty and unpredictability has been linked to catecholamine release (Mason, 1968) and increased skin conductance levels (Epstein & Roupenian, 1970). Physiological markers with high temporal resolution, such as heart rate and skin conductance, are particularly sensitive to anticipatory anxiety (Armario et al., 2020; Boesch et al., 2014). In Study 1, participants were aware that the second exposure would again involve standing on the wooden cross preceding the anxiety-inducing environment, making it likely that the elevated heart rate observed in the 30 seconds before task onset reflected anticipatory anxiety (Butler & Mathews, 1987). Anticipatory refers to changes measurable before an event, whereas sensitization reflects a heightened responsiveness of psychophysiological or subjective anxiety markers during the event itself. The elevated subjective anxiety ratings and skin conductance levels observed across repeated exposure may therefore also indicate sensitization. These findings underscore that while behavioral readouts of the EPM are highly reliable, certain physiological and subjective indices may be more sensitive to contextual factors.

Regarding validity, the results of Study 1 also showed that two newly designed environments (a naturalistic desert scene and a stylized video-game-like scene) did not differ significantly in their ability to elicit behavioral, physiological, and subjective responses. Ecological validity is commonly defined as the degree to which experimental findings generalize to real-world behavior (Schmuckler, 2001). In the virtual EPM validation study, Biedermann et al. (2017) argued that mixed-reality adaptations of the EPM already possess ecological validity because they are more life-like than conventional laboratory tasks. Earlier VR research has emphasized the role of presence and immersion in evoking emotional responses in VR (Felnhofer et al., 2015; Kisker, Gruber, et al., 2021; Kisker, Lange, et al., 2021; Slater & Wilbur, 1997). While some studies reported that presence differs across emotional states (Baños et al., 2004, 2008; Riva et al., 2007), others suggested that presence may rather serve as a prerequisite for affective responses to occur (Felnhofer et al., 2014), although experimental findings remain mixed (Gromer et al., 2018, 2019). Moreover, prior work proposed that 360° video VR environments enhance the sense of presence and thereby promote more realistic behavior (Schöne, Kisker, et al., 2021). This raises the question of whether more visually realistic environments confer higher ecological validity compared to those with less real-word resemblance. Though Study 1 did not compare 360° video-based VR environments with computer-generated VR environments, the study found that a stylized video-game scene was as reliable in eliciting anxiety-related behavior on the EPM as a naturalistic, therefore more visually realistic, scene. These findings suggest that ecological validity of the virtual EPM may not rely on strict visually realistic resemblance to real-world settings, provided that core characteristics of the virtual EPM are kept (e.g., open arms above abyss) and presence and immersion are achieved. Presence and immersion could be achieved due to the mixed-reality setup, where additional tactile cues (e.g., wind stimulation) supported immersion, which is consistent with studies showing that sensorimotor cues can enhance presence and fear responses in VR (Deligiannidis & Jacob, 2006; Gromer et al., 2018; Hoffman et al., 2003; Hülsmann et al., 2014; Peperkorn & Mühlberger, 2013).

While Study 1 primarily addressed reliability and ecological validity of the virtual EPM, **Study 4** further extended this line of research by examining the validity of the virtual DLB. Face validity (whether the test appears to measure what it's supposed to measure) of the DLB was supported by consistent findings across measures. Specific subjective anxiety ratings increased from bright to dark areas. Compared to the EPM, participants reported higher presence and immersion in the virtual DLB (EPM: spatial presence = 8.1 ± 0.6, involvement =

2.3 ± 0.8, experienced realism = 4.8 ± 0.7; DLB: 14.62 ± 3.7, 5.9 ± 4.3, 8.7 ± 3.5). Moreover, participants showed careful movement in the dark, paralleling rodent behavior in the aversive compartment. Video analyses revealed seemingly protective postures (arms raised in front of the torso when entering dark areas, see https://osf.io/e7n3z/?view_only=2ddaae1993d14c88bfdcc7e6dbbf8881, "Example 1"), which may be considered a human analogue of risk-assessment behaviors described in rodent models, thereby underscoring the ecological credibility of the paradigm. Together, these findings indicate that the DLB successfully elicited the intended anxiogenic effects.

Content validity (whether the test captures a full range of features relevant to a given construct) of the DLB was addressed by combining behavioral, physiological, endocrine, and self-report measures, thus covering a broad spectrum of anxiety-relevant features. Subjective ratings increased in the dark areas, and participants showed a clear stress response with heart rate, skin conductance, and respiratory rate rising in the first minute of exposure. This was mirrored by sympathetic activation indexed by an immediate α-amylase increase, but not by a corresponding cortisol rise. This difference in endocrine responses likely reflects the distinctive reaction pathways. Salivary α-amylase closely tracks adrenergic drive to the salivary glands and reacts within minutes (Nater & Rohleder, 2009), whereas cortisol reflects activation of the hypothalamic-pituitary-adrenal (HPA) axis. In this cascade, the hypothalamus releases corticotropin releasing hormone, which stimulates the anterior pituitary to secrete adrenocorticotropin, prompting the adrenal cortex to release cortisol into the bloodstream. As a result, peak cortisol concentrations typically occur 21-40 minutes after stressor onset (Dickerson & Kemeny, 2004). Prior findings likewise report immediate α-amylase increases but delayed cortisol responses (Gordis et al., 2006; Takai et al., 2004), and the virtual EPM validation showed a similar pattern with α-amylase peaking at test end and cortisol 15 minutes later (Biedermann et al., 2017). Therefore, the DLB likely elicited a lower stressor intensity, such as reduced uncontrollability (Dickerson & Kemeny, 2004), which may have limited a cortisol response. Methodologically, the last cortisol sample was collected 15 minutes after exposure (about 20 minutes from stressor onset), placing it at the lower bound of the cortisol peak window. Even a modest later HPA axis response could therefore have been missed.

On the behavioral level, content validity was only partially supported. Multivariate regression showed that subjective anxiety significantly explained variance across the behavioral domain, suggesting that the task overall taps anxiety-related exploration. At the level of single outcomes, however, only the ethological marker arms-before-body duration was robustly associated with subjective anxiety and differentiated high- from low-anxiety participants after correction. Classical indices such as time in the dark, total distance, and latency of first visit trended in the expected direction but did not survive correction. At trend level, subjective anxiety was linked to reduced total distance, trait anxiety to increased latency of first visit, and claustrophobia-related anxiety to reduced latency. Group comparisons suggested that high-anxiety participants tended to spend less time in the dark and covered less distance, and exploratory post hoc analyses indicated reduced looking-up duration in this group. Although none of these effects remained significant after correction, their directions fit theoretical expectations of avoidance. Thus, the DLB demonstrated robust sensitivity to risk-assessment-like behavior but did not consistently capture classic approach-avoidance indices, limiting the strength of its behavioral content validity. Taken together, content validity was supported at the physiological, endocrine, and subjective levels, but only partly on the behavioral level.

The lack of a robust avoidance behavior within the DLB also limited concurrent validity. Concurrent validity (whether the test outcomes are consistent with other measures of the same construct measured at the same time) was supported by the convergence of physiological and endocrine measures during DLB testing, which showed significant arousal increases, and by alignment with specific subjective reports of anxiety (assessed retrospectively but explicitly referring to feelings during the DLB). Multivariate analyses further showed that subjective anxiety explained variance across the behavioral domain, but robust effects were again limited to the ethological marker arms-before-body duration. Classic avoidance measures such as time in the dark and total distance showed only trend-level associations in the expected direction. Thus, concurrent validity rests mainly on the agreement between physiological/endocrine and subjective indices.

Predictive validity (whether the test can predict behavior on a related measure) of the DLB was partly supported. At the multivariate level, behavioral outcomes were significantly associated with sensation seeking and subjective anxiety. On the level of individual outcomes, sensation seeking predicted greater total distance and more frequent re-examinations, while subjective anxiety was linked to increased arms-before-body duration. These findings partly parallel results from the EPM, where sensation seeking also predicted approach behavior, while acrophobia predicted avoidance (Biedermann et al., 2017), suggesting convergent risk-taking/approach tendency across paradigms. In contrast, in the DLB avoidance measures showed no clear associations with either trait measures or subjective anxiety. Thus, predictive validity in the DLB appears to rely mainly on risk-assessment-like and approach behavior. Pharmacological sensitivity, considered the benchmark for predictive validity in animal models, has yet to be established for the virtual DLB and could strengthen its predictive validity.

Lastly, construct validity (whether the test is grounded and measures in the intended theoretical construct) of the virtual DLB is conceptually well-grounded, as the paradigm directly derives from theoretical models of anxiety as an approach-avoidance conflict and closely parallels rodent analogues. However, in applied terms, construct validity relies on the joint evidence from content and criterion (concurrent and predictive) validity (Biedermann et al., 2017). Because the present findings indicate that these domains are only partially fulfilled, the construct validity of the DLB can be regarded as conceptually strong but empirically incomplete, underscoring the need for further validation.

Overall, the evaluation of the DLB supports its status as a proof of concept rather than a fully validated paradigm. Evidence comes from the consistent signal in the ethological marker arms-before-body duration, the omnibus effect between subjective anxiety and overall behavior, and the convergence with physiological and subjective measures. Yet the lack of robust avoidance effects and the absence of pharmacological validation limit the strength of claims that can be made about the test's validity. Taken together, the findings suggest that while the DLB is not yet fully validated, it shows sufficient promise to justify further refinement and continued validation efforts.

Further Evidence From the DLB: Clinical and Cross-Paradigm Analyses

An essential step in validating experimental paradigms for anxiety research is testing their sensitivity in clinical populations. While studies in healthy samples can establish face, content, and criterion validity, the translational value of the DLB ultimately depends on whether it captures clinically meaningful variation. The DLB was therefore examined in a transdiagnostic patient sample including anxiety disorders, post-traumatic stress disorder (AD/PTSD), and borderline personality disorder (BPD). This approach was chosen because avoidance and enhanced risk assessment are transdiagnostic features of multiple psychiatric conditions, consistent with the Research Domain Criteria (RDoC) framework, an initiative by the U.S. National Institute of Mental Health that promotes dimensional rather than categorical approaches to psychopathology. Within this framework, constructs such as fear and anxiety are emphasized across traditional diagnostic categories (Cuthbert, 2014; Insel et al., 2010;

National Institute of Mental Health). Testing whether the DLB distinguishes patients from matched controls, and whether diagnostic subgroups show distinct behavioral patterns, allows an initial evaluation of its clinical sensitivity and transdiagnostic validity.

Paired-samples *t*-tests for transdiagnostic comparison of patients with individually matched controls showed no statistically significant differences in DLB behavior, and effect sizes were small across variables. The largest observed effect, for re-examination frequency ($dz \approx -0.28$), did not reach significance. A post hoc power analysis indicated that a sample of ~100 matched pairs would be required to detect an effect of this size with 80% power. Given the small effect sizes, patient-control differences appear subtle and their clinical relevance even for larger samples remains to be determined. The heterogeneity of the patient group may have further reduced sensitivity to diagnosis-specific differences. Exploratory subgroup analyses (were based on small samples ($n = 20$ per group) and uncorrected p-values and must be interpreted with caution. Within these limitations, AD/PTSD patients showed a relatively consistent pattern of lower behavioral engagement, most clearly reflected in reduced total distance compared to controls and trend-level reductions in exploratory markers. By contrast, BPD patients exhibited a more heterogeneous response profile, characterized by increased time spent in the dark alongside mixed trends in other measures. These patterns suggest that with both increased and decreased avoidance relative to controls, the BPD group may differ from controls in both directions. Although these subgroup findings were uncorrected and based on a small sample size, they show medium to large effect sizes, which point to potentially diagnosis-specific behavioral profiles that warrant follow-up in larger, stratified cohorts. Overall, the clinical comparison did not provide robust evidence for transdiagnostic patient-control differences but suggests that the DLB may be sensitive to subtle, diagnosis-related variations in behavior.

Beyond the validation of the DLB in isolation, cross-paradigm analysis was conducted to examine the convergence between the virtual DLB and EPM in order to evaluate whether overlapping behavioral tendencies are captured across these independent anxiety tasks. Pearson correlations between core behavioral measures of the EPM and DLB revealed mostly small to moderate associations, with no strong effects exceeding $r = 0.50$. Time spent in the dark, re-examination behavior, and total distance in the DLB were positively correlated with exploration of the open arms, number of open-arm entries, and overall distance in the EPM. This pattern suggests that both tasks tap into overlapping aspects of exploratory and approach-related behavior, consistent with their shared basis in approach-avoidance conflict. Latencies in both paradigms were also positively correlated, suggesting that they capture a shared hesitation or avoidance tendency, possibly reflecting anxiety-related decision-making.

However, the overall absence of strong correlations also indicates that the paradigms are not redundant but rather complementary. Furthermore, ethological markers that were assessed only in the DLB, such as arms-before-body and lower-body behaviors, showed only weak or negligible associations with EPM parameters, indicating that they do not have a meaningful linear relationship with the core behavioral measures of the EPM. This is expected because these variables are derived from detailed video analyses that capture subtle postural and risk-assessment patterns, thereby sensitizing the test and providing more fine-grained information. Similar to rodent EPM studies, where the inclusion of ethological measures such as risk assessment and defensive postures has been shown to enhance the sensitivity of the paradigm and reveal drug-specific effects beyond conventional spatiotemporal metrics (Rodgers et al., 1997; Rodgers & Dalvi, 1997; Rodgers & Johnson, 1995), arms-before-body may likewise serve as a complementary marker that enriches the behavioral repertoire captured in the virtual DLB. Although such ethological coding could in principle also be applied to the virtual EPM, the specific markers that emerge may depend on the task context, suggesting that each paradigm could offer task-specific ethological markers in humans. This complementarity in correlation patterns of the virtual DLB and EPM is consistent with the animal literature, where different approach-avoidance paradigms are known to capture partially distinct aspects of anxiety-related behavior (Calhoon & Tye, 2015; Cryan & Holmes, 2005; La-Vu et al., 2020).

It is worth noting that, while both paradigms are linked to sensation seeking, the EPM has been shown to be influenced by fear of heights (Biedermann et al., 2017; Madeira et al., 2021), whereas the DLB appears to be also shaped by different anxiety-related traits. At the trend level, trait anxiety was associated with increased avoidance (longer latency of first visit, reduced total distance), whereas claustrophobia was linked to reduced avoidance (shorter latency of first visit). The latter trend may be explained by the task setup, where participants started in the smaller bright compartment and moved toward the larger dark area, making entry less aversive for those higher in claustrophobic traits. Taken together, these findings suggest that the paradigms converge on a shared approach-avoidance core but may be influenced by different anxiety-related traits, thereby providing complementary information. While the DLB is not yet fully validated, its partial overlap with EPM variables together with the presence of additional ethological markers identified in the present study, illustrates its potential to add complementary value beyond what is captured by either paradigm alone.

Trait- and Experience-Related Variability in VR Anxiety Behavior

Trait- and experience-related variability emerged as one of the central themes across the present set of studies, underscoring that VR-based anxiety paradigms do not elicit uniform responses but are shaped by stable personality characteristics and prior experiences. Rather than being mere sources of noise, these factors provide important insights into the validity and explanatory power of the paradigms: participants bring their traits and experiences into the virtual environment, thereby enriching the observed behavioral patterns. Evidence from the current work highlights two domains of variability: traits with clinical relevance such as psychopathy, experiential factors such as gaming expertise.

Study 2 demonstrated a clear association between psychopathic personality traits and anxiety-related behavior in the virtual EPM. Higher overall psychopathy scores were linked to reduced avoidance behavior and lower levels of subjective anxiety. Within the dimensional structure of psychopathy measured by the FPP, the subscales Fearlessness and Lack of Empathy emerged as the most consistent predictors of anxiety-related behavior on the EPM, remaining significant across both the multivariate test and the separate regression models. Fearlessness was related to nearly all behavioral outcomes, while Lack of Empathy specifically predicted the number of entries into the open arms and the latency to first open arm entry. The EPM is an established paradigm to detect anxiety-related behavior, and while fear and anxiety are considered conceptually distinct, they are also closely related adaptive emotions (Beckers et al., 2023; Daniel-Watanabe & Fletcher, 2022; Hamm, 2020). Against this background, the observed association between Fearlessness and reduced avoidance behavior appears plausible. The FPP Fearlessness subscale captures reduced anticipation of and sensitivity to negative consequences of behavior (Etzler & Rohrmann, 2017), which may explain its strong predictive value for diminished avoidance in the EPM. In addition, some items of the Fearlessness subscale also reflect physical risk-taking, suggesting that increased approach motivation may further contribute to the observed pattern of results. Moreover, in Study 2, sensation-seeking correlated positively with psychopathic traits and measures of approach behavior on the EPM (see Supplementary Table S1). This is consistent with prior work linking risk-taking tendencies to psychopathic traits (Book et al., 2022).

By contrast, the predictive role of Lack of Empathy is less straightforward. One possibility is that reduced emotional empathy, as captured by the FPP (e.g., diminished resonance with others' fear; Etzler & Rohrmann, 2017), reflects a broader affective detachment that may extend to one's own protective impulses. This could translate into less hesitation and greater approach behavior in anxiogenic contexts such as the EPM. Converging evidence from neuroimaging, psychophysiological, and behavioral studies on psychopathy indicates reduced responsivity to others' negative emotional signals, such as fear and pain (Blair, 2008; Brook & Kosson, 2013; Decety et al., 2013; Kim & Kim, 2024). Yet, whether such mechanisms also impact the processing of one's own anxiety signals remains to be further researched. Impulsivity was related to shorter latencies in open arm end-exploration in the univariate regression, but this effect did not remain consistent across outcomes once other psychopathy subscales were considered in the multivariate model. Nevertheless, this finding is noteworthy, as impulsivity represents a core feature of the antisocial-lifestyle component of psychopathy (Factor 2). While meta-analytic evidence has shown that Factor 2 is often positively associated with self-reported anxiety (Derefinko, 2015), this study indicates the opposite for anxiety-related behavior. The apparent discrepancy between our behavioral findings and these self-report data may reflect differences between subjective reports of anxious affect and experimental measures of approach-avoidance conflict.

Together these findings are in line with previous work (N. E. Anderson et al., 2021), and with the well-documented "lack of anxiety" hypothesis in psychopathy (Lilienfeld & Andrews, 1996; Neumann et al., 2013; Patrick et al., 1993) as well as possible deficits in the inhibition of potentially inappropriate approach behavior (Blair et al., 2004; Newman & Kosson, 1986). The present results extend this line of research by demonstrating that fearlessness as a component of psychopathy is also expressed at the behavioral level, as higher psychopathy scores were associated with less subjective anxiety and reduced avoidance in the EPM. Moreover, addressing actual behavior and not just proxy measures (e.g., questionnaires or computer tasks) allows a more direct assessment of the interplay between psychopathic traits and anxiety. At the same time, the findings should not be taken as evidence for a general absence of anxiety in psychopathy, but rather as indication of psychopathic trait-related differences in approach-avoidance behavior measured on the EPM.

Beyond these immediate results, the pattern of findings also resonates with broader theoretical frameworks. The subscale Impulsivity links conceptually to the ICD-11 (World Health Organization, 2022) trait domain Disinhibition as one of the five trait domain specifiers, including impulsive behavior, which, together with Dissociality, forms part of the dimensional approach to personality disorders (Bach & First, 2018). Moreover, the current study's observation that Fearlessness and Lack of Empathy jointly predicted anxiety-related behavior in the EPM is consistent with the ICD-11 framework that emphasizes domain-level traits as the basis for diagnosing personality disorders. In addition, the complex relationship between anxiety, antisocial tendencies, and delinquent behavior remains debated in the literature. Some evidence points towards less chronic offending of individuals with antisocial personality disorder who also show higher levels of anxiety and depression, indicating a potential protective effect of anxiety on criminal behavior (Moffitt et al., 2002; Polier et al., 2012). Other studies show a strong relationship between anxiety, depression, and antisocial behavior in broader populations, suggesting that these symptom domains often co-occur (Goodwin & Hamilton, 2003; Lenzenweger et al., 2007; Sareen et al., 2004). From this perspective, replication of the current findings in forensic or clinical populations would provide important insights into the role of anxiety-related behavior in psychopathy and its potential impact on explicitly criminal behavior.

An important aspect of this work is that the associations were not confined to subjective reports but also emerged at the level of actual behavior, demonstrating that the EPM is sensitive to stable personality traits as they manifest in approach-avoidance conflict. This indicates that VR-based anxiety paradigms can be suitable to investigate how stable traits shape anxiety-related behavior. Moreover, the findings suggest that the virtual EPM can serve a broader transdiagnostic function: rather than being limited to the study of anxiety disorders, it can be used to capture anxiety-related processes on a behavioral level that cut across multiple psychiatric conditions. This highlights the paradigm's potential as a methodological bridge between experimental psychopathology and clinical assessment.

While Study 2 focused on stable personality traits with clinical relevance, **Study 3** examined the role of prior experiences in shaping behavior in the virtual EPM. Specifically, we asked whether gaming experience influences approach-avoidance behavior on the virtual EPM and whether such effects could explain differences in anxiety-related behavior between men and women. The main finding was that women with gaming experience showed less anxiety-related behavior than those without such experience on the EPM. In contrast, no corresponding effect was observed in men. Although the interaction between gender and gaming experience was not significant across all behavioral parameters, subgroup analyses indicate that effects of gaming were primarily evident in women. In this study the subgroup of men without gaming experience was small ($n = 27$), possibly accounting for the lack of difference in anxiety-related behavior in this group. However, this gender bias in gaming experience is also reflected by the literature, with men typically engaging in gaming more than women (Lopez-Fernandez et al., 2019; Veltri et al., 2014). Effect size estimates likewise indicated no substantial effects of gaming in men, whereas effects in women reached weak to medium range, suggesting that the absence of effects in men is unlikely to be attributable to sample size alone. Sex and gender differences in anxiety-related behavior have been documented in both rodent models and humans, although with opposing patterns (Bangasser & Cuarenta, 2021; Donner & Lowry, 2013). Rodent studies often report lower anxiety-like behavior in females, although this can depend on strain or estrous cycle stage (Donner & Lowry, 2013; Fernandes et al., 1999). By contrast, in humans anxiety disorders are more prevalent in women compared to men (Donner & Lowry, 2013). Previous EPM studies in humans showed that men behaved in a way that suggests less anxiety than women (Biedermann et al., 2017, 2024; Nouri et al., 2022). In addition, approach-avoidance behavior and subjective anxiety were not affected by estradiol treatment (Nouri et al., 2022). The present findings therefore converge with the gender-difference in behavior of previous human literature, while also suggesting that prior experiences such as gaming may affect behavioral responses differently across genders.

Furthermore, the results also connect to animal research showing that prior task experience in the rodent EPM can alter anxiety-related behavior. Animals previously exposed to the EPM show increased behavioral indices of anxiety up to complete avoidance of the open arms when retested. These effects are not reversed by anxiolytics, an effect referred to as one-trial tolerance (Carobrez & Bertoglio, 2005; File, 1993; Litvin et al., 2008). By contrast, Study 1 applied common rodent re-test strategies of changing the testing environment and a 28-day inter-trial interval, thereby demonstrating high reliability of virtual EPM measures across repeated exposures in humans. Study 3, in turn, suggests that broader experiential factors such as gaming can affect behavior in the virtual EPM. Thus, highlighting that the impact of experience may depend on its nature, with controlled laboratory repetition yielding stable behavioral patterns, whereas prior gaming history even outside of a VR setting could alter behavioral tendencies. Importantly, gaming experience in the present study did not reflect prior VR familiarity but rather regular exposure to video and computer games in general. Such experience may nevertheless influence behavior in a VR-based anxiety paradigm by fostering general learning mechanisms, including enhanced spatial navigation and attentional control (Bavelier, Green, et al., 2012; Kühn et al., 2014; Spence & Feng, 2010), or shaping more strategic and exploratory approaches to novel environments (Adachi & Willoughby, 2013), and modulating (e.g. suppressing) neural activation patterns related to affective and arousal processing in virtual environments (Mathiak & Weber, 2006).

One possible mechanism by which gaming experience could influence behavior in the virtual EPM is desensitization. While most often discussed in the context of aggression and violent video games, the evidence base is mixed (C. A. Anderson et al., 2003; C. A. Anderson & Bushman, 2001; Ferguson, 2010, 2015; Kühn et al., 2018; Lacko et al., 2024; Olejarnik & Romano, 2023). Ambivalence is also found in immersive contexts: greater presence in 3D video gaming has been shown to mediate anger responses following violent play (Lull &

Bushman, 2016), embodiment of a virtual robotic avatar reduced pain sensitivity (Weger & Loughnan, 2014), yet other work reported no generalized desensitization effect of immersive violent games (Wagener & Melzer, 2022). The findings of Study 3 are consistent with the notion that prior gaming experience, at least in women, may contribute to a desensitization to VR stimuli in an anxiety-related context. Such processes could attenuate anxiety-related behavioral responses in VR, yet future studies are needed to reliably confirm the presence of such desensitization regarding anxiety-related behavior and to clarify potential gender differences in this effect.

To date, apart from work on pathological gaming (Marino et al., 2020), no consistent link between gaming and anxiety has been established (Alsaad et al., 2022). This may in part reflect the scarcity of studies directly targeting this association. An exception is the study by Geslin and colleagues (2011), who investigated gaming effects on emotions such as fear, surprise, and anxiety in a VR environment. They found no influence of gaming experience on anxiety. However, the study's limitations were categorization based solely on self-report (gamer/non-gamer), recruitment at a VR trade fair with VR-affine participants, and the use of a first-person shooter paradigm designed to elicit fear rather than anxiety, which restrict the generalizability of this finding. In contrast, the present paradigm was explicitly developed to translate a well-established animal assay of anxiety, the EPM, into VR (Biedermann et al., 2017). Against this background, the current results provide initial evidence that prior gaming experience may influence anxiety-related behavior in a VR setting.

Extending this perspective, personality traits may also interact with gaming habits in shaping behavior on the EPM. Research on gaming and personality points to a complex interplay between individual traits and gaming habits: non-gamers and individuals with problematic gaming behavior typically score higher on neuroticism, with the latter group also showing lower extraversion and conscientiousness, whereas regular gamers tend to report lower neuroticism. Genre preferences are likewise associated with trait differences, with action game players scoring higher in extraversion and lower in neuroticism, and non-gamers tending toward greater conscientiousness (Braun et al., 2016). Further work indicates that frequent gaming is linked to reduced extraversion and conscientiousness but elevated narcissistic tendencies such as exploitativeness and entitlement. Specific genres further correlate with distinct personality dimensions, for example, sports games with extraversion and narcissistic traits, role-playing games with openness, and strategy games with feelings of uniqueness or specialness. Moreover, distinct gaming profiles (casual, challenge, hardcore, and arousal gamers) have been identified, each with characteristic personality signatures; for instance, casual gamers generally score higher on extraversion and conscientiousness and lower on neuroticism compared to hardcore and arousal gamers, who are more emotionally reactive and lower in openness (Potard et al., 2020). These findings converge with Study 2, which investigated the influence of a personality trait (Psychopathic traits) on EPM behavior, underscoring that gaming habits, together with personality traits, likely contribute to the modulation of approach-avoidance tendencies in VR.

Methodological Considerations and Limitations

Several methodological factors shaped the scope of the present findings. The following sections outline key limitations for each study, focusing on aspects such as sample composition, construct operationalization, physiological and endocrine measures, and task design, with different emphases depending on the study.

In **Study 1**, anticipatory anxiety was not assessed separately, limiting interpretation of altered physiological patterns at retest. This would be necessary to confirm whether the absence of distinct autonomic activation in some parameters reflected genuine stability or was masked by anticipatory processes. The design also involved only two test sessions with a fixed four-week interval; more frequent repetitions and comparable environments are known to increase the likelihood of habituation in core behavioral measures of anxiety (Schrader et al., 2018), so alternative testing schedules may have yielded different outcomes. Moreover, no direct comparison was made between repeated use of the same virtual environment and the introduction of a novel one, which would have clarified whether environmental novelty is essential to preserve reliable behavioral responses. Another constraint was the lack of endocrine assessments. Prior research shows that the hypothalamic-pituitary-adrenal (HPA) axis shows habituation across repeated exposures, whereas autonomic responses typically do not (Gerra et al., 2001; Schommer et al., 2003; Wüst et al., 2005). Cortisol is therefore considered more sensitive to endocrine habituation than α-amylase, yet it was not measured in the present study. Finally, although gaming experience was recorded, the lack of variability within the sample prevented any systematic analysis of its role in repeated exposure. This limitation links directly to Study 3, which examined gaming experience as a potential source of individual differences in approach-avoidance behavior.

In **Study 3**, one limitation is the small number of men without gaming experience (n = 27), which reduced the power to detect differences in this subgroup. However, effect size estimates indicated that associations in this subgroup were minimal, suggesting that the absence of findings cannot be solely attributed to sample size limitations. Furthermore, no subjective anxiety measures and physiological (e.g., autonomic arousal) or endocrinological (e.g., cortisol) measures were included. These additional measures could have extended this study by providing evidence on whether anxiety-related behaviors were accompanied by subjective anxiety and physical stress responses, and whether such responses differed between gamers and non-gamers. In addition, definitions of “gamers” varied across studies, with some classifying based on short-term exposure ranging from minutes to hours, complicating cross-study comparisons and highlighting the need for more standardized, long-term operationalizations (Kühn et al., 2018). Most importantly, gaming experience in the present study was assessed only by a binary question, limiting interpretability. These constraints highlight the need for more detailed assessments of experiential and personality-related influences. Study 2 addressed this issue more directly by examining psychopathic traits as a source of stable individual differences in approach-avoidance behavior.

In Study 2, the use of a healthy student population limited the range of psychopathy scores, precluding insights into individuals at the extreme end of the spectrum, such as offender samples. Psychopathy was further assessed with a self-report instrument rather than the Psychopathy Checklist-Revised (PCL-R; Hare, 2003), which is widely regarded as the standard and requires structured clinical interviews. Moreover, as in Study 3, the study focused on behavioral measures, without complementary physiological or endocrine data (e.g., heart rate, skin conductance, cortisol) that could have provided a broader understanding of the relation between psychopathic traits and anxiety-related responses.

In addition to these methodological constraints, focusing on a single paradigm provides only a partial view of anxiety-related behavior. **Study 4** therefore explored the Dark-Light Box (DLB) in both healthy participants and psychiatric patients, aiming to provide complementary insights alongside the EPM. On the physiological level, cortisol sampling ended 15 minutes after exposure, placing it at the lower bound of the expected HPA axis peak window and raising the possibility that later responses were missed. In the behavioral domain, robust effects were confined to the ethological marker arms-before-body, whereas classical avoidance indices such as time in the dark, distance, and latency showed only trend-level associations, limiting the strength of concurrent and content validity.

A design-related factor may have affected the assessment of classical avoidance indices: the bright and dark compartments were not of equal size, complicating comparisons of locomotor activity. Predictive validity also remains incomplete, as pharmacological sensitivity has not yet been established for the virtual DLB. A further consideration concerns the specific subjective anxiety ratings. These were analyzed only for participants who provided complete data across all positions ($n = 40$). This ensured comparability within the ANOVA but limits generalizability, as individuals who consistently avoided the dark compartment did not contribute full ratings. The observed increase in anxiety across darker areas should therefore be interpreted as a conservative estimate, as the most avoidant individuals did not contribute complete ratings.

With respect to motion sickness, the Simulator Sickness Questionnaire (SSQ) appears suboptimal for anxiety paradigms, as visual discomfort is inherent to alternating bright and dark conditions, and physiological responses such as sweating or salivation are more likely to reflect anxiety than simulator sickness. Regarding the clinical sample, the inclusion of 40 patients and 40 matched controls would have provided sufficient power to detect medium effects, but the exploratory subgroup analyses (AD/PTSD vs. BPD) were based on smaller samples of 20 patients per group and uncorrected *p*-values. Furthermore, patient data were not controlled for comorbidity in the statistical analyses. Because several patients presented with overlapping comorbidities, interpretation of the results is limited. While the transdiagnostic comparison partly accounts for such overlap, the exploratory subgroup analyses cannot fully disentangle effects specific to the primary diagnostic categories from those influenced by comorbid conditions. These subgroup findings should therefore be regarded as preliminary. Taken together, the DLB reached proof-of-concept status, but several domains of validity are only partially supported, underscoring the need for continued refinement and validation.

Conclusions and Implications for Translational Anxiety Research

The present set of studies demonstrates that virtual reality adaptations of the EPM and the DLB can capture anxiety-related behavior in humans in ways that parallel established animal paradigms. The virtual EPM proved to be reliable, with behavioral indices showing temporal stability across repeated testing. By contrast, the DLB produced clear subjective and physiological effects, but behavioral outcomes were less robust, suggesting that further refinement is needed before it can reach the same level of validation as the EPM. Importantly, the two paradigms showed both convergence and divergence. Correlations between behavioral measures indicated partial convergence across paradigms, but the absence of strong associations also suggests that each task captures distinct facets of approach-avoidance conflict. In addition, the DLB contributed ethological markers identified in the present study that have not yet been assessed in the virtual EPM, thereby providing complementary information. This pattern supports the combined use of multiple paradigms to obtain a more comprehensive assessment of human anxiety. Furthermore, the findings underline the role of individual differences in shaping behavior. Regular gaming experience was linked to reduced anxiety responses, and psychopathic traits were associated with reduced avoidance and lower subjective anxiety on the EPM. These results illustrate that VR paradigms can be sensitive to trait- and experience-related variability, thereby advancing our understanding of factors shaping behavioral responses in VR anxiety tasks. Finally, the DLB study included psychiatric patients, offering first evidence for clinical applicability. While overall patient-control differences were subtle, exploratory subgroup analyses could suggest diagnosis-related patterns, though these remain tentative and require further research. Together, these findings highlight the translational value of VR paradigms, while also pointing to areas where further validation and clinical testing remain necessary.

Virtual reality thus provides a powerful methodological bridge between animal and human anxiety research. By adapting classical paradigms such as the EPM and DLB into immersive environments, it becomes possible to establish direct cross-species comparisons at the paradigm level. Moreover, this method allows approach-avoidance behavior to be studied directly in humans under conditions that maintain high ecological validity. In addition, VR combines the advantages of high standardization with a strong sense of presence, enabling both reproducibility and lifelike (authentic) engagement under controlled conditions. VR therefore enables the integration of authentic behavioral, psychophysiological, endocrine, and subjective indices, offering a more comprehensive construct representation than is typically possible in rodent paradigms or previous standard human assessments. At the same time, the heterogeneity of human samples, shaped by traits, prior experiences, and clinical subgroups, emerges as a crucial factor shaping responses in VR. This translational potential is not unidirectional: insights from VR studies in humans can also inform animal research. Novel relationships between traits, experiences, and anxiety-related behavior identified in VR can be systematically tested in rodent models, where underlying neurobiological mechanisms can be examined with high experimental control. Such bidirectional exchange strengthens both preclinical and human research, ultimately refining the validity of paradigms across species.

VR paradigms also hold considerable promise for clinical application. By capturing nuanced patterns of approach-avoidance behavior, they may support more differentiated diagnostics, for example by identifying distinct anxiety profiles within a diagnosis or distinguishing between primary disorders and comorbid constellations. Beyond categorical diagnoses, such paradigms also align with transdiagnostic frameworks such as Research Domain Criteria (RDoC), as they probe core dimensions of avoidance and threat processing across disorders (Cuthbert, 2014; Insel et al., 2010). Standardized immersive environments further provide experimental systems in which novel interventions can be tested. While the virtual EPM has already demonstrated pharmacological sensitivity in earlier validation work (Biedermann et al., 2017), the present DLB study provided first indications of feasibility in both healthy and clinical participants. Although these patient data remain preliminary, they highlight the potential of VR paradigms to be extended into clinical populations and point to the importance of further validation in larger, diagnostically well-characterized cohorts. More broadly, the present findings show that VR paradigms are applicable in clinical samples and sensitive to individual differences, thereby opening avenues toward personalized and precision approaches to assessment and treatment.

Beyond their role in anxiety research, the present findings also point to broader societal and forensic implications. The observation that psychopathic traits were linked to reduced avoidance and lower subjective anxiety in the virtual EPM underscores that approach-avoidance paradigms can capture processes relevant not only to anxiety disorders but also to antisociality and delinquent behavior. This aligns with long-standing theoretical accounts of altered threat sensitivity in psychopathy and suggests that VR-based paradigms may provide an ecologically valid way to study these mechanisms. Such an approach could be of particular value in forensic psychology and criminology, where impaired fear processing and diminished avoidance tendencies are central to understanding violent or antisocial behavior. While the present findings do not yet justify direct application in forensic or security contexts, they illustrate the potential of VR paradigms to contribute to these domains if further validation is achieved in offender and high-risk samples. Extending research in this direction will be essential for establishing whether VR-based assessments can meaningfully inform areas of societal relevance, including forensic risk assessment, violence prevention, and the development of targeted interventions. At the same time, the prospect of using behavioral readouts in forensic or criminological assessment raises important ethical concerns. Because such measures reflect impairments that individuals cannot (or only hardly) control, there is a risk of overinterpreting them as deterministic indicators of risk potential. This misuse and reducing offenders to behavioral markers without integrating multiple complementary sources of information could result in biased or overly narrow judgments of risk. Future applications, and especially their users in forensic or clinical practice will therefore need to carefully balance the potential value of VR-based paradigms with the necessity of multimethod assessment strategies to ensure fairness, proportionality, and ethical responsibility.

Future Directions

As touched on in the conclusions of this dissertation, future work will require larger and more diverse samples to strengthen the generalizability of findings. This includes classical clinical cohorts (e.g., patients with anxiety disorders or borderline personality disorder), transdiagnostic samples that capture processes across diagnostic boundaries, and forensic groups such as individuals with high levels of psychopathy. For clinical applications in particular, diagnostically stratified cohorts with systematic comorbidity assessments will be essential to disentangle disorder-specific effects from those driven by overlapping conditions. Larger and more heterogeneous samples would also enable robust subgroup analyses, including targeted oversampling of underrepresented groups such as men without gaming experience, to clarify potential gender-specific influences that were underrepresented in the present work. Moreover, the exploratory subgroup patterns observed in Study 4 will need replication in larger, diagnostically well-characterized cohorts to establish whether diagnosis-specific behavioral differences can be reliably detected. Future research should also consider international, multicenter designs. Such approaches would allow testing whether VR-based anxiety paradigms yield consistent results across different laboratories, technical setups, and cultural contexts. This would not only strengthen external validity but also ensure that potential applications in clinical and forensic settings are based on paradigms that are robust, reproducible, and transferable beyond a single research environment.

A particularly important extension concerns the systematic inclusion of different age groups. To date, many VR-based anxiety paradigms have predominantly been tested in adult samples, yet initial evidence suggests that adolescents may respond differently to VR scenarios. For example, VR exposure has been used to elicit school-related anxiety in 14-17 year-olds, producing measurable state anxiety and physiological activation (Beele et al., 2024). Likewise, in pediatric medicine VR has already been applied to reduce procedural fear and pain in children, demonstrating both feasibility and strong emotional effects (Tas et al., 2022). These findings indicate that extending VR-based anxiety paradigms to under-18 populations could be particularly valuable for prevention and early intervention (Xu et al., 2025), and would also allow systematic investigation of developmental differences in anxiety-related behavior and VR responses.

Another important direction concerns the refinement of study and paradigm design. Future research could build richer datasets by systematically incorporating more detailed assessments of prior experience (e.g., gaming frequency, duration, genre) and by explicitly examining personality traits as potential influences on VR-based anxiety behavior. Systematically including such measures across studies will also enable meta-analytic integration and clarify how specific gaming styles and genres relate to personality traits and anxiety-related behavior, as suggested by prior work (Braun et al., 2016; Potard et al., 2020). For the DLB in particular, several design optimizations are conceivable. Using equally sized bright and dark compartments, as in some rodent implementations (La-Vu et al., 2020), would not only facilitate more balanced exploration but also improve the comparability of locomotor indices, thereby strengthening the reliability of behavioral measures. Moreover, adding sensory cues may increase immersion and strengthened participants' perception of threat. For example, Kisker, Lange, et al. (2021) enhanced a virtual cave paradigm by implementing tactile wall structures, which could inspire similar refinements in the DLB. Beyond optimizing individual tasks, pharmacological validation remains an important benchmark, particularly for the DLB, to establish predictive validity comparable to animal models. Finally, moving toward standardized test batteries that combine multiple paradigms, such as the EPM, DLB, and complementary VR-based tasks, would parallel established practice in animal research and allow the construction of multidimensional anxiety profiles that capture a broader spectrum of approach-avoidance behavior.

A further extension concerns the systematic integration of ethological markers across paradigms. In Study 4, such markers were coded only in the DLB, where they revealed additional behavioral dimensions not captured by spatiotemporal indices. These markers are not restricted to the DLB, but were introduced through detailed video analysis which could in principle also be assessed in the EPM. Future studies should therefore examine which ethological markers emerge in each paradigm and to what extent they overlap or remain task-specific. To make such analyses more feasible, advances in automated tracking and AI-based behavior recognition will be essential, enabling higher-throughput assessment and systematic incorporation of ethological measures into VR-based anxiety research.

Building on such behavioral refinements, next-generation research should also take advantage of emerging technologies that allow richer and more automatic readouts during VR exposures. In particular, neuroimaging in VR is becoming increasingly feasible: immersive VR combined with mobile functional near infrared spectroscopy (fNIRS) can track brain responses during authentic behavior and shows clear potential for cognitive-affective paradigms, including threat and avoidance contexts (Peng et al., 2024). In parallel, mobile EEG-VR pipelines are being deployed not only for measurement but also for real-time neurofeedback, allowing patients to practice self-regulation and therapists to adapt interventions dynamically. Such systems are further extending toward remote interventions, highlighting the increasing practicality of brain-signal integration in therapeutic VR (Castanho et al., 2025; Tacca et al., 2024). In addition, real-world extended reality approaches can push ecological embedding beyond the laboratory. Smartphone-based augmented reality (AR) exposure for specific phobias has already shown preliminary clinical benefits, suggesting that graded, contextually anchored exposures could be delivered directly in everyday environments and made accessible to a broad range of patients (Jurcik et al., 2024; Rajkumar, 2024). Field-deployable VR protocols are also emerging to test behavior outside traditional laboratory settings while retaining experimental control (Quirós-Ramírez et al., 2025).

Another promising avenue concerns the application of machine learning (ML) methods to VR-based anxiety research. In the present studies, classical statistical approaches such as regression and multivariate models were used to test specific, hypothesis-driven associations (e.g., the influence of psychopathy or gaming experience on avoidance behavior). These methods are well suited for identifying targeted effects and providing interpretable estimates of their size, but they are limited when it comes to handling larger sets of predictors or non-linear relationships. ML approaches, by contrast, are data-driven and allow for the simultaneous integration of many behavioral, physiological, and endocrine measures. Instead of testing one predictor at a time, ML algorithms can detect complex, multivariate patterns and optimize predictive accuracy across a wide range of input features. For example, combining ethological markers, movement trajectories, gaze, heart rate, electrodermal activity, respiratory rate, and salivary markers could reveal latent profiles of anxiety responses that are not evident from single variables alone. Recent studies have trained machine learning models on multimodal data from VR therapy sessions for social anxiety and on real-time physiological signals recorded during VR exposure, demonstrating that ML can predict not only overall anxiety severity but also specific symptom dimensions with viable accuracy (Chun et al., 2022; Park et al., 2025). While these models often trade some interpretability for predictive power, new methods for feature attribution increasingly allow insight into the relative contribution of different variables. An example of these advancements are Shapley Additive Explanations (SHAP) values, which estimate each variable's contribution to model predictions. In this way, ML complements hypothesis-driven analyses by enabling the discovery of novel patterns, the prediction of clinical outcomes such as diagnostic group membership or treatment response, and ultimately the development of personalized models of anxiety behavior in VR.

Ultimately, these developments open the door to personalized modeling. By integrating behavioral data from VR paradigms with physiological and neural measures, it becomes possible to derive individual response profiles rather than relying solely on group averages. Such profiles could reveal whether a person primarily shows heightened physiological arousal (e.g., increased heart rate and sweating), pronounced avoidance behavior, or elevated subjective anxiety ratings. In turn, this information could guide treatment selection and monitoring. For example, indicating who might benefit most from exposure-based therapy versus pharmacological augmentation. Importantly, profiles can also be tracked over time to evaluate whether interventions are producing measurable change. Emerging research in digital mental health has already pointed this direction, emphasizing the need for multimodal, dynamic markers (Hammelrath et al., 2025) and underscoring broader implementation issues for such precision approaches in psychiatry (Torous et al., 2025).

Taken together, the studies presented in this dissertation illustrate how Virtual Reality can serve as a methodological bridge between the controlled precision of animal models and the complexity of human behavior. By translating paradigms such as the Elevated Plus Maze and Dark-Light Box into immersive, interactive simulations, VR enables the systematic study of defensive behaviors under ecologically valid yet standardized conditions. This duality of experimental control and experiential realism creates the conditions for integrative approaches that link behavioral data with psychophysiological measures and, in the future, computational models or machine learning frameworks. In this sense, the present work contributes to a growing movement toward translational, data-rich anxiety research that connects biological mechanisms, subjective experience, and predictive modeling. Ultimately, such interdisciplinary frameworks hold promise for advancing our understanding of human anxiety and refining the tools through which we assess, predict, and intervene in maladaptive fear and anxiety processes.

Chapter 5 References

Adachi, P. J. C., & Willoughby, T. (2013). More than just fun and games: The longitudinal relationships between strategic video games, self-reported problem solving skills, and academic grades. Journal of Youth and Adolescence, 42(7), 1041–1052. https://doi.org/10.1007/s10964-013-9913-9

Alsaad, F., Binkhamis, L., Alsalman, A., Alabdulqader, N., Alamer, M., Abualait, T., Khalil, M. S., & Al Ghamdi, K. S. (2022). Impact of action video gaming behavior on attention, anxiety, and sleep among university students. Psychology Research and Behavior Management, 15, 151–160. https://doi.org/10.2147/PRBM.S347694

Alsina-Jurnet, I., & Gutiérrez-Maldonado, J. (2010). Influence of personality and individual abilities on the sense of presence experienced in anxiety triggering virtual environments. International Journal of Human-Computer Studies, 68(10), 788–801. https://doi.org/10.1016/j.ijhcs.2010.07.001

Anderson, C. A., Berkowitz, L., Donnerstein, E., Huesmann, L. R., Johnson, J. D., Linz, D., Malamuth, N. M., & Wartella, E. (2003). The influence of media violence on youth. Psychological Science in the Public Interest, 4(3), 81–110. https://doi.org/10.1111/j.1529-1006.2003.pspi_1433.x

Anderson, C. A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature. Psychological Science, 12(5), 353–359. https://doi.org/10.1111/1467-9280.00366

Anderson, N. E., & Kiehl, K. A. (2012). The psychopath magnetized: Insights from brain imaging. Trends in Cognitive Sciences, 16(1), 52–60. https://doi.org/10.1016/j.tics.2011.11.008

Anderson, N. E., & Stanford, M. S. (2012). Demonstrating emotional processing differences in psychopathy using affective ERP modulation. Psychophysiology, 49(6), 792–806. https://doi.org/10.1111/j.1469-8986.2012.01369.x

Anderson, N. E., Widdows, M., Maurer, J. M., & Kiehl, K. A. (2021). Clarifying fearlessness in psychopathy: An examination of thrill-seeking and physical risk-taking. Journal of Psychopathology and Behavioral Assessment, 43(1), 21–32. https://doi.org/10.1007/s10862-020-09847-y

Armario, A., Labad, J., & Nadal, R. (2020). Focusing attention on biological markers of acute stressor intensity: Empirical evidence and limitations. Neuroscience & Biobehavioral Reviews, 111, 95–103. https://doi.org/10.1016/j.neubiorev.2020.01.013

Augustsson, H., & Meyerson, B. J. (2004). Exploration and risk assessment: A comparative study of male house mice (Mus musculus musculus) and two laboratory strains. Physiology & Behavior, 81(4), 685–698. https://doi.org/10.1016/j.physbeh.2004.03.014

Aupperle, R. L., & Paulus, M. P. (2010). Neural systems underlying approach and avoidance in anxiety disorders. Dialogues in Clinical Neuroscience, 12(4), 517–531. https://doi.org/10.31887/DCNS.2010.12.4/raupperle

Aupperle, R. L., Sullivan, S., Melrose, A. J., Paulus, M. P., & Stein, M. B. (2011). A reverse translational approach to quantify approach-avoidance conflict in humans. Behavioural Brain Research, 225(2), 455–463. https://doi.org/10.1016/j.bbr.2011.08.003

Bach, B., & First, M. B. (2018). Application of the ICD-11 classification of personality disorders. BMC Psychiatry, 18(1), 351. https://doi.org/10.1186/s12888-018-1908-3

Balk, S. A., Bertola, D. B., & Inman, V. W. (2013). Simulator sickness questionnaire: Twenty years later. Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, 257–263. https://doi.org/10.17077/drivingassessment.1498

Bangasser, D. A., & Cuarenta, A. (2021). Sex differences in anxiety and depression: Circuits and mechanisms. Nature Reviews Neuroscience, 22(11), 674–684. https://doi.org/10.1038/s41583-021-00513-0

Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., & Rey, B. (2004). Immersion and emotion: Their impact on the sense of presence. CyberPsychology & Behavior, 7(6), 734–741. https://doi.org/10.1089/cpb.2004.7.734

Baños, R. M., Botella, C., Rubió, I., Quero, S., García-Palacios, A., & Alcañiz, M. (2008). Presence and emotions in virtual environments: The influence of stereoscopy. CyberPsychology & Behavior, 11(1), 1–8. https://doi.org/10.1089/cpb.2007.9936

Bateson, M., Desire, S., Gartside, S. E., & Wright, G. A. (2011). Agitated honeybees exhibit pessimistic cognitive biases. Current Biology, 21(12), 1070–1073. https://doi.org/10.1016/j.cub.2011.05.017

Bavelier, D., Achtman, R. L., Mani, M., & Föcker, J. (2012). Neural bases of selective attention in action video game players. Vision Research, 61, 132–143. https://doi.org/10.1016/j.visres.2011.08.007

Bavelier, D., Green, C. S., Pouget, A., & Schrater, P. (2012). Brain plasticity through the life span: Learning to learn and action video games. Annual Review of Neuroscience, 35, 391–416. https://doi.org/10.1146/annurev-neuro-060909-152832

Beckers, T., Hermans, D., Lange, I., Luyten, L., Scheveneels, S., & Vervliet, B. (2023). Understanding clinical fear and anxiety through the lens of human fear conditioning.

Nature Reviews Psychology, 2(4), 233–245. https://doi.org/10.1038/s44159-023-00156-1

Beele, G., Liesong, P., Bojanowski, S., Hildebrand, K., Weingart, M., Asbrand, J., Correll, C. U., Morina, N., & Uhlhaas, P. J. (2024). Virtual reality exposure therapy for reducing school anxiety in adolescents: Pilot study. JMIR Mental Health, 11, e56235. https://doi.org/10.2196/56235

Bekoff, M. (2000). Animal emotions: Exploring passionate natures. BioScience, 50(10), 861. https://doi.org/10.1641/0006-3568(2000)050%5B0861:AEEPN%5D2.0.CO;2

Belzung, C., & Griebel, G. (2001). Measuring normal and pathological anxiety-like behaviour in mice: A review. Behavioural Brain Research, 125(1), 141–149. https://doi.org/10.1016/S0166-4328(01)00291-1

Bernardo, P. D., Bains, A., Westwood, S., & Mograbi, D. C. (2021). Mood induction using virtual reality: A systematic review of recent findings. Journal of Technology in Behavioral Science, 6(1), 3–24. https://doi.org/10.1007/s41347-020-00152-9

Bertini, C., Cecere, R., & Làdavas, E. (2013). I am blind, but I “see” fear. Cortex, 49(4), 985–993. https://doi.org/10.1016/j.cortex.2012.02.006

Biedermann, S. V., Biedermann, D. G., Wenzlaff, F., Kurjak, T., Nouri, S., Auer, M. K., Wiedemann, K., Briken, P., Haaker, J., Lonsdorf, T. B., & Fuss, J. (2017). An elevated plus-maze in mixed reality for studying human anxiety-related behavior. BMC Biology, 15(1), 125. https://doi.org/10.1186/s12915-017-0463-6

Biedermann, S. V., Roth, L., Biedermann, D., & Fuss, J. (2024). Reliability of repeated exposure to the human elevated plus-maze in virtual reality: Behavioral, emotional, and autonomic responses. Behavior Research Methods, 56(1), 187–198. https://doi.org/10.3758/s13428-022-02046-5

Birbaumer, N., Veit, R., Lotze, M., Erb, M., Hermann, C., Grodd, W., & Flor, H. (2005). Deficient fear conditioning in psychopathy: A functional magnetic resonance imaging study. Archives of General Psychiatry, 62(7), 799. https://doi.org/10.1001/archpsyc.62.7.799

Blair, R. J. R. (2008). Fine cuts of empathy and the amygdala: Dissociable deficits in psychopathy and autism. Quarterly Journal of Experimental Psychology, 61(1), 157–170. https://doi.org/10.1080/17470210701508855

Blair, R. J. R., Mitchell, D. G. V., Leonard, A., Budhani, S., Peschardt, K. S., & Newman, C. (2004). Passive avoidance learning in individuals with psychopathy: Modulation by reward but not by punishment. Personality and Individual Differences, 37(6), 1179–1192. https://doi.org/10.1016/j.paid.2003.12.001

Blanchard, D. C. (2001). Mouse defensive behaviors: Pharmacological and behavioral assays for anxiety and panic. Neuroscience & Biobehavioral Reviews, 25(3), 205–218. https://doi.org/10.1016/S0149-7634(01)00009-4

Blanchard, D. C. (2017). Translating dynamic defense patterns from rodents to people. Neuroscience & Biobehavioral Reviews, 76, 22–28. https://doi.org/10.1016/j.neubiorev.2016.11.001

Blanchard, D. C., & Blanchard, R. J. (2008). Defensive behaviors, fear, and anxiety. In R. J. Blanchard, D. C. Blanchard, G. Griebel, & D. Nutt (Eds.), Handbook of anxiety and fear (Vol. 17, pp. 63–79). Elsevier. https://doi.org/10.1016/S1569-7339(07)00005-7

Blanchard, D. C., Hynd, A. L., Minke, K. A., Minemoto, T., & Blanchard, R. J. (2001). Human defensive behaviors to threat scenarios show parallels to fear- and anxiety-related defense patterns of non-human mammals. Neuroscience & Biobehavioral Reviews, 25(7–8), 761–770. https://doi.org/10.1016/S0149-7634(01)00056-2

Blanchard, R. J., & Blanchard, D. C. (1989). Attack and defense in rodents as ethoexperimental models for the study of emotion. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 13, S3–S14. https://doi.org/10.1016/0278-5846(89)90105-X

Blanchard, R. J., Blanchard, D. C., Griebel, G., & Nutt, D. (2008). Introduction to the handbook on fear and anxiety. In R. J. Blanchard, D. C. Blanchard, G. Griebel, & D. Nutt (Eds.), Handbook of anxiety and fear (Vol. 17, pp. 3–7). Elsevier. https://doi.org/10.1016/S1569-7339(07)00001-X

Blanchard, R. J., Blanchard, D. C., Weiss, S. M., & Meyer, S. (1990). The effects of ethanol and diazepam on reactions to predatory odors. Pharmacology Biochemistry and Behavior, 35(4), 775–780. https://doi.org/10.1016/0091-3057(90)90357-N

Blanchard, R. J., Flannelly, K. J., & Blanchard, D. C. (1986). Defensive behaviors of laboratory and wild Rattus norvegicus. Journal of Comparative Psychology, 100(2), 101.

Boesch, M., Sefidan, S., Ehlert, U., Annen, H., Wyss, T., Steptoe, A., & La Marca, R. (2014). Mood and autonomic responses to repeated exposure to the Trier Social Stress Test for Groups (TSST-G). Psychoneuroendocrinology, 43, 41–51. https://doi.org/10.1016/j.psyneuen.2014.02.003

Book, A., Visser, B., & Wattam, T. (2022). The effect of fear-inducing stimuli on risk taking in people with psychopathic traits. Cognition and Emotion, 36(7), 1313–1326. https://doi.org/10.1080/02699931.2022.2106944

Bouchard, S., Robillard, G., & Renaud, P. (2007). Revising the factor structure of the simulator sickness questionnaire. Annual Review of CyberTherapy and Telemedicine, 5, 128–137.

Braun, B., Stopfer, J. M., Müller, K. W., Beutel, M. E., & Egloff, B. (2016). Personality and video gaming: Comparing regular gamers, non-gamers, and gaming addicts and differentiating between game genres. Computers in Human Behavior, 55, 406–412. https://doi.org/10.1016/j.chb.2015.09.041

Brook, M., & Kosson, D. S. (2013). Impaired cognitive empathy in criminal psychopathy: Evidence from a laboratory measure of empathic accuracy. Journal of Abnormal Psychology, 122(1), 156–166. https://doi.org/10.1037/a0030261

Brunborg, G. S., Mentzoni, R. A., Melkevik, O. R., Torsheim, T., Samdal, O., Hetland, J., Andreassen, C. S., & Palleson, S. (2013). Gaming addiction, gaming engagement, and psychological health complaints among Norwegian adolescents. Media Psychology, 16(1), 115–128. https://doi.org/10.1080/15213269.2012.756374

Budhani, S., Richell, R. A., & Blair, R. J. R. (2006). Impaired reversal but intact acquisition: Probabilistic response reversal deficits in adult individuals with psychopathy. Journal of Abnormal Psychology, 115(3), 552–558. https://doi.org/10.1037/0021-843x.115.3.552

Butler, G., & Mathews, A. (1987). Anticipatory anxiety and risk perception. Cognitive Therapy and Research, 11(5), 551–565. https://doi.org/10.1007/BF01183858

Calhoon, G. G., & Tye, K. M. (2015). Resolving the neural circuits of anxiety. Nature Neuroscience, 18(10), 1394–1404. https://doi.org/10.1038/nn.4101

Canteras, N. S., & Blanchard, D. C. (2008). A behavioral and neural systems comparison of unconditioned and conditioned defensive behavior. In R. J. Blanchard, D. C. Blanchard, G. Griebel, & D. Nutt (Eds.), Handbook of anxiety and fear (Vol. 17, pp. 141–153). Elsevier. https://doi.org/10.1016/S1569-7339(07)00009-4

Carobrez, A. P., & Bertoglio, L. J. (2005). Ethological and temporal analyses of anxiety-like behavior: The elevated plus-maze model 20 years on. Neuroscience & Biobehavioral Reviews, 29(8), 1193–1205. https://doi.org/10.1016/j.neubiorev.2005.04.017

Carranza-Pinedo, V., Krohs, U., & Richter, S. H. (2025). Towards a scientific definition of animal emotions: Integrating innate, appraisal, and network mechanisms. Neuroscience & Biobehavioral Reviews, 172, 106127. https://doi.org/10.1016/j.neubiorev.2025.106127

Castanho, L., Martinho, D. V., Saial, A. C., Gouveia, B. R., Gouveia, É. R., & Ribeiro, F. (2025). The efficacy of virtual reality-based EEG neurofeedback in health-related symptoms relief: A systematic review. Applied Psychophysiology and Biofeedback. https://doi.org/10.1007/s10484-025-09730-0

Centifanti, L. C. M., Gillespie, S. M., & Thomson, N. D. (2022). Skin conductance responses to a discrete threat in virtual reality: Associations with psychopathy and anxiety. Journal of Psychopathology and Behavioral Assessment, 44(1), 39–50. https://doi.org/10.1007/s10862-021-09943-7

Chambless, D. L., Caputo, G. C., Bright, P., & Gallagher, R. (1984). Assessment of fear of fear in agoraphobics: The Body Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire. Journal of Consulting and Clinical Psychology, 52(6), 1090–1097. https://doi.org/10.1037/0022-006X.52.6.1090

Chun, J. Y., Kim, H.-J., Hur, J.-W., Jung, D., Lee, H.-J., Pack, S. P., Lee, S., Kim, G., Cho, C.-Y., Lee, S.-M., Lee, H., Choi, S., Cheong, T., & Cho, C.-H. (2022). Prediction of Specific Anxiety Symptoms and Virtual Reality Sickness Using In Situ Autonomic Physiological Signals During Virtual Reality Treatment in Patients With Social Anxiety Disorder: Mixed Methods Study. JMIR Serious Games, 10(3), e38284. https://doi.org/10.2196/38284

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cleckley, H. M. (1988). The mask of sanity: An attempt to clarify some issues about the so-called psychopathic personality (5th ed.). Author.

Cohen, D. C. (1977). Comparison of self-report and overt-behavioral procedures for assessing acrophobia. Behavior Therapy, 8(1), 17–23. https://doi.org/10.1016/S0005-7894(77)80116-0

Cohen, S., Hamrick, N. M. S., Rodriguez, M. S., Feldman, P. J., Rabin, B. S., & Manuck, S. B. (2000). The stability of and intercorrelations among cardiovascular, immune, endocrine, and psychological reactivity. Annals of Behavioral Medicine, 22(3), 171–179. https://doi.org/10.1007/BF02895111

Crawley, J. N. (1981). Neuropharmacologic specificity of a simple animal model for the behavioral actions of benzodiazepines. Pharmacology Biochemistry and Behavior, 15(5), 695–699. https://doi.org/10.1016/0091-3057(81)90007-1

Crawley, J. N. (1985). Exploratory behavior models of anxiety in mice. Neuroscience & Biobehavioral Reviews, 9(1), 37–44. https://doi.org/10.1016/0149-7634(85)90030-2

Crawley, J. N., & Paylor, R. (1997). A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. Hormones and Behavior, 31(3), 197–211. https://doi.org/10.1006/hbeh.1997.1382

Cryan, J., & Holmes, A. (2005). The ascent of mouse: Advances in modelling human depression and anxiety. Nature Reviews Drug Discovery, 4, 775–790. https://doi.org/10.1038/nrd1825

Cuthbert, B. N. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology: Forum - The Research Domain Criteria Project. World Psychiatry, 13(1), 28–35. https://doi.org/10.1002/wps.20087

Daniel-Watanabe, L., & Fletcher, P. C. (2022). Are fear and anxiety truly distinct? Biological Psychiatry Global Open Science, 2(4), 341–349. https://doi.org/10.1016/j.bpsgos.2021.09.006

Darwin, C. (1897). The expression of the emotions in man and animals (3rd ed.). D. Appleton and Company.

Davis, M. (1992). The role of the amygdala in fear and anxiety. Annual Review of Neuroscience, 15, 353–375. https://doi.org/10.1146/annurev.ne.15.030192.002033

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: Role of the extended amygdala in fear vs anxiety. Neuropsychopharmacology, 35(1), 105–135. https://doi.org/10.1038/npp.2009.109

Decety, J., Chen, C., Harenski, C., & Kiehl, K. A. (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://doi.org/10.3389/fnhum.2013.00489

Deligiannidis, L., & Jacob, R. J. K. (2006). The VR scooter: Wind and tactile feedback improve user performance. In Proceedings of the IEEE Symposium on 3D User Interfaces (3DUI'06) (pp. 143–150). IEEE Computer Society. https://doi.org/10.1109/VR.2006.131

Derefinko, K. J. (2015). Psychopathy and Low Anxiety: Meta-Analytic Evidence for the Absence of Inhibition, Not Affect. Journal of Personality, 83(6), 693–709. https://doi.org/10.1111/jopy.12124

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. Psychological Bulletin, 130(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Domonkos, E., Borbélyová, V., Csongová, M., Bosý, M., Kačmárová, M., Ostatníková, D., Hodosy, J., & Celec, P. (2017). Sex differences and sex hormones in anxiety-like behavior of aging rats. Hormones and Behavior, 93, 159–165. https://doi.org/10.1016/j.yhbeh.2017.05.019

Donner, N. C., & Lowry, C. A. (2013). Sex differences in anxiety and emotional behavior. Pflügers Archiv - European Journal of Physiology, 465(5), 601–626. https://doi.org/10.1007/s00424-013-1271-7

Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. Neuropsychologia, 47(8–9), 1780–1789. https://doi.org/10.1016/j.neuropsychologia.2009.02.002

Edmunds, M. (1974). Defence in animals: A survey of anti-predator defences. Longman.

Eilam, D., Izhar, R., & Mort, J. (2011). Threat detection: Behavioral practices in animals and humans. Neuroscience & Biobehavioral Reviews, 35(4), 999–1006. https://doi.org/10.1016/j.neubiorev.2010.08.002

Ekman, P., Freisen, W. V., & Ancoli, S. (1980). Facial signs of emotional experience. Journal of Personality and Social Psychology, 39(6), 1125–1134. https://doi.org/10.1037/h0077722

Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. Motivation and Emotion, 10(2), 159–168. https://doi.org/10.1007/BF00992253

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. Science, 221(4616), 1208–1210. https://doi.org/10.1126/science.6612338

Epstein, S., & Roupenian, A. (1970). Heart rate and skin conductance during experimentally induced anxiety: The effect of uncertainty about receiving a noxious stimulus. Journal of Personality and Social Psychology, 16(1), 20–28. https://doi.org/10.1037/h0029786

Espejo, E. F. (1997). Effects of weekly or daily exposure to the elevated plus-maze in male mice. Behavioural Brain Research, 87(2), 233–238. https://doi.org/10.1016/S0166-4328(97)02286-9

Esteller, À., Poy, R., & Moltó, J. (2016). Deficient aversive-potentiated startle and the triarchic model of psychopathy: The role of boldness. Biological Psychology, 117, 131–140. https://doi.org/10.1016/j.biopsycho.2016.03.012

Etzler, S., & Rohrmann, S. (2017). The development and preliminary validation of a brief questionnaire of psychopathic personality traits. Frontiers in Psychology, 8. https://doi.org/10.3389/fpsyg.2017.01471

Fanselow, M. S. (1980). Conditional and unconditional components of post-shock freezing. The Pavlovian Journal of Biological Science, 15(4), 177–182. https://doi.org/10.1007/BF03001163

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39(2), 175–191. https://doi.org/10.3758/BF03193146

Felnhofer, A., Kothgassner, O. D., Hetterle, T., Beutl, L., Hlavacs, H., & Kryspin-Exner, I. (2014). Afraid to be there? Evaluating the relation between presence, self-reported anxiety, and heart rate in a virtual public speaking task. Cyberpsychology, Behavior, and Social Networking, 17(5), 310–316. https://doi.org/10.1089/cyber.2013.0472

Felnhofer, A., Kothgassner, O. D., Schmidt, M., Heinzle, A.-K., Beutl, L., Hlavacs, H., & Kryspin-Exner, I. (2015). Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. International Journal of Human-Computer Studies, 82, 48–56. https://doi.org/10.1016/j.ijhcs.2015.05.004

Ferguson, C. J. (2010). Blazing angels or resident evil? Can violent video games be a force for good? Review of General Psychology, 14(2), 68–81. https://doi.org/10.1037/a0018941

Ferguson, C. J. (2015). Do angry birds make for angry children? A meta-analysis of video game influences on children's and adolescents' aggression, mental health, prosocial behavior, and academic performance. Perspectives on Psychological Science, 10(5), 646–666. https://doi.org/10.1177/1745691615592234

Fernandes, C., González, M. I., Wilson, C. A., & File, S. E. (1999). Factor analysis shows that female rat behaviour is characterized primarily by activity, male rats are driven by sex and anxiety. Pharmacology Biochemistry and Behavior, 64(4), 731–736. https://doi.org/10.1016/S0091-3057(99)00139-2

File, S. E. (1993). The interplay of learning and anxiety in the elevated plus-maze. Behavioural Brain Research, 58(1), 199–202. https://doi.org/10.1016/0166-4328(93)90103-W

Fowles, D. C. (1980). The three arousal model: Implications of Gray's two-factor learning theory for heart rate, electrodermal activity, and psychopathy. Psychophysiology, 17(2), 87–104. https://doi.org/10.1111/j.1469-8986.1980.tb00117.x

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? Personality and Individual Differences, 17(6), 791–802. https://doi.org/10.1016/0191-8869(94)90048-5

Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. Methods in Ecology and Evolution, 7(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

Gamito, P., Oliveira, J., Morais, D., Baptista, A., Santos, N., Soares, F., Saraiva, T., & Rosa, P. (2010). Training presence: The importance of virtual reality experience on the "sense of being there." Studies in Health Technology and Informatics, 154, 128–133.

Gamito, P., Oliveira, J., Santos, P., Morais, D., Saraiva, T., Pombal, M., & Mota, B. (2008). Presence, immersion and cybersickness assessment through a test anxiety virtual environment. Annual Review of CyberTherapy and Telemedicine, 6, 83–90.

Gerra, G., Zaimovic, A., Mascetti, G. G., Gardini, S., Zambelli, U., Timpano, M., Raggi, M. A., & Brambilla, F. (2001). Neuroendocrine responses to experimentally-induced psychological stress in healthy humans. Psychoneuroendocrinology, 26(1), 91–107. https://doi.org/10.1016/S0306-4530(00)00046-9

Geslin, E., Bouchard, S., & Richir, S. (2011). Gamers' versus non-gamers' emotional response in virtual reality. Journal of CyberTherapy and Rehabilitation, 4(4), 489–494.

Goodwin, R. D., & Hamilton, S. P. (2003). Lifetime comorbidity of antisocial personality disorder and anxiety disorders among adults in the community. Psychiatry Research, 117(2), 159–166. https://doi.org/10.1016/S0165-1781(02)00320-7

Gordis, E., Granger, D., Susman, E., & Trickett, P. (2006). Asymmetry between salivary cortisol and α-amylase reactivity to stress: Relation to aggressive behavior in adolescents. Psychoneuroendocrinology, 31(8), 976–987. https://doi.org/10.1016/j.psyneuen.2006.05.010

Gray, J. A. (1982). Précis of The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system. Behavioral and Brain Sciences, 5(3), 469–484. https://doi.org/10.1017/S0140525X00013066

Gray, J. A., & McNaughton, N. (2000). The neuropsychology of anxiety: An enquiry into the function of the septo-hippocampal system (2nd ed.). Oxford Univ. Press.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. Nature, 423(6939), 534–537. https://doi.org/10.1038/nature01647

Green, C. S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. Journal of Experimental Psychology. Human Perception and Performance, 32(6), 1465–1478. https://doi.org/10.1037/0096-1523.32.6.1465

Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. Psychological Science, 18(1), 88–94. https://doi.org/10.1111/j.1467-9280.2007.01853.x

Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. Current Biology, 22(6), R197–R206. https://doi.org/10.1016/j.cub.2012.02.012

Gromer, D., Kiser, D. P., & Pauli, P. (2021). Thigmotaxis in a virtual human open field test. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-85678-5

Gromer, D., Madeira, O., Gast, P., Nehfischer, M., Jost, M., Müller, M., Mühlberger, A., & Pauli, P. (2018). Height simulation in a virtual reality CAVE system: Validity of fear responses and effects of an immersion manipulation. Frontiers in Human Neuroscience, 12. https://doi.org/10.3389/fnhum.2018.00372

Gromer, D., Reinke, M., Christner, I., & Pauli, P. (2019). Causal interactive links between presence and fear in virtual reality height exposure. Frontiers in Psychology, 10. https://doi.org/10.3389/fpsyg.2019.00141

Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. Psychological Review, 77(5), 419–450. https://doi.org/10.1037/h0029810

Gujjar, K. R., van Wijk, A., Kumar, R., & de Jongh, A. (2019). Efficacy of virtual reality exposure therapy for the treatment of dental phobia in adults: A randomized controlled trial. Journal of Anxiety Disorders, 62, 100–108. https://doi.org/10.1016/j.janxdis.2018.12.001

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. Tutorials in Quantitative Methods for Psychology, 8(1), 23–34. https://doi.org/10.20982/tqmp.08.1.p023

Hamm, A. O. (2020). Fear, anxiety, and their disorders from the perspective of psychophysiology. Psychophysiology, 57(2), e13474. https://doi.org/10.1111/psyp.13474

Hammelrath, L., Brose, A., Heinrich, M., Zagorscak, P., Burchert, S., Langhammer, T., & Knaevelsrud, C. (2025). PREACT-digital: Study protocol for a longitudinal, observational multicentre study on digital phenotypes of non-response to cognitive behavioural therapy for internalising disorders. BMJ Open, 15(7), e102392. https://doi.org/10.1136/bmjopen-2025-102392

Handley, S. L., & Mithani, S. (1984). Effects of alpha-adrenoceptor agonists and antagonists in a maze-exploration model of 'fear'-motivated behaviour. Naunyn-Schmiedeberg's Archives of Pharmacology, 327(1), 1–5. https://doi.org/10.1007/BF00504983

Hare, R. D. (1978). Psychopathy and electrodermal responses to nonsignal stimulation. Biological Psychology, 6(4), 237–246. https://doi.org/10.1016/0301-0511(78)90026-1

Hare, R. D. (2003). The manual for the hare psychopathy checklist-revised (2nd ed.). Multi-Health Systems.

Hare, R. D., Frazelle, J., & Cox, D. N. (1978). Psychopathy and Physiological Responses to Threat of an Aversive Stimulus. Psychophysiology, 15(2), 165–172. https://doi.org/10.1111/j.1469-8986.1978.tb01356.x

Hare, R. D., & Neumann, C. S. (2010). The role of antisociality in the psychopathy construct: Comment on Skeem and Cooke (2010). Psychological Assessment, 22(2), 446–454. https://doi.org/10.1037/a0013635

Hoehfurtner, T., Wilkinson, A., Moszuti, S. A., & Burman, O. H. P. (2025). Evidence of mood states in reptiles. Animal Cognition, 28(1), 52. https://doi.org/10.1007/s10071-025-01973-y

Hoffman, H. G., Garcia-Palacios, A., Carlin, A., Furness III, T. A., & Botella-Arbona, C. (2003). Interfaces that heal: Coupling real and virtual objects to treat spider phobia.

International Journal of Human–Computer Interaction, 16(2), 283–300. https://doi.org/10.1207/S15327590IJHC1602_08

Holmes, A., Parmigiani, S., Ferrari, P. F., Palanza, P., & Rodgers, R. J. (2000). Behavioral profile of wild mice in the elevated plus-maze test for anxiety. Physiology & Behavior, 71(5), 509–516. https://doi.org/10.1016/S0031-9384(00)00373-5

Holmes, A., & Rodgers, R. J. (1998). Responses of Swiss–Webster Mice to Repeated Plus-Maze Experience: Further Evidence for a Qualitative Shift in Emotional State? Pharmacology Biochemistry and Behavior, 60(2), 473–488. https://doi.org/10.1016/S0091-3057(98)00008-2

Hoppenbrouwers, S. S., Bulten, B. H., & Brazil, I. A. (2016). Parsing fear: A reassessment of the evidence for fear deficits in psychopathy. Psychological Bulletin, 142(6), 573–600. https://doi.org/10.1037/bul0000040

Hripcsak, G., & Heitjan, D. F. (2002). Measuring agreement in medical informatics reliability studies. Journal of Biomedical Informatics, 35(2), 99–110. https://doi.org/10.1016/S1532-0464(02)00500-2

Hubert-Wallander, B., Green, C. S., Sugarman, M., & Bavelier, D. (2011). Changes in search rate but not in the dynamics of exogenous attention in action videogame players. Attention, Perception, & Psychophysics, 73(8), 2399–2412. https://doi.org/10.3758/s13414-011-0194-7

Huck, U. W., & Price, E. O. (1975). Differential effects of environmental enrichment on the open-field behavior of wild and domestic Norway rats. Journal of Comparative and Physiological Psychology, 89(8), 892–898. https://doi.org/10.1037/h0077160

Hülsmann, F., Fröhlich, J., Mattar, N., & Wachsmuth, I. (2014). Wind and warmth in virtual reality: Implementation and evaluation. Laval Virtual VRIC'14. https://doi.org/10.1145/2617841.2620712

igroup Project Consortium. (n.d.). Igroup presence questionnaire (IPQ) factor analysis. Igroup Project Consortium. Retrieved October 28, 2025, from https://www.igroup.org/pq/ipq/factor.php

Imhof, J. T., Coelho, Z. M., Schmitt, M. L., Morato, G. S., & Carobrez, A. P. (1993). Influence of gender and age on performance of rats in the elevated plus maze apparatus. Behavioural Brain Research, 56(2), 177–180. https://doi.org/10.1016/0166-4328(93)90036-P

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. American Journal of Psychiatry, 167(7), 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379

Jarius, S., & Wildemann, B. (2015). And Pavlov still rings a bell: Summarising the evidence for the use of a bell in Pavlov's iconic experiments on classical conditioning. Journal of Neurology, 262(9), 2177–2178. https://doi.org/10.1007/s00415-015-7858-5

Johnston, A. L., & File, S. E. (1991). Sex differences in animal tests of anxiety. Physiology & Behavior, 49(2), 245–250. https://doi.org/10.1016/0031-9384(91)90039-Q

Jönsson, P., Wallergård, M., Österberg, K., Hansen, Å. M., Johansson, G., & Karlson, B. (2010). Cardiovascular and cortisol reactivity and habituation to a virtual reality version of the Trier Social Stress Test: A pilot study. Psychoneuroendocrinology, 35(9), 1397–1403. https://doi.org/10.1016/j.psyneuen.2010.04.003

Jurcik, T., Zaremba-Pike, S., Kosonogov, V., Mohammed, A.-R., Krasavtseva, Y., Sawada, T., Samarina, I., Buranova, N., Adu, P., Sergeev, N., Skuratov, A., Demchenko, A., & Kochetkov, Y. (2024). The efficacy of augmented reality exposure therapy in the treatment of spider phobia—A randomized controlled trial. Frontiers in Psychology, 15, 1214125. https://doi.org/10.3389/fpsyg.2024.1214125

Kalueff, A. V., Wheaton, M., & Murphy, D. L. (2007). What's wrong with my mouse model?: Advances and strategies in animal modeling of anxiety and depression. Behavioural Brain Research, 179(1), 1–18. https://doi.org/10.1016/j.bbr.2007.01.023

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. The International Journal of Aviation Psychology, 3(3), 203–220. https://doi.org/10.1207/s15327108ijap0303_3

Kim, A. Y., & Kim, Y. Y. (2024). Reduced late positive potentials to distress in individuals with high psychopathic traits during pain judgment tasks. Biological Psychology, 190, 108810. https://doi.org/10.1016/j.biopsycho.2024.108810

Kirlic, N., Young, J., & Aupperle, R. L. (2017). Animal to human translational paradigms relevant for approach avoidance conflict decision making. Behaviour Research and Therapy, 96, 14–29. https://doi.org/10.1016/j.brat.2017.04.010

Kisker, J., Gruber, T., & Schöne, B. (2021). Behavioral realism and lifelike psychophysiological responses in virtual reality by the example of a height exposure. Psychological Research, 85(1), 68–81. https://doi.org/10.1007/s00426-019-01244-9

Kisker, J., Lange, L., Flinkenflügel, K., Kaup, M., Labersweiler, N., Tetenborg, F., Ott, P., Gundler, C., Gruber, T., Osinsky, R., & Schöne, B. (2021). Authentic fear responses in virtual reality: A mobile EEG study on affective, behavioral and electrophysiological correlates of fear. Frontiers in Virtual Reality, 2. https://doi.org/10.3389/frvir.2021.716318

Knight, M. M., & Arns, L. L. (2006). The relationship among age and other factors on incidence of cybersickness in immersive environment users. Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization, 162. https://doi.org/10.1145/1140491.1140539

Knight, P., Chellian, R., Wilson, R., Behnood-Rod, A., Panunzio, S., & Bruijnzeel, A. (2021). Sex differences in the elevated plus-maze test and large open field test in adult Wistar rats. Pharmacology, Biochemistry, and Behavior, 204, 173168. https://doi.org/10.1016/j.pbb.2021.173168

Kremer, L., Klein Holkenborg, S. E. J., Reimert, I., Bolhuis, J. E., & Webb, L. E. (2020). The nuts and bolts of animal emotion. Neuroscience & Biobehavioral Reviews, 113, 273–286. https://doi.org/10.1016/j.neubiorev.2020.01.028

Kühn, S., Gleich, T., Lorenz, R. C., Lindenberger, U., & Gallinat, J. (2014). Playing Super Mario induces structural brain plasticity: Gray matter changes resulting from training with a commercial video game. Molecular Psychiatry, 19(2), 265–271. https://doi.org/10.1038/mp.2013.120

Kühn, S., Kugler, D., Schmalen, K., Weichenberger, M., Witt, C., & Gallinat, J. (2018). The myth of blunted gamers: No evidence for desensitization in empathy for pain after a violent video game intervention in a longitudinal fMRI study on non-gamers. Neurosignals, 26(1), 22–30. https://doi.org/10.1159/000487217

Lacko, D., Machackova, H., & Smahel, D. (2024). Does violence in video games impact aggression and empathy? A longitudinal study of Czech adolescents to differentiate within- and between-person effects. Computers in Human Behavior, 159, 108341. https://doi.org/10.1016/j.chb.2024.108341

Lange, B., & Pauli, P. (2019). Social anxiety changes the way we move—A social approach-avoidance task in a virtual reality CAVE system. PLOS ONE, 14(12), e0226805. https://doi.org/10.1371/journal.pone.0226805

Larson, C. L., Baskin-Sommers, A. R., Stout, D. M., Balderston, N. L., Curtin, J. J., Schultz, D. H., Kiehl, K. A., & Newman, J. P. (2013). The interplay of attention and emotion: Top-down attention modulates amygdala activation in psychopathy. Cognitive, Affective, & Behavioral Neuroscience, 13(4), 757–770. https://doi.org/10.3758/s13415-013-0172-8

La-Vu, M., Tobias, B. C., Schuette, P. J., & Adhikari, A. (2020). To approach or avoid: An introductory overview of the study of anxiety using rodent assays. Frontiers in Behavioral Neuroscience, 14. https://doi.org/10.3389/fnbeh.2020.00145

LeDoux, J. (2000). Emotion circuits in the brain. Annual Review of Neuroscience, 23, 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155

LeDoux, J. (2012). Rethinking the emotional brain. Neuron, 73(4), 653–676. https://doi.org/10.1016/j.neuron.2012.02.004

Lenzenweger, M. F., Lane, M. C., Loranger, A. W., & Kessler, R. C. (2007). DSM-IV personality disorders in the National Comorbidity Survey Replication. Biological Psychiatry, 62(6), 553–564. https://doi.org/10.1016/j.biopsych.2006.09.019

Leonhardt, M., & Overå, S. (2021). Are there differences in video gaming and use of social media among boys and girls?—A mixed methods approach. International Journal of Environmental Research and Public Health, 18(11), 6085. https://doi.org/10.3390/ijerph18116085

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal population. Journal of Personality Assessment, 66(3), 488–524. https://doi.org/10.1207/s15327752jpa6603_3

Ling, Y., Nefs, H. T., Brinkman, W.-P., Qu, C., & Heynderickx, I. (2013). The relationship between individual characteristics and experienced presence. Computers in Human Behavior, 29(4), 1519–1530. https://doi.org/10.1016/j.chb.2012.12.010

Lister, R. G. (1987). The use of a plus-maze to measure anxiety in the mouse. Psychopharmacology, 92(2), 180–185. https://doi.org/10.1007/BF00177912

Litvin, Y., Pentkowski, N. S., Pobbe, R. L., Blanchard, D. C., & Blanchard, R. J. (2008). Unconditioned models of fear and anxiety. In R. J. Blanchard, D. C. Blanchard, G. Griebel, & D. Nutt (Eds.), Handbook of anxiety and fear (Vol. 17, pp. 81–99). Elsevier. https://doi.org/10.1016/S1569-7339(07)00006-9

Lopez-Fernandez, O., Williams, A. J., Griffiths, M. D., & Kuss, D. J. (2019). Female gaming, gaming addiction, and the role of women within gaming culture: A narrative literature review. Frontiers in Psychiatry, 10. https://doi.org/10.3389/fpsyt.2019.00454

Lull, R. B., & Bushman, B. J. (2016). Immersed in violence: Presence mediates the effect of 3D violent video gameplay on angry feelings. Psychology of Popular Media Culture, 5(2), 133–144. https://doi.org/10.1037/ppm0000062

Lykken, D. T. (1957). A study of anxiety in the sociopathic personality. The Journal of Abnormal and Social Psychology, 55(1), 6–10. https://doi.org/10.1037/h0047232

Lykken, D. T. (1995). The antisocial personalities (1st ed.). Psychology Press. https://doi.org/10.4324/9780203763551

Madeira, O., Gromer, D., Latoschik, M. E., & Pauli, P. (2021). Effects of acrophobic fear and trait anxiety on human behavior in a virtual elevated plus-maze. Frontiers in Virtual Reality, 2. https://doi.org/10.3389/frvir.2021.635048

Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. Annual Review of Neuroscience, 24(1), 897–931. https://doi.org/10.1146/annurev.neuro.24.1.897

Marino, C., Canale, N., Vieno, A., Caselli, G., Scacchi, L., & Spada, M. M. (2020). Social anxiety and Internet gaming disorder: The role of motives and metacognitions. Journal of Behavioral Addictions, 9(3), 617–628. https://doi.org/10.1556/2006.2020.00044

Mason, J. W. (1968). A review of psychoendocrine research on the sympathetic-adrenal medullary system. Biopsychosocial Science and Medicine, 30(5), 631.

Mathiak, K., & Weber, R. (2006). Toward brain correlates of natural behavior: fMRI during violent video games. Human Brain Mapping, 27(12), 948–956. https://doi.org/10.1002/hbm.20234

McEwen, B. S. (2020). Hormones and behavior and the integration of brain-body science. Hormones and Behavior, 119, 104619. https://doi.org/10.1016/j.yhbeh.2019.104619

McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: Fear/anxiety and defensive distance. Neuroscience & Biobehavioral Reviews, 28(3), 285–305. https://doi.org/10.1016/j.neubiorev.2004.03.005

McNaughton, N., & Zangrossi, H. (2008). Theoretical approaches to the modeling of anxiety in animals. In R. J. Blanchard, D. C. Blanchard, G. Griebel, & D. Nutt (Eds.), Handbook of anxiety and fear (Vol. 17, pp. 11–27). Elsevier. https://doi.org/10.1016/S1569-7339(07)00002-1

Miller, N. E. (1944). Experimental studies of conflict. In J. M. Hunt (Ed.), Personality and the behavior disorders (pp. 431–465). Ronald Press.

Mineka, S., & Öhman, A. (2002). Phobias and preparedness: The selective, automatic, and encapsulated nature of fear. Biological Psychiatry, 52(10), 927–937. https://doi.org/10.1016/S0006-3223(02)01669-4

Moffitt, T. E., Caspi, A., Harrington, H., & Milne, B. J. (2002). Males on the life-course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. Development and Psychopathology, 14(1), 179–207. https://doi.org/10.1017/S0954579402001104

Montgomery, K. C. (1955). The relation between fear induced by novel stimulation and exploratory drive. Journal of Comparative and Physiological Psychology, 48(4), 254–260. https://doi.org/10.1037/h0043788

Montgomery, K. C., & Monkman, J. A. (1955). The relation between fear and exploratory behavior. Journal of Comparative and Physiological Psychology, 48(2), 132–136. https://doi.org/10.1037/h0048596

Nater, U. M., & Rohleder, N. (2009). Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: Current state of research.

Psychoneuroendocrinology, 34(4), 486–496. https://doi.org/10.1016/j.psyneuen.2009.01.014

National Institute of Mental Health. (n.d.). RDoC matrix. National Institute of Mental Health. https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/constructs/rdoc-matrix.shtml

Nestler, E. J., Gould, E., & Manji, H. (2002). Preclinical models: Status of basic research in depression. Biological Psychiatry, 52(6), 503–528. https://doi.org/10.1016/S0006-3223(02)01405-1

Nestler, E. J., & Hyman, S. E. (2010). Animal models of neuropsychiatric disorders. Nature Neuroscience, 13(10), 1161–1169. https://doi.org/10.1038/nn.2647

Neumann, C. S., Johansson, P. T., & Hare, R. D. (2013). The Psychopathy Checklist-Revised (PCL-R), low anxiety, and fearlessness: A structural equation modeling analysis. Personality Disorders: Theory, Research, and Treatment, 4(2), 129–137. https://doi.org/10.1037/a0027886

Newman, J. P., & Kosson, D. S. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. Journal of Abnormal Psychology, 95(3), 252–256. https://doi.org/10.1037/0021-843X.95.3.252

Nouri, S., Biedermann, S. V., Joue, G., Auer, M. K., Sommer, T., & Fuss, J. (2022). Effects of circulating estradiol on physiological, behavioural, and subjective correlates of anxiety: A double-blind, randomized, placebo-controlled trial. Psychoneuroendocrinology, 138, 105682. https://doi.org/10.1016/j.psyneuen.2022.105682

Nutt, D., De Miguel, B. G., & Davies, S. J. C. (2008). Phenomenology of anxiety disorders. In R. J. Blanchard, D. C. Blanchard, G. Griebel, & D. Nutt (Eds.), Handbook of anxiety and fear (Vol. 17, pp. 365–393). Elsevier. https://doi.org/10.1016/S1569-7339(07)00017-3

Olejarnik, S. Z., & Romano, D. (2023). Is playing violent video games a risk factor for aggressive behaviour? Adding narcissism, self-esteem and PEGI ratings to the debate. Frontiers in Psychology, 14, 1155807. https://doi.org/10.3389/fpsyg.2023.1155807

Palaus, M., Marron, E. M., Viejo-Sobera, R., & Redolar-Ripoll, D. (2017). Neural basis of video gaming: A systematic review. Frontiers in Human Neuroscience, 11. https://doi.org/10.3389/fnhum.2017.00248

Panksepp, J. (2010). Affective neuroscience of the emotional BrainMind: Evolutionary perspectives and implications for understanding depression. Dialogues in Clinical Neuroscience, 12(4), 533–545. https://doi.org/10.31887/DCNS.2010.12.4/jpanksepp

Panksepp, J. (2011). Cross-species affective neuroscience decoding of the primal affective experiences of humans and related animals. PLoS ONE, 6(9), e21236. https://doi.org/10.1371/journal.pone.0021236

Park, J.-H., Shin, Y.-B., Jung, D., Hur, J.-W., Pack, S. P., Lee, H.-J., Lee, H., & Cho, C.-H. (2025). Machine learning prediction of anxiety symptoms in social anxiety disorder: Utilizing multimodal data from virtual reality sessions. Frontiers in Psychiatry, 15, 1504190. https://doi.org/10.3389/fpsyt.2024.1504190

Pastor, M. C., Moltó, J., Vila, J., & Lang, P. J. (2003). Startle reflex modulation, affective ratings and autonomic reactivity in incarcerated Spanish psychopaths. Psychophysiology, 40(6), 934–938. https://doi.org/10.1111/1469-8986.00111

Patrick, C. J. (1994). Emotion and psychopathy: Startling new insights. Psychophysiology, 31(4), 319–330. https://doi.org/10.1111/j.1469-8986.1994.tb02440.x

Patrick, C. J., Bradley, M. M., & Lang, P. J. (1993). Emotion in the criminal psychopath: Startle reflex modulation. Journal of Abnormal Psychology, 102(1), 82–92. https://doi.org/10.1037/0021-843x.102.1.82

Pavlov, I. P. (1927). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex (G. V. Anrep, Trans.). Oxford University Press. http://psychclassics.yorku.ca/Pavlov/

Pellow, S., Chopin, P., File, S. E., & Briley, M. (1985). Validation of open: Closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. Journal of Neuroscience Methods, 14(3), 149–167. https://doi.org/10.1016/0165-0270(85)90031-7

Peng, K., Moussavi, Z., Karunakaran, K. D., Borsook, D., Lesage, F., & Nguyen, D. K. (2024). iVR-fNIRS: Studying brain functions in a fully immersive virtual environment. Neurophotonics, 11(2), 020601. https://doi.org/10.1117/1.NPh.11.2.020601

Peperkorn, H. M., & Mühlberger, A. (2013). The impact of different perceptual cues on fear and presence in virtual reality. Studies in Health Technology and Informatics, 191, 75–79.

Perkins, A. M., & Corr, P. J. (2006). Reactions to threat and personality: Psychometric differentiation of intensity and direction dimensions of human defensive behaviour. Behavioural Brain Research, 169(1), 21–28. https://doi.org/10.1016/j.bbr.2005.11.027

Perusini, J. N., & Fanselow, M. S. (2015). Neurobehavioral perspectives on the distinction between fear and anxiety. Learning & Memory, 22(9), 417–425. https://doi.org/10.1101/lm.039180.115

Polier, G. G., Vloet, T. D., Herpertz-Dahlmann, B., Laurens, K. R., & Hodgins, S. (2012). Comorbidity of conduct disorder symptoms and internalising problems in children:

Investigating a community and a clinical sample. European Child & Adolescent Psychiatry, 21(1), 31–38. https://doi.org/10.1007/s00787-011-0229-6

Potard, C., Henry, A., Boudoukha, A.-H., Courtois, R., Laurent, A., & Lignier, B. (2020). Video game players' personality traits: An exploratory cluster approach to identifying gaming preferences. Psychology of Popular Media, 9(4), 499–512. https://doi.org/10.1037/ppm0000245

Preuss, T. M., & Wise, S. P. (2022). Evolution of prefrontal cortex. Neuropsychopharmacology, 47(1), 3–19. https://doi.org/10.1038/s41386-021-01076-5

Price, E. O. (1999). Behavioral development in animals undergoing domestication. Applied Animal Behaviour Science, 65(3), 245–271. https://doi.org/10.1016/s0168-1591(99)00087-8

Publication manual of the American Psychological Association (7th ed.). (2020). American Psychological Association. https://doi.org/10.1037/0000165-000

Quirós-Ramírez, M. A., Feineisen, A., Streuber, S., & Reips, U.-D. (2025). Virtual reality experiments in the field. PLOS ONE, 20(4), e0318688. https://doi.org/10.1371/journal.pone.0318688

Radomsky, A. S., Rachman, S., Thordarson, D. S., McIsaac, H. K., & Teachman, B. A. (2001). The Claustrophobia Questionnaire. Journal of Anxiety Disorders, 15(4), 287–297. https://doi.org/10.1016/S0887-6185(01)00064-0

Rajkumar, R. P. (2024). Augmented reality as an aid to behavior therapy for anxiety disorders: A narrative review. Cureus, 16(9), e69454. https://doi.org/10.7759/cureus.69454

Ramos, A. (2008). Animal models of anxiety: Do I need multiple tests? Trends in Pharmacological Sciences, 29(10), 493–498. https://doi.org/10.1016/j.tips.2008.07.005

Ramos, A., Kangerski, A. L., Basso, P. F., Da Silva Santos, J. E., Assreuy, J., Vendruscolo, L. F., & Takahashi, R. N. (2002). Evaluation of Lewis and SHR rat strains as a genetic model for the study of anxiety and pain. Behavioural Brain Research, 129(1–2), 113–123. https://doi.org/10.1016/s0166-4328(01)00337-0

Ramos, A., & Mormède, P. (1997). Stress and emotionality: A multidimensional and genetic approach. Neuroscience & Biobehavioral Reviews, 22(1), 33–57. https://doi.org/10.1016/S0149-7634(97)00001-8

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., Coppola, G., Geyer, M. A., Glanzman, D. L., Marsland, S., McSweeney, F. K., Wilson, D. A., Wu, C.-F., & Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. Neurobiology of Learning and Memory, 92(2), 135–138. https://doi.org/10.1016/j.nlm.2008.09.012

Regenbrecht, H., & Schubert, T. (2002). Real and illusory interactions enhance presence in virtual environments. Presence: Teleoperators and Virtual Environments, 11(4), 425–434. https://doi.org/10.1162/105474602760204318

Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., Gaggioli, A., Botella, C., & Alcañiz, M. (2007). Affective interactions using virtual reality: The link between presence and emotions. CyberPsychology & Behavior, 10(1), 45–56. https://doi.org/10.1089/cpb.2006.9993

Rodgers, R. J., & Dalvi, A. (1997). Anxiety, defence and the elevated plus-maze. Neuroscience & Biobehavioral Reviews, 21(6), 801–810. https://doi.org/10.1016/S0149-7634(96)00058-9

Rodgers, R. J., & Johnson, N. J. T. (1995). Factor analysis of spatiotemporal and ethological measures in the murine elevated plus-maze test of anxiety. Pharmacology Biochemistry and Behavior, 52(2), 297–303. https://doi.org/10.1016/0091-3057(95)00138-M

Rodgers, R. J., Perrault, G., Sanger, D. J., & Griebel, G. (1997). Risk Assessment Behaviour: Evaluation of Utility in the Study of 5-HT-Related Drugs in the Rat Elevated Plus-Maze Test. Pharmacology Biochemistry and Behavior, 57(4), 817–827. https://doi.org/10.1016/S0091-3057(96)00402-9

Rodgers, R. J., & Shepherd, J. K. (1993). Influence of prior maze experience on behaviour and response to diazepam in the elevated plus-maze and light/dark tests of anxiety in mice. Psychopharmacology, 113(2), 237–242. https://doi.org/10.1007/BF02245704

Rosenthal, N., & Brown, S. (2007). The mouse ascending: Perspectives for human-disease models. Nature Cell Biology, 9(9), 993–999. https://doi.org/10.1038/ncb437

Salekin, R. T., Leistico, A.-M. R., Trobst, K. K., Schrum, C. L., & Lochman, J. E. (2005). Adolescent psychopathy and personality theory—the interpersonal circumplex: Expanding evidence of a nomological net. Journal of Abnormal Child Psychology, 33(4), 445–460. https://doi.org/10.1007/s10802-005-5726-y

Sareen, J., Stein, M. B., Cox, B. J., & Hassard, S. T. (2004). Understanding comorbidity of anxiety disorders with antisocial behavior: Findings from two large community surveys. The Journal of Nervous and Mental Disease, 192(3), 178. https://doi.org/10.1097/01.nmd.0000116460.25110.9f

Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M., & Dymond, S. (2016). The tipping point: Value differences and parallel dorsal–ventral frontal circuits gating human approach–avoidance behavior. NeuroImage, 136, 94–105. https://doi.org/10.1016/j.neuroimage.2016.04.070

Schmitt, W. A., & Newman, J. P. (1999). Are all psychopathic individuals low-anxious? Journal of Abnormal Psychology, 108(2), 353–358. https://doi.org/10.1037/0021-843x.108.2.353

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. Infancy, 2(4), 419–436. https://doi.org/10.1207/S15327078IN0204_02

Schneider, P., Ho, Y.-J., Spanagel, R., & Pawlak, C. R. (2011). A novel elevated plus-maze procedure to avoid the one-trial tolerance problem. Frontiers in Behavioral Neuroscience, 5. https://doi.org/10.3389/fnbeh.2011.00043

Scholl, J. L., Afzal, A., Fox, L. C., Watt, M. J., & Forster, G. L. (2019). Sex differences in anxiety-like behaviors in rats. Physiology & Behavior, 211, 112670. https://doi.org/10.1016/j.physbeh.2019.112670

Schommer, N. C., Hellhammer, D. H., & Kirschbaum, C. (2003). Dissociation between reactivity of the hypothalamus-pituitary-adrenal axis and the sympathetic-adrenal-medullary system to repeated psychosocial stress. Psychosomatic Medicine, 65(3), 450–460. https://doi.org/10.1097/01.PSY.0000035721.12441.17

Schöne, B., Kisker, J., Lange, L., Gruber, T., Sylvester, S., & Osinsky, R. (2023). The reality of virtual reality. Frontiers in Psychology, 14, 1093014. https://doi.org/10.3389/fpsyg.2023.1093014

Schöne, B., Kisker, J., Sylvester, R. S., Radtke, E. L., & Gruber, T. (2021). Library for universal virtual reality experiments (LUVRE): A standardized immersive 3D/360° picture and video database for VR based research. Current Psychology. https://doi.org/10.1007/s12144-021-01841-1

Schöne, B., Sylvester, R. S., Radtke, E. L., & Gruber, T. (2021). Sustained inattentional blindness in virtual reality and under conventional laboratory conditions. Virtual Reality, 25(1), 209–216. https://doi.org/10.1007/s10055-020-00450-w

Schrader, A. J., Taylor, R. M., Lowery-Gionta, E. G., & Moore, N. L. T. (2018). Repeated elevated plus maze trials as a measure for tracking within-subjects behavioral performance in rats (Rattus norvegicus). PLOS ONE, 13(11), e0207804. https://doi.org/10.1371/journal.pone.0207804

Schubert, T., Friedmann, F., & Regenbrecht, H. (1999). Embodied presence in virtual environments. In R. Paton & I. Neilson (Eds.), Visual representations and interpretations (pp. 269–278). Springer. https://doi.org/10.1007/978-1-4471-0563-3_30

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. Presence: Teleoperators and Virtual Environments, 10(3), 266–281. https://doi.org/10.1162/105474601300343603

Schultz, D. H., Balderston, N. L., Baskin-Sommers, A. R., Larson, C. L., & Helmstetter, F. J. (2016). Psychopaths show enhanced amygdala activation during fear conditioning. Frontiers in Psychology, 7. https://doi.org/10.3389/fpsyg.2016.00348

Sevinc, V., & Berkman, M. I. (2020). Psychometric evaluation of simulator sickness questionnaire and its variants as a measure of cybersickness in consumer virtual environments. Applied Ergonomics, 82, 102958. https://doi.org/10.1016/j.apergo.2019.102958

Shuhama, R., Del-Ben, C. M., Loureiro, S. R., & Graeff, F. G. (2007). Animal defense strategies and anxiety disorders. Anais Da Academia Brasileira de Ciências, 79(1), 97–109. https://doi.org/10.1590/S0001-37652007000100012

Shuhama, R., Del-Ben, C. M., Loureiro, S. R., & Graeff, F. G. (2008). Defensive responses to threat scenarios in Brazilians reproduce the pattern of Hawaiian Americans and non-human mammals. Brazilian Journal of Medical and Biological Research, 41(4), 324–332. https://doi.org/10.1590/S0100-879X2008000400011

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. Presence: Teleoperators and Virtual Environments, 6(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Solvi, C., Baciadonna, L., & Chittka, L. (2016). Unexpected rewards induce dopamine-dependent positive emotion–like state changes in bumblebees. Science, 353(6307), 1529–1531. https://doi.org/10.1126/science.aaf4454

Spence, I., & Feng, J. (2010). Video games and spatial cognition. Review of General Psychology, 14(2), 92–104. https://doi.org/10.1037/a0019491

Spielberger, C., Gorsuch, R., & Lushene, R. (1970). Manual for the State-Trait Anxiety Inventory. Consulting Psychologists Press.

Steinman, S. A., & Teachman, B. A. (2014). Reaching new heights: Comparing interpretation bias modification to exposure therapy for extreme height fear. Journal of Consulting and Clinical Psychology, 82(3), 404–417. https://doi.org/10.1037/a0036023

Tacca, C., Kerr, B. A., McLamb, C., Ridgway, K. L., & Friis, E. A. (2024). Efficacy of a remote virtual reality and EEG enabled psychotherapy system for the treatment of depressive symptoms. Frontiers in Virtual Reality, 5. https://doi.org/10.3389/frvir.2024.1281017

Tak, S., & Catsambis, S. (2023). “Video games for boys and chatting for girls?”: Gender, screen time activities and academic achievement in high school. Education and Information Technologies, 28(11), 15415–15443. https://doi.org/10.1007/s10639-023-11638-3

Takai, N., Yamaguchi, M., Aragaki, T., Eto, K., Uchihashi, K., & Nishikawa, Y. (2004). Effect of psychological stress on the salivary cortisol and amylase levels in healthy young adults.

Archives of Oral Biology, 49(12), 963–968. https://doi.org/10.1016/j.archoralbio.2004.06.007

Tamietto, M., & De Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. Nature Reviews Neuroscience, 11(10), 697–709. https://doi.org/10.1038/nrn2889

Tas, F. Q., van Eijk, C. A. M., Staals, L. M., Legerstee, J. S., & Dierckx, B. (2022). Virtual reality in pediatrics, effects on pain and anxiety: A systematic review and meta-analysis update. Paediatric Anaesthesia, 32(12), 1292–1304. https://doi.org/10.1111/pan.14546

Thomson, N. D., Aboutanos, M., Kiehl, K. A., Neumann, C., Galusha, C., & Fanti, K. A. (2019). Physiological reactivity in response to a fear-induced virtual reality experience: Associations with psychopathic traits. Psychophysiology, 56(1). https://doi.org/10.1111/psyp.13276

Toet, A., van Welie, M., & Houtkamp, J. (2009). Is a dark virtual environment scary? CyberPsychology & Behavior, 12(4), 363–371. https://doi.org/10.1089/cpb.2008.0293

Torous, J., Linardon, J., Goldberg, S. B., Sun, S., Bell, I., Nicholas, J., Hassan, L., Hua, Y., Milton, A., & Firth, J. (2025). The evolving field of digital mental health: Current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality. World Psychiatry, 24(2), 156–174. https://doi.org/10.1002/wps.21299

Treit, D., Menard, J., & Royan, C. (1993). Anxiogenic stimuli in the elevated plus-maze. Pharmacology Biochemistry and Behavior, 44(2), 463–469. https://doi.org/10.1016/0091-3057(93)90492-C

Vaidyanathan, U., Hall, J. R., Patrick, C. J., & Bernat, E. M. (2011). Clarifying the role of defensive reactivity deficits in psychopathy and antisocial personality using startle reflex methodology. Journal of Abnormal Psychology, 120(1), 253–258. https://doi.org/10.1037/a0021224

van der Staay, F. J. (2006). Animal models of behavioral dysfunctions: Basic concepts and classifications, and an evaluation strategy. Brain Research Reviews, 52(1), 131–159. https://doi.org/10.1016/j.brainresrev.2006.01.006

Veltri, N., Krasnova, H., Baumann, A., & Kalayamthanam, N. (2014). Gender differences in online gaming: A literature review. AMCIS 2014 Proceedings. https://aisel.aisnet.org/amcis2014/SocioTechnicalIssues/GeneralPresentations/5

Voulgaris, A., Biedermann, S. V., Biedermann, D., Bründl, S., Roth, L., Wiessner, C., Briken, P., & Fuss, J. (2024). The impact of psychopathic traits on anxiety-related behaviors in a mixed reality environment. Scientific Reports, 14(1), 11832. https://doi.org/10.1038/s41598-024-62438-9

Wagener, G. L., & Melzer, A. (2022). Comfortably numb? Violent video games and their effects on aggression, mood, and pain-related responses. Entertainment Computing – ICEC 2022, 13477, 18–38. https://doi.org/10.1007/978-3-031-20212-4_2

Walz, N., Mühlberger, A., & Pauli, P. (2016). A human open field test reveals thigmotaxis related to agoraphobic fear. Biological Psychiatry, 80(5), 390–397. https://doi.org/10.1016/j.biopsych.2015.12.016

Weech, S., Kenny, S., Lenizky, M., & Barnett-Cowan, M. (2020). Narrative and gaming experience interact to affect presence and cybersickness in virtual reality. International Journal of Human-Computer Studies, 138, 102398. https://doi.org/10.1016/j.ijhcs.2020.102398

Weger, U. W., & Loughnan, S. (2014). Virtually numbed: Immersive video gaming alters real-life experience. Psychonomic Bulletin & Review, 21(2), 562–565. https://doi.org/10.3758/s13423-013-0512-2

Wieser, M. J., Pauli, P., Grosseibl, M., Molzow, I., & Mühlberger, A. (2010). Virtual social interactions in social anxiety—The impact of sex, gaze, and interpersonal distance. Cyberpsychology, Behavior, and Social Networking, 13(5), 547–554. https://doi.org/10.1089/cyber.2009.0432

Willner, P. (1984). The validity of animal models of depression. Psychopharmacology, 83(1), 1–16. https://doi.org/10.1007/BF00427414

World Health Organization. (2022). International classification of diseases for mortality and morbidity statistics (11th revision) [ICD-11]. World Health Organization. https://icd.who.int/

Würbel, H. (2002). Behavioral phenotyping enhanced – Beyond (environmental) standardization. Genes, Brain and Behavior, 1(1), 3–8. https://doi.org/10.1046/j.1601-1848.2001.00006.x

Wüst, S., Federenko, I. S., Van Rossum, E. F. C., Koper, J. W., & Hellhammer, D. H. (2005). Habituation of cortisol responses to repeated psychosocial stress—Further characterization and impact of genetic factors. Psychoneuroendocrinology, 30(2), 199–211. https://doi.org/10.1016/j.psyneuen.2004.07.002

Xu, D., Liu, Y., Zeng, Y., & Liu, D. (2025). Virtual reality in adolescent mental health management under the new media communication environment. Humanities and Social Sciences Communications, 12(1), 201. https://doi.org/10.1057/s41599-025-04528-1

Zinbarg, R. E. (1998). Concordance and synchrony in measures of anxiety and panic reconsidered: A hierarchical model of anxiety and panic. Behavior Therapy, 29(2), 301–323. https://doi.org/10.1016/S0005-7894(98)80009-9

Zuckerman, M., Bone, R. N., Neary, R., Mangelsdorff, D., & Brustman, B. (1972). What is the sensation seeker? Personality trait and experience correlates of the sensation-seeking scales. Journal of Consulting and Clinical Psychology, 39(2), 308–321. https://doi.org/10.1037/h0033398

Zuckerman, M., Eysenck, S. B., & Eysenck, H. J. (1978). Sensation seeking in England and America: Cross-cultural, age, and sex comparisons. Journal of Consulting and Clinical Psychology, 46(1), 139–149. https://doi.org/10.1037/0022-006X.46.1.139

Zuckerman, M., & Link, K. (1968). Construct validity for the sensation-seeking scale. Journal of Consulting and Clinical Psychology, 32(4), 420–426. https://doi.org/10.1037/h0026047

Zych, A. D., & Gogolla, N. (2021). Expressions of emotions across species. Current Opinion in Neurobiology, 68, 57–66. https://doi.org/10.1016/j.conb.2021.01.003

Chapter 6 Acknowledgements

I would like to express my heartfelt gratitude to the many people who have supported and accompanied me throughout the course of my doctoral journey.

First and foremost, I would like to express my sincere gratitude to my supervisor, Johannes Fuss, who made this project possible and thereby enabled me to pursue my PhD. His continuous guidance, constructive feedback, and encouragement throughout this dissertation have been invaluable, and his expertise and commitment have had a profound impact on both this work and my academic development. I am also sincerely thankful to Sarah Biedermann for her support in developing and implementing the clinical aspects of the project, as well as for her valuable feedback that contributed substantially to this dissertation. In addition, I would like to thank Peer Birken for his role as a committee member, for providing support and resources that made this project possible, and for his thoughtful feedback that enriched this work.

I also thank my dearest former colleagues at the Institute of Forensic Psychiatry and Sex Research in Essen: Basti, Maria, Juliette, Inga, Michael, Moncsi, and Deborah, who often listened to my project hurdles and progress and provided insightful perspectives and suggestions. I am especially thankful to Susi, who has been not only a great colleague but also a true friend. Your ability to both challenge and support me has been inspiring and has contributed greatly to the completion of this dissertation.

I owe heartfelt thanks to my family for always standing by me and encouraging me to pursue my goals. Particularly, I am grateful to my sister for reminding me of the importance of boundaries and resilience, and to my mother, whose deep support and belief in me have accompanied me throughout this journey.

I am deeply appreciative of my friends, whose companionship and encouragement have carried me through this and many other journeys of my life. I thank Sylvia, whose generosity, humor, and thoughtful advice have supported me in countless ways and always made me feel understood. I am grateful to Joy, who never failed to make me laugh and lift my spirits, whether its citing memes, playing uno, or simply reminding me that everything will be alright. Their friendship has meant more to me than words can express.

I am intensely grateful to Francine and Sophie, who continue to make my life truly special. Francine's spontaneity, unapologetic spirit, and dependability have inspired me and reminded me of the greatness that true friendship brings. Sophie's steady support, thoughtful advice, and the effortless understanding between us over so many years have embodied for me the very essence of a lasting friendship and has been a constant source of strength and encouragement. I am deeply thankful for you both and would not have come this far without our friendship.

Last but definitely not least, I am extremely thankful to my partner Kari, whose presence has been a profound source of stability and encouragement. Their kind love and support in both everyday life and during the most demanding phases, together with their genuine trust in me, was instrumental not only in bringing this work to completion, but also in restoring confidence in myself and everything around me. Thank you so much!

This dissertation would not have been possible without the collective support, patience, and kindness of all these individuals. I am truly thankful.

Chapter 7 Supplementary Material

Study 2: The Impact of Psychopathic Traits on Anxiety-Related Behaviors on the Virtual EPM

Table S1: Pearson correlations between psychopathic traits, behavioral measures on the EPM, subjective anxiety, acrophobia and sensation seeking. Voulgaris et al., Sci Rep 14, 11832 (2024).

	Lack of empathy	**Fearlessness**	**Narcissistic egocentricity**	**Impulsivity**	**Social manipulation**	**Power**	**Latency 1st visit (s)**	**Latency end exploration (s)**	**Time on open arms (s)**	**Number of entries open arm**	**Subjective anxiety**	**Acrophobia**	**Sensation seeking**
Lack of empathy	1	0.235**	0.266**	-0.148	0.001	0.139	-0.198*	-0.124	0.175*	0.297**	-0.060	-0.091	0.203*
Fearlessness		1	0.303**	0.151*	0.326**	0.347**	-0.359**	-0.221**	0.336**	0.299**	-0.232**	-0.283**	0.325**
Narcissistic egocentricity			1	0.269**	0.437**	0.577**	-0.138	-0.157*	0.169*	0.182*	-0.007	-0.032	0.146
Impulsivity				1	0.153*	0.364**	-0.106	-0.194*	0.155	0.104	0.042	0.150	0.107
Social manipulation					1	0.564**	-0.164*	-0.175*	0.151	0.141	-0.097	-0.147	0.241**
Power						1	-0.142	-0.205**	0.184*	0.174*	-0.056	0.023	0.141
Latency 1st visit (s)							1	0.540**	-0.704**	-0.746**	0.563**	0.430**	-0.280**
Latency end exploration (s)								1	-0.645**	-0.451**	0.325**	0.194*	-0.114
Time on open arms (s)									1	0.633**	-0.469**	-0.323**	0.276**
Number of entries open arm										1	-0.397**	-0.299**	0.312**
Subjective anxiety											1	0.460**	-0.201*
Acrophobia												1	-0.290**
Sensation seeking													1

*: Correlation is significant at the 0.05 level (2-tailed)
**: Correlation is significant at the 0.01 level (2-tailed)

Study 4: Validation of the Dark-Light Box in Virtual Reality

Figure S1: Flyer for recruitment of the healthy participant sample in Study 4.

Figure S2: Flyer for recruitment of the patient sample in Study 4 at the Department of Psychiatry and Psychotherapy at the University Medical Center Hamburg-Eppendorf.

Table S2: Fear of Darkness Questionnaire (20 items, unpublished internal document).

Fear of Darkness

Nachstehend sind 20 Punkte aufgeführt, die sich auf das Maß Ihrer Angst beziehen. Bitte beantworten Sie den folgenden Teil in Bezug auf Ihre Gefühle bezüglich Dunkelheit (zum Beispiel: als würden Sie in einem dunklen Raum sitzen oder draußen bei Nacht umhergehen). Jedes Szenario hat vier mögliche Antworten: „manchmal oder einen kleinen Teil der Zeit“, „einen Teil der Zeit“, „einen großen Teil der Zeit“, „die meiste oder die ganze Zeit“. Bitte geben Sie an, welche der Auswahlmöglichkeiten am besten zu Ihrem Empfinden passt, indem Sie den entsprechenden Bereich markieren. Es gibt keine richtige oder falsche Antwort.	**manchmal oder einen kleinen Teil der Zeit**	**einen Teil der Zeit**	**einen großen Teil der Zeit**	**Die meiste oder die ganze Zeit**
1. Ich fühle mich nervöser und ängstlicher als normalerweise.	1	2	3	4
2. Ich fühle mich verängstigt völlig ohne Grund.	1	2	3	4
3. Ich werde leicht verärgert oder fühle mich panisch.	1	2	3	4
4. Ich fühle mich als ob ich auseinander falle und in die Brüche gehe.	1	2	3	4
5. Ich fühle, dass alles in Ordnung ist und nichts Böses passieren wird.	1	2	3	4
6. Meine Arme und Beine zittern und beben.	1	2	3	4
7. Ich bin geplagt von Kopfschmerzen, Nacken- und Rückenschmerzen.	1	2	3	4
8. Ich fühle mich schwach und werde leicht müde.	1	2	3	4
9. Ich fühle mich ruhig und kann problemlos stillsitzen.	1	2	3	4
10. Ich kann mein Herz spüren, wie es schnell schlägt.	1	2	3	4
11. Ich bin geplagt von Schwindelanfällen.	1	2	3	4
12. Ich habe Ohnmachtsanfälle oder fühle mich so, als ob ich welche hätte.	1	2	3	4
13. Ich kann problemlos ein- und ausatmen.	1	2	3	4
14. Ich kriege Taubheits- und Kribbelgefühle in meinen Fingern und Zehen.	1	2	3	4
15. Ich werde von Magenschmerzen oder Verdauungsstörungen geplagt.	1	2	3	4
16. Ich muss meine Blase häufig entleeren.	1	2	3	4
17. Meine Hände sind normalerweise trocken und warm.	1	2	3	4
18. Mein Gesicht wird heiß und errötet.	1	2	3	4

19. Ich kann problemlos einschlafen und kriege eine gute Nachtruhe.	1	2	3	4
20. Ich habe Albträume.	1	2	3	4

Table S3: The Intraclass Correlation Coefficients for behaviors or Behavioral Coding Reliability were evaluated based on absolute agreement. ICC values are interpreted based on Cicchetti (1994) criteria: Poor (<0.40), Moderate (0.40-0.59), Good (0.60-0.74), and Excellent (>0.75). Missing CI values (.) indicate perfect agreement.

Behavior	Measure Type	Single ICC (95% CI)	Interpretation	Average ICC (95% CI)	Interpretation
Arms bent	Duration	0.468 (-0.182, 0.835)	Moderate	0.639 (-0.444, 0.910)	Moderate
	Frequency	0.776 (0.332, 0.939)	Good	0.874 (0.498, 0.969)	Excellent
Bouncing	Duration	0.465 (-0.177, 0.832)	Moderate	0.634 (-0.431, 0.909)	Moderate
	Frequency	0.564 (-0.041, 0.870)	Moderate	0.721 (-0.086, 0.930)	Good
Crossed arms	Duration	0.989 (0.958, 0.997)	Excellent	0.994 (0.979, 0.999)	Excellent
	Frequency	1.000 (., .)	Excellent	1.000 (., .)	Excellent
Crossed arms disguised	Duration	0.993 (0.973, 0.998)	Excellent	0.996 (0.986, 0.999)	Excellent
	Frequency	0.209 (-0.543, 0.733)	Poor	0.346 (-2.373, 0.846)	Poor
Crossed legs	Duration	0.215 (-0.429, 0.721)	Poor	0.354 (-1.505, 0.838)	Poor
	Frequency	0.166 (-0.473, 0.696)	Poor	0.284 (-1.795, 0.821)	Poor
Exploration	Duration	0.665 (0.533, 0.859)	Moderate	0.695 (-0.206, 0.924)	Moderate
	Frequency	0.821 (0.461, 0.951)	Good	0.902 (0.631, 0.975)	Excellent
Fiddling	Duration	-0.006 (-0.608, 0.599)	Poor	-0.013 (-3.107, 0.749)	Poor
	Frequency	-0.044 (-0.578, 0.558)	Poor	-0.092 (-2.740, 0.717)	Poor
Foot edging/foot shifting	Duration	0.962 (0.863, 0.990)	Excellent	0.999 (0.926, 0.995)	Excellent
	Frequency	0.659 (0.143, 0.901)	Moderate	0.877 (0.618, 0.969)	Excellent
Immobile	Duration	0.999 (0.994, 1.000)	Excellent	0.999 (0.997, 1.000)	Excellent

	Frequency	1.000 (., .)	Excellent	1.000 (., .)	Excellent
Looking up	Duration	0.655 (0.133, 0.899)	Moderate	0.792 (0.235, 0.947)	Good
	Frequency	0.877 (0.603, 0.967)	Good	0.934 (0.752, 0.983)	Excellent
One arm crossed	Duration	0.989 (0.958, 0.997)	Excellent	0.994 (0.979, 0.999)	Excellent
	Frequency	1.000 (., .)	Excellent	1.000 (., .)	Excellent
One leg	Duration	0.997 (0.987, 0.999)	Excellent	0.998 (0.993, 1.000)	Excellent
	Frequency	0.883 (0.618, 0.969)	Excellent	0.938 (0.764, 0.984)	Excellent
Out of sight	Duration	0.982 (0.933, 0.995)	Excellent	0.991 (0.966, 0.998)	Excellent
	Frequency	0.635 (0.082, 0.893)	Moderate	0.776 (0.152, 0.944)	Good
Re-examination	Duration	0.984 (0.966, 0.998)	Excellent	0.996 (0.993, 1.000)	Excellent
	Frequency	0.831 (0.476, 0.955)	Good	0.908 (0.645, 0.977)	Excellent
Shuffling feet	Duration	0.981 (0.792, 0.924)	Excellent	0.981 (0.821, 0.902)	Excellent
	Frequency	0.795 (0.250, 0.948)	Good	0.924 (0.792, 0.924)	Excellent

Table S4: Frequencies and relative percentages of all coded behaviors in the Dark-Light Box paradigm (N = 114). Behaviors occurring in fewer than 5% of observations were generally considered rare and excluded. However, conceptually related rare behaviors were collapsed into broader categories to retain information while reducing redundancy "arms before body behaviors", "lower body behaviors").

Behavior	**Count**	**Percent %**
arms bent	118	4.01%
crossed arms	8	0.27%
crossed arms disguised	71	2.41%
one arm crossed	26	0.88%
Arms before body behaviors	***223***	***7.58%***
bouncing	15	0.51%
crossed legs	2	0.07%
foot edging/ foot shifting	130	4.42%
one leg	379	12.88%
shuffling feet	33	1.12%
Lower body behaviors	***559***	***18.99%***
exploration	477	16.21%
fiddling	10	0.34%
immobile	12	0.41%
looking up	1475	50.12%

out of sight	17	0.58%
reexamination	170	5.78%
Total	2943	100.00%

Table S5: T-tests of coefficients for the effects of questionnaire sum scores and subjective anxiety on behavioral variables, with robust standard errors and adjusted *p*-values (Benjamini-Hochberg). Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; . $p \leq 0.1$. Abbreviations: ACQ (Agoraphobic Cognitions Questionnaire), STAI-T (State-Trait Anxiety Inventory: trait), SSSV (Sensation Seeking Scale Form V), IUS (Intolerance of Uncertainty Scale), FOD (Fear of Darkness Questionnaire), CLQ (Claustrophobia Questionnaire), STAI-S (State-Trait Anxiety Inventory: state)."

Effect	Estimate	*SE*	*t*	*p*	Adjusted *p*	sr^2
Looking up(n):(Intercept)	6.273	14.155	0.443	0.659	0.805	0.003
Looking up(n):ACQ	-0.109	0.245	-0.444	0.658	0.765	0.003
Looking up(n):STAI-T	-0.252	0.128	-1.971	0.053	0.238	0.056
Looking up(n):SSSV	0.364	0.186	1.960	0.054	0.238	0.056
Looking up(n):IUS	0.065	0.076	0.858	0.394	0.549	0.011
Looking up(n):FOD	-0.315	0.182	-1.727	0.089	0.277	0.044
Looking up(n):CLQ	0.091	0.077	1.187	0.240	0.468	0.021
Looking up(n):STAI-S	0.101	0.162	0.622	0.536	0.652	0.006
Looking up(n):SubjectiveAnxiety	-0.669	0.423	-1.582	0.119	0.336	0.038
Time in dark(s):(Intercept)	76.479	130.600	0.586	0.560	0.805	0.005
Time in dark(s):ACQ	-2.249	2.096	-1.072	0.287	0.503	0.017
Time in dark(s):STAI-T	-2.014	1.320	-1.526	0.132	0.336	0.035
Time in dark(s):SSSV	3.963	1.770	2.238	0.033	0.178	0.072
Time in dark(s):IUS	0.798	0.701	1.140	0.259	0.470	0.020
Time in dark(s):FOD	-2.30	1.758	-1.307	0.196	0.408	0.027
Time in dark(s):CLQ	0.953	0.711	1.339	0.185	0.408	0.027
Time in dark(s):STAI-S	-0.037	1.353	-0.027	0.978	0.978	0.000
Time in dark(s):SubjectiveAnxiety	-8.927	4.671	1.911	0.060	0.238	0.053

Total distance(m): (Intercept)	-7.105	30.885	-0.230	0.819	0.819	0.001
Total distance(m):ACQ	-0.609	0.536	-1.143	0.260	0.470	0.020
Total distance(m):STAI-T	-0.707	0.320	-2.143	0.036	0.201	0.066
Total distance(m):SSSV	1.213	0.382	3.177	0.002	0.046*	0.134
Total distance(m):IUS	0.022	0.193	0.113	0.911	0.945	0.000
Total distance(m):FOD	0.035	0.560	0.062	0.950	0.968	0.000
Total distance(m):CLQ	0.325	0.209	1.559	0.124	0.336	0.036
Total distance(m):STAI-S	0.654	0.365	1.794	0.078	0.255	0.047
Total distance(m):SubjectiveAnxiety	-2.832	1.032	-2.744	0.008	0.077 .	0.104
Log First Visit(s):(Intercept)	2.104	0.685	3.071	0.003	0.022*	0.127
Log First Visit(s):ACQ	0.019	0.010	1.874	0.065	0.238	0.051
Log First Visit(s):STAI-T	0.019	0.007	2.725	0.008	0.077.	0.103
Log First Visit(s):SSSV	-0.022	0.001	-2.309	0.024	0.169	0.076
Log First Visit(s):IUS	-0.003	0.004	-0.764	0.448	0.582	0.009
Log First Visit(s):FOD	0.016	0.008	1.876	0.065	0.238	0.051
Log First Visit(s):CLQ	-0.011	0.004	-2.734	0.008	0.077.	0.103
Log First Visit(s):STAI-S	-0.007	0.008	-0.864	0.391	0.549	0.011
Log First Visit(s):SubjectiveAnxiety	0.024	0.024	0.996	0.323	0.549	0.015
Log reexamination(n):(Intercept)	-1.096	0.458	-2.390	0.020	0.069.	0.081
Log reexamination(n):ACQ	0.006	0.007	0.924	0.359	0.549	0.013
Log reexamination(n):STAI-T	-0.006	0.005	-1.304	0.197	0.408	0.026
Log reexamination(n):SSSV	0.022	0.006	3.688	0.000	0.026*	0.173
Log reexamination(n):IUS	0.003	0.003	0.843	0.402	0.549	0.011
Log reexamination(n):FOD	0.008	0.008	0.954	0.344	0.549	0.014
Log reexamination(n):CLQ	-0.005	0.003	-1.531	0.131	0.336	0.035

Log reexamination(n):STAI-S	0.002	0.006	0.361	0.719	0.790	0.002
Log reexamination(n):SubjectiveAnxiety	-0.001	0.019	-0.515	0.609	0.725	0.004
Log arms before body(s):(Intercept)	-1.467	1.365	-1.074	0.286	0.668	0.017
Log arms before body(s):ACQ	-0.022	0.029	-0.744	0.459	0.585	0.008
Log arms before body(s):STAI-T	-0.027	0.015	-1.855	0.068	0.238	0.050
Log arms before body(s):SSSV	0.028	0.020	1.388	0.170	0.396	0.029
Log arms before body(s):IUS	0.008	0.009	0.826	0.412	0.549	0.010
Log arms before body(s):FOD	0.016	0.025	0.633	0.529	0.652	0.006
Log arms before body(s):CLQ	0.002	0.013	0.130	0.897	0.945	0.000
Log arms before body(s):STAI-S	0.008	0.019	0.429	0.669	0.765	0.003
Log arms before body(s):SubjectiveAnxiety	0.184	0.059	3.148	0.002	0.046*	0.132
Log lower body behaviors(s):(Intercept)	-0.515	1.287	-0.400	0.690	0.805	0.002
Log lower body behaviors(s):ACQ	-0.015	0.018	-0.827	0.411	0.549	0.010
Log lower body behaviors(s):STAI-T	0.010	0.012	0.846	0.401	0.549	0.011
Log lower body behaviors(s):SSSV	0.035	0.014	2.457	0.017	0.134	0.085
Log lower body behaviors(s):IUS	-0.006	0.006	-0.964	0.339	0.549	0.014
Log lower body behaviors(s):FOD	-0.022	0.016	-1.419	0.161	0.391	0.030
Log lower body behaviors(s):CLQ	0.007	0.006	1.180	0.242	0.468	0.021
Log lower body behaviors(s):STAI-S	-0.004	0.011	-0.380	0.705	0.789	0.002
Log lower body behaviors(s): SubjectiveAnxiety	-0.007	0.039	-0.172	0.864	0.931	0.000

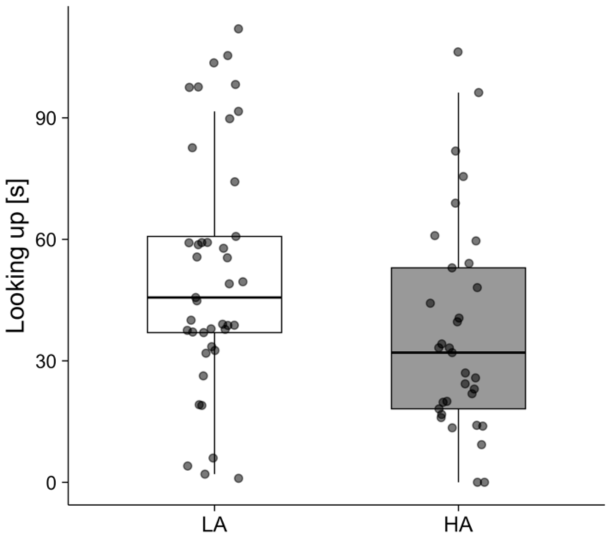


Figure S3: Group comparisons of Looking up duration in the virtual DLB between low-anxiety (LA; n = 41) and high-anxiety (HA; n = 33) participants, classified using a median split of subjective anxiety ratings (1-9, median = 3). Boxplots show medians, interquartile ranges, and individual data points.

Curriculum Vitae

Der Lebenslauf wurde aus datenschutzrechtlichen Gründen entfernt

Der Lebenslauf wurde aus datenschutzrechtlichen Gründen entfernt

Eidesstattliche Erklärung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe, insbesondere ohne entgeltliche Hilfe von Vermittlungs- und Beratungsdiensten, verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe. Das gilt insbesondere auch für alle Informationen aus Internetquellen. Soweit beim Verfassen der Dissertation KI-basierte Tools („Chatbots") verwendet wurden, versichere ich ausdrücklich, den daraus generierten Anteil deutlich kenntlich gemacht zu haben. Die „Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG" aus September 2023 wurde dabei beachtet. Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich damit einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Datum Unterschrift