Psychophysiological measures related to aggressive behaviour and self-harm in people with mild intellectual disabilities:

Potential, pitfalls and future directions

Marlieke van Swieten

Psychophysiological measures related to aggressive behaviour and self-harm in people with mild intellectual disabilities: Potential, pitfalls and future directions

Marlieke J. van Swieten

Cover design: Marlieke J. van Swieten
llustrations for the cover and chapter pages were created at Trajectum
Layout & Printed by HAVEKA

ISBN: 978 90 361 0841 6

Psychophysiological measures related to aggressive behaviour and self-harm in people with mild intellectual disabilities: Potential, pitfalls and future directions

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 23 januari 2026
om 10.30 uur precies

door

Marlieke Johanna van Swieten
geboren op 12 september 1994
in Leidschendam

Promotor: Prof. dr. H.C.M. Didden

Copromotoren: Dr. P.C. de Looff
Dr. J. van der Nagel (Universiteit Twente)

Manuscriptcommissie: Prof. dr. J.I.M. Egger
Prof. dr. R. Otten
Prof. dr. M. Keulen-de Vos (Maastricht University)
Prof. dr. L. Nauta-Jansen (Amsterdam UMC)
Dr. Y. Derks (GGNet)

TABLE OF CONTENTS

Chapter 1

General introduction

Jason[1]

Jason, a 28-year-old man, is currently living at Trajectum, a secure forensic psychiatric treatment facility. He is undergoing treatment following his conviction and prison sentence for a violent robbery and extortion. He has been diagnosed with borderline intellectual functioning, an antisocial personality disorder and cannabis use disorder. His early life was marked by significant adversity, including financial instability, parental alcohol addiction, and neglect.

At the age of 12, Jason began using cannabis, initially in a social context with peers but later as a means of self-medication to reduce negative thoughts and stress. By the age of 15, he dropped out of secondary school. His first police contact occurred during adolescence, and he subsequently engaged in repeated minor offenses. By the age of 26, his criminal behaviour had escalated, culminating in his conviction for the aforementioned violent crimes.

Jason's current treatment plan is multifaceted and includes trauma-focused therapy, cognitive behavioural anger management treatment (called 'Grip op Agressie', GoA), skills training, and pharmacotherapy. These interventions are embedded in a therapeutic group climate, which emphasises daily routines and meaningful activities in a secure environment. He is talented in woodworking, which he practices on weekdays in the facility's workshop. To assess and manage his risk of aggressive behaviour and recidivism, his treatment team conducts regular evaluations using structured risk-assessment tools such as the Dynamic Risk Outcome Scales (DROS) and the HKT-R scale (Historisch, Klinisch, Toekomst – Revisie [Historical Clinical Future-Revised]). Jason shows commitment in his treatment program and is motivated to change his behaviour. His coping skills improved during the GoA training and aggressive behaviour has decreased. However, incidents of both verbal and physical aggressive behaviour still occur on the ward. It is believed that these incidents are often caused by increased stress in combination with limited interoceptive awareness, and inability to generalize learned adaptive coping skills. Caregivers report that some of these incidents are preceded by observable signs of heightened tension; other incidents, however, arose without apparent warning.

1 Jason is a fictitious case build on several cases involved in this thesis.

INTRODUCTION

Aggressive behaviour and self-harm are a great concern for many treatment and assisted living facilities for people with mild intellectual disabilities or borderline intellectual functioning (MID-BIF), and have detrimental consequences for both clients and caregivers. Psychophysiological measures with wearable technology offer innovative opportunities to understand mechanisms and processes preceding and following aggressive behaviour and self-harm from a biopsychosocial perspective. These measures might also be useful for supporting more personalized interventions and to study the effectiveness of interventions.

This thesis focuses on the relation between physiological arousal (e.g., heart rate), aggressive behaviour and self-harm. First, we explored the association between aggressive behaviour, self-harm and shared risk factors in this population from a psychosocial perspective. Second, a broad overview of literature on the relation between physiological arousal and self-harm is provided. Third, using wearables we investigated changes in physiological arousal after aggressive behaviour and self-harm in the living group. Finally, we tested an application of wearables for stress-reduction in daily life.

This introductory chapter starts with a short description of the group of individuals with whom this research was conducted. Then, the rationale of the studies is provided, including definitions, frequencies and the impact of aggressive behaviour and self-harm. Different (psychosocial) models and theories for understanding aggressive behaviour and self-harm are described, followed by a focus on arousal and wearable technology as this is an innovative and increasingly important area of research in relation to aggressive behaviour and self-harm. The introduction section concludes with a description of the thesis layout.

People with mild intellectual disabilities or borderline intellectual functioning

People with mild intellectual disabilities or borderline intellectual functioning encounter serious challenges in multiple domains of life. Intellectual developmental disorder (intellectual disability) is defined in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR) as a disorder with an onset during the developmental period, and is characterised by intellectual and adaptive functioning deficits in conceptual, social, and practical domains (American Psychiatric Association 2022). Four levels of severity of intellectual disability are distinguished: mild, moderate, severe and profound. As a shift away from traditional views which defined severity levels based solely on IQ (intelligence quotient), the DSM-5-TR defines the level of severity on the basis of adaptive functioning, and not on IQ scores alone. Adaptive functioning comprises a conceptual, social and practical domain (American Psychiatric Association, 2022). In the conceptual domain, adults with a mild intellectual disability (MID) have impaired abstract thinking, executive function, short-term memory, as well as impairments in the functional use of academic skills (e.g., reading, money management; Schalock, 2011). In the social domain people with MID can experience problems in social interactions, with difficulties in perceiving social cues and understanding complex social situations. Furthermore, they can have difficulties in regulating emotions and behaviour, understanding risks in social situations and are at risk of being manipulated (Wagemaker et al., 2020). In the practical domain, people with MID can function in daily life with some support but may struggle with relatively complex tasks such as shopping, paying bills, taking care of their children, or taking public transportation (Tassé et al., 2012). Furthermore, people with MID are less often employed (Landgren et al., 2024) and are at higher risk of substance use-related problems (Påhlsson-Notini et al., 2024). The percentage of adults with a MID is estimated at 2.1 (however, this estimation is only based on IQ; Woittiez et al., 2019).

Besides intellectual disability, the DSM-5-TR includes the V-code borderline Intellectual functioning. This is a category that may be used when borderline intellectual functioning is the focus of clinical attention or has an impact on the individual's treatment or prognosis. The DSM-5-TR does not provide criteria regarding IQ range but an IQ range of 70 to 85 is often reported in literature (Hassiotis et al., 2022). Approximately 30% of adults with an IQ in this range experiences adaptive functioning deficits in conceptual, social, and practical domains, resulting in an estimation of 4.1% of Dutch adults meeting the description of borderline intellectual functioning (BIF) (Woittiez et al., 2019). Persons with BIF are at increased risk of experiencing physical problems, poverty and have limited social support. Several studies show increased risk for the development of psychiatric disorders, including substance abuse and personality disorders (Gigi et al., 2014; Hassiotis et al., 2008; Nieuwenhuis et al., 2021; Peltopuro et al., 2014). In the Netherlands, people with BIF and comorbid psychiatric disorders have access to the same specialized mental health care facilities as people with MID and are therefore often studied as one group (i.e., MID-BIF). Trajectum is an example of an organization in the Netherlands, and part of De Borg, providing (forensic) treatment and care to people with MID-BIF. It served as the setting for the research presented in this thesis (for more information, see Box 1).

Box 1 Trajectum and De Borg
De Borg organizations provide specialized care at the intersection of forensic care, mental health care, and disability care. De Borg is a cooperation of four facilities in the Netherlands: Trajectum, Ipse de Bruggen, Stevig Dichterbij and Fivoor. Clients are placed in the facility under criminal law, civic law or on a voluntary basis. Aggressive behaviour is often one of the main reasons for referral (Delforterie et al., 2020). Persons without offences but who are a danger for themselves or show self-destructive behaviour are also admitted. Besides MID-BIF and complex behavioural or psychiatric disorders clients often have an (history of) addiction to substances (van Duijvenbode et al., 2015), experience problems in multiple areas of life such as relationships with family, partners or children and have financial problems. All four facilities of De Borg provide treatment and care aimed at a diverse range of behavioural and psychiatric disorders, as well as the prevention of recidivism. The average treatment duration at Trajectum for example is 2 years and 2 months (Neimeijer et al., 2021). Treatment is divided into stages with gradually reducing security levels (with clients typically moving from high security settings to open units) while offering a combination of cognitive-behavioural programmes, art, drama-, and psychomotor therapy, occupational therapy, skills training and pharmacotherapy. Treatment programmes are embedded in a therapeutic group climate with a daily structure, activities and care that is adapted to the emotional, adaptive and intellectual abilities of clients (Neimeijer et al., 2021). The study described in Chapter 2 was conducted at Trajectum and the studies in Chapter 4 and 5 were conducted at all facilities of De Borg.

Aggressive behaviour: definition, prevalence and impact

There are many different definitions for aggressive behaviour (Rippon, 2000; Warburton & Anderson, 2015). For example, aggressive behaviour has been defined as: "a range of behaviours or actions that can result in harm, hurt or injury to *another* person, regardless of whether the violence or aggression is physically or verbally expressed, physical harm is sustained or the intention is clear" (p. 6; National Collaborating Centre for Mental Health, 2015). Other studies have also included auto-aggression (self-harm) as a subtype of aggressive behaviour (Crocker et al., 2006; De Looff et al., 2019). The World Health Organization uses a broad definition of a severe form of aggressive behaviour, namely violence: ''the intentional use of physical force or power, threatened or actual, against *oneself, another person, or against a group or community*, that either results in or has a high likelihood of resulting in injury, death, psychological harm, maldevelopment, or deprivation'' (Krug et al., 2002). There is currently no consensus on the definition of aggressive behaviour, with varying definitions applied depending on the theoretical framework, study design, and setting. This variability is also reflected in the present thesis and will be elaborated upon in the subsequent chapters.

Aggressive behaviour is relatively common in many treatment and assisted living facilities. A systematic review found that the prevalence of any form of aggressive behaviour towards staff members of psychiatric wards (for adults with or without intellectual disabilities) throughout their careers ranged from 65% to 99%, with 38% to 82% having experienced physical aggressive behaviour (Weltens et al., 2021). The percentages of clients that were involved in aggressive behaviour on psychiatric wards ranged from 7.5% up to 75.9%, with a weighted mean of 23% (Weltens et al., 2021). Compared to the general population, people with intellectual disabilities and people with psychiatric disorders (e.g. affective disorders and personality disorders) are at an increased risk for engaging in aggressive behaviour (Girasek et al., 2022; Nieuwenhuis et al., 2022). In institutional settings for people with intellectual disabilities, aggressive behaviour is the most common reason for admittance and re-admittance (Delforterie et al., 2020; Didden et al., 2019). For example, in Trajectum, one of De Borg facilities (see Box 1), 65% (n = 161) of the clients were involved in a total of 1003 aggressive incidents during a one-year study period (Neimeijer et al., 2021).

Aggressive behaviour may have severe negative psychological and physical consequences for clients engaging in this behaviour, other clients as well as caregivers. Aggressive behaviour often leads to seclusion (Cowman et al., 2017; Vruwink et al., 2022) and it can disrupt the rehabilitation of the client because of coercive measures, conviction, and transfer to another facility (Neimeijer et al., 2021). Victims of aggressive behaviour are at risk of negative psychological effects (e.g. anxiety, sleep disturbance, fear and anger) and physical injury (Knotter, 2019). For caregivers, clients' aggressive behaviour can lead to an unsafe working environment, feelings of anger and fear, posttraumatic stress symptoms and physical injury (Verstegen et al., 2024). Caregivers exposed to more frequent or severe aggressive behaviour report more sick leave and more burnout symptoms (De Looff et al., 2018; Hensel et al., 2014; Kind et al., 2018). This can increase staff shortage, workload, working overtime and impede an organization's ability to provide continuous high-quality care (Brown & Kalaitzidis, 2013; Fanneran et al., 2015; Luther et al., 2017; Robson & Attard, 2019).

Self-harm: definition, prevalence and impact

Self-harm is an umbrella term that includes both self-injurious behaviour with the intent to end one's life and self-injurious behaviour without the intent to die (Shafti et al., 2021). The latter is called non-suicidal self-injury (NSSI). In the DSM-5-TR (American Psychiatric Association, 2022) NSSI is defined as behaviour involving intentionally inflicting damage to one's body that will likely induce bleeding, bruising or pain. Self-harm is estimated to affect 14.6 million individuals each year (Global Burden of Disease Collaborative Network, GBD, 2019). A recent meta-analysis on self-harm in adolescents in different countries reported an estimated global lifetime prevalence of 19% (Lucena et al., 2022). Compared to the general population people with intellectual disabilities and people with psychiatric disorders (e.g. affective disorders and personality disorders) are at an increased risk for engaging in self-harm (Bøe et al., 2022; Calver et al., 2021; Flygare Wallén et al., 2023). Similar to aggressive behaviour, self-harm is also associated with multiple adverse consequences for the client engaging in self-harm, other clients and caregivers. Self-harm causes physical harm (Matson & Turygin, 2012), is associated with feelings of regret, embarrassment and/or shame (Muller et al., 2020) and an increased risk of suicide (Carroll et al., 2014). Furthermore, witnessing or hearing about this behaviour may trigger self-harm in other clients (James et al., 2012). It may also affect emotional wellbeing of caregivers as it can lead to feelings of anger, inadequacy and guilt in caregivers (Fish, 2000). They may also feel fearful, overwhelmed and powerless when witnessing self-harm (James et al., 2012).

Understanding aggressive behaviour and self-harm

Aggressive behaviour has been studied from many perspectives (also see, Warburton & Anderson, 2015). Through the years different theories have been developed to understand aggressive behaviour such as psychoanalytic and psychodynamic approaches (Freud, 1915; Teodosio & Ugo, 2019), the frustration-aggression hypothesis (Dollard et al., 1939), learning theories (Eron et al., 1971), cognitive theories such as social information processing (Dodge, 1980) and script theories (Huesmann, 1982), and the cognitive neoassociation theory (Berkowitz, 1989). These theories mainly included psychosocial components and do not, or only to a limited extend, include biological aspects. One of the broadest and most often cited models of aggressive behaviour to date is the general aggression model (GAM; Anderson & Bushman, 2002, 2018). The GAM integrates several theories into an integrative framework describing the role of both personal and situational variables on aggressive behaviour, ranging from biological to cultural variables. Although the GAM is often cited and used as a theoretical framework it is not without criticism (Ferguson & Dyck, 2012). It has been described as being too narrow in scope, and focusing to a large degree on social learning and the social cognitive paradigm.

To date, the development and expression of aggressive behaviour is generally seen as a multifactorial process with a complex interplay of (dynamic and static) individual, situational, environmental and biological factors risk factors explaining this behaviour (Cohen & Tsiouris, 2020; Weltens et al., 2021). Similarly, the development and expression of self-harm is also seen as a multifactorial process with a complex interplay of (dynamic and static) risk factors (Hulsmans et al., 2024). Research showed that the same set of risk factors (e.g. heightened impulsivity and impaired coping skills) are associated with both behaviours, which is known as multifinality (O'Donnell et al., 2015).

While aggressive behaviour and self-harm are considered different types of behaviour, research across various populations has consistently shown that these behaviours frequently co-occur, which is known as 'dual-harm' (Slade et al., 2020). For example, the study by Verstegen et al. (2020) reported that inpatients of a forensic psychiatric hospital with self-harm had an almost nine times higher risk of exhibiting physical aggressive behaviour than patients without self-harm. Shafti et al. (2021) presented the cognitive-emotional model of dual-harm, a theoretical framework that accounts for why people may engage in dual-harm. This model states that a combination of biological factors and adverse environmental factors result in the development of a certain personality style. The personality traits predispose the person to dual-harm through its effects on cognition, arousal and affect. Dual harm can have an interpersonal function or emotion regulating function. The type of behaviour (i.e., aggressive behaviour versus self-harm) is influenced by the social context and expectancies the person has regarding the harmful behaviour, which are influenced by earlier effects of the behaviour. Shafti et al. (2021) concluded that the proposed model should be tested across different populations to assess its generalisability. Unfortunately, clients with MID-BIF are often not included in studies on the association between self-harm and aggressive behaviour (dual harm) or only form a small minority of the study's sample. Furthermore, Verstegen et al. (2020) stated that more research is needed in forensic settings, which should further explore if there are shared risk factors for self-harm and aggressive behaviour.

Arousal, aggressive behaviour and self-harm

Besides affect and cognition, certain arousal characteristics may form a predisposed risk to both aggressive behaviour and self-harm, as described in the cognitive-emotional model of dual-harm. One useful theoretical model to understand the association between stressors and the stress response is a dynamic systems model (Everly & Lating Jr, 2019). Stressors are mediated by physiological mechanisms, and may result in increased arousal by activation of target organ systems (e.g. cardiovascular system, the gastrointestinal system, the skin, the immune system; Everly & Lating Jr, 2019). Altered (increased or decreased) arousal has repeatedly been linked to aggressive behaviour and self-harm (Bellato et al., 2023; Cohen & Tsiouris, 2020; Hooley & Franklin, 2018; Warburton & Anderson, 2015). A meta-analysis on the association between aggressive behaviour and physiological arousal even reported that some effect sizes of physiological measures of arousal were larger than several psychosocial risk factors, suggesting that integrating psychosocial and biological information could be useful in determining treatment strategies for individuals with aggressive behaviour (De Looff, Cornet, et al., 2022b). Implementing biopsychosocial information is seen as an important direction for future research into aggressive and other antisocial behaviour (Jansen, 2022). Increased arousal may play a particularly significant role in the development of aggressive behaviour and self-harm in people with MID-BIF because they are at increased risk to experience (high levels of) stress resulting in increased arousal, compared to the general population (e.g., Forte et al., 2011; Griffith et al., 2013; Rouleaux et al., 2024). In addition, living in a residential setting is associated with multiple stressors (Griffith et al., 2013; Neimeijer et al., 2021). When confronted with stressful situations, people with MID-BIF are more prone to use maladaptive coping strategies in comparison to people without MID-BIF (Hartley & MacLean Jr, 2008; Plotner et al., 2020). Deficits in intellectual functioning can impair the development of adaptive strategies to cope with stress and with which arousal can be reduced (Didden et al., 2019; Everly & Lating Jr, 2019; Taylor & Novaco, 2005).

Arousal was originally and still is often measured using subjective measures such as self-reports or observational studies. However, the relation between arousal and aggressive behaviour and self-harm has increasingly been studied by arguably more objective physiological measures (see below). Multiple systematic reviews and meta-analyses provided an overview of the literature on the relation between physiological arousal and aggressive behaviour (De Looff, Cornet, et al., 2022b; Lorber, 2004; Ortiz & Raine, 2004; Portnoy & Farrington, 2015). Until recently, no overview on the relation between self-harm and physiological arousal existed. Goreis et al. (2023) conducted a meta-analysis on specifically physiological stress reactivity and self-harm, while Bellato et al. (2023) conducted a review and meta-analysis on the autonomic dysregulation and self-harm thoughts and behaviours in children and young people. These meta-analyses found evidence for significantly altered stress reactivity (e.g. lower heart rate variability in rest and recovery) in people who self-harm and reduced heart rate variability in children and young people who self-harm, respectively. However, both meta-analyses excluded case studies and included mainly laboratory studies. Much less is known about the relationship between self-harm and physiological parameters when measured in natural environments, as well as about physiology related to actual or imagined self-harm incidents. Moreover, persons with (moderate, severe and profound) intellectual disabilities were not represented[2] in these studies. It is important to include persons with intellectual disabilities because of the relatively high prevalence of self-harm in this population. In addition, some differences in self-reported reasons of self-harm have been found in persons with intellectual disabilities compared to people without intellectual disabilities (Samways et al., 2024), which might be associated with different physiological patterns related to self-harm.

2 The included studies often excluded persons with intellectual disabilities or did not report on intellectual functioning and recruited from the community, students or outpatient and inpatient clinics not specialized at people with intellectual disabilities. It is therefore unknown if people with BIF were included.

Measuring physiological arousal

Arousal can be measured based on physiological parameters reflecting the activation of the autonomic nervous system (ANS). The ANS can be subdivided into two branches, the sympathetic and the parasympathetic branch. The sympathetic branch prepares the body for action while the parasympathetic branch of the ANS regulates restorative functions and relaxation of the body (Everly & Lating Jr, 2019). Many physiological parameters that reflect ANS functioning can be measured (for an overview see Behnke et al., 2022). Heart rate (HR) and electrodermal activity (EDA) are most often studied, but also heart rate variability (HRV) is receiving increasing interest (De Looff, Cornet, et al., 2022b) (See Box 2 for more information about these physiological parameters). In the present thesis, arousal is measured with HR, HRV and EDA. The stress response also has an effect on other target organs (e.g. pupil dilation), but these effects are, to varying degrees, more difficult or uncomfortable to measure in daily life.

The golden standard for measuring HR is electrocardiography (ECG) which involves placing multiple electrodes on the skin, connected by wires to a monitor (Nezamabadi et al., 2022). For instance, The VU-AMS is a valid and reliable lightweight ECG device for ambulatory assessment (De Geus et al., 1995). Although it provides opportunities for ambulatory measurements in real-life contexts, application of the electrodes and setup of the device needs to be done by an expert (Schuurmans et al., 2020), and can be quite tedious. In recent years wearable technology has been developed, using for example photoplethysmography with a wristband to measure HR, enabling simpler, continuous and less invasive measurement of physiological parameters of the ANS. It can be used by both researchers and lay-people (Schuurmans et al., 2020), outside a laboratory environment and in daily life (Johnson & Picard, 2020). Wearable devices therefore might be an important solution for measuring physiology continuously and noninvasively in daily life, for example in relation to aggressive behaviour and self-harm. However, physiological measurements from daily life are more prone to artefacts and typically result in an increased amount of missing data in comparison to golden standard reference devices (Boucsein, 2012; De Looff et al., 2019; Van der Mee et al., 2021). Wearable devices can be worn by individuals and exist in different forms such as: wristwatches (see Figure 2 for an example of a wristwatch which was used for the study described in Chapter 5), glasses, chest straps, rings, and prosthetic sockets (Sabry et al., 2022). These devices can measure a variety of health-related factors such as HR[3], acceleration or sweat, but are increasingly being used to estimate 'stress' levels, sleep, cognitive load, or breathing, see Figure 3. To measure these human signals, wearables often feature a variety of sensors, such as temperature sensors, movement sensors (via accelerometers), optical sensors, and biometric sensors (Sabry et al., 2022). It is however important to note that not all measurements and estimates are equally useful and reliable (Altini, 2024).

3 Also referred to as Pulse rate when measured on the wrist.

Box 2 Explanation of physiological parameters of the autonomic nervous system

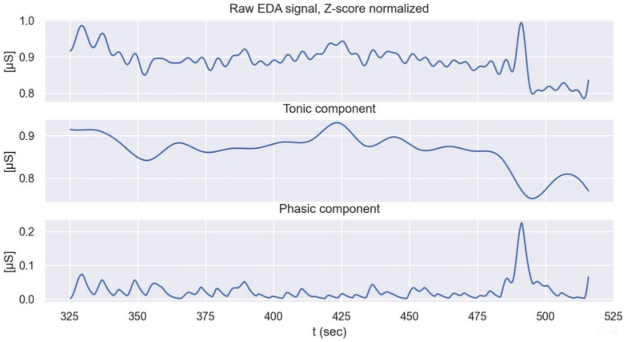
Heart rate (HR) is the frequency of contractions (beats) of the heart per minute. It is a measure of autonomic arousal, associated with dually innervated SNS and PNS activity. Increase in HR is an indicator of increased motivational intensity, action readiness, and engagement (Behnke et al., 2022).

Heart rate variability (HRV) includes changes in the time intervals between consecutive heartbeats. High HRV is an indicator of healthy regulatory systems that can effectively adapt to sudden environmental and psychological challenges (Pham et al., 2021). Many different parameters of HRV exist, such as respiratory sinus arrhythmia (RSA). This includes the variation in duration of consecutive interbeat intervals that follow respiratory cycles. RSA is associated with PNS activity. High (but not excessive) RSA reflects greater PNS regulation of HR and is associated with adaptive social and emotional functioning, positive mental health outcomes, and strong emotion regulation (e.g. Beauchaine, 2015).

Electrodermal activity (EDA) includes changes in the electrical conductivity of the skin. When the body responds to a stressor, the SNS activates the sweat glands, causing an increase in EDA (Pattyn et al., 2023). EDA has a tonic and phasic component (Boucsein, 2012), see Figure 1. The tonic component varies slowly and is referred to as skin conductance level (SCL). The phasic component reflects rapid responses following a stimulus and is known as skin conductance response (SCR).

Figure 1

Example of Decomposed Tonic and Phasic Components from Raw EDA Signal



Note. From 'To Drive or to Be Driven? The Impact of Autopilot, Navigation System, and Printed Maps on Driver's Cognitive Workload and Spatial Knowledge', 2021, by Brishtel et al., International Journal of Geo-Information, p. 12, (https://doi.org/10.3390/ijgi10100668). CC-BY 4.0.

Figure 2

An Example of a Wearable Technology in the Form of a Wristwatch: The Embrace-Plus

Figure 3

Summary of Forms of Wearables and Targeted Measures

FIGURE 1 | Summary of current technologies for monitoring health/performance and targeted physical measurements.

Note. *From 'A Critical Review of Consumer Wearables, Mobile Applications, and Equipment for Providing Biofeedback, Monitoring Stress, and Sleep in Physically Active Populations,' 2018, by Peake et al., Frontiers in Physiology, p. 3 (https://doi.org/ 10.3389/fphys.2018.00743). CC-BY 4.0.*

For a variety of reasons, wearable technology might be especially promising for obtaining arousal related measurements such as HR and EDA (in relation to aggressive behaviour and/or self-harm) in people with MID-BIF. For instance, it has been widely observed in clinical practice that people with MID-BIF often have reduced interoceptive awareness, implying a reduced ability to perceive internal bodily sensations (e.g. Bellemans et al., 2022). This impairs the ability of using self-report measures of arousal. Moreover, informant reports based on observations are less suitable for measuring arousal because arousal is an internal state, and caregivers are not consistently present in close proximity to clients which limits their ability to adequately monitor arousal in their clients (Noordzij & Laroy-Noordzij, 2013). Indeed, caregivers often report that they are surprised by the clients' aggressive behaviour or self-harm and did not detect a visible build-up in arousal preceding the behaviour, as was the case with Jason (Bowers et al., 2011; Simons et al., 2021). Studies using wearable technology could enhance the validity of measuring arousal (in terms of HR and EDA) in relation to aggressive behaviour and self-harm in people with MID-BIF.

Despite advantages of wearable technology, most studies to date studied the relationship between aggressive behaviour, self-harm and physiological parameters in laboratory environments (with community samples). These studies for example investigate physiology during experiments simulating mild aggressive behaviour such as applying noise bursts or mild shocks to participants or compare physiology during cognitive or social stress tasks in groups with and without a history of self-harm (for an overview see; De Looff, Cornet, et al., 2022b; Goreis et al., 2023). The conclusions of these studies are often heterogeneous and cannot be generalized to actual aggressive behaviour or self-harm in daily life situations. Generalization is limited because these studies typically compare various antisocial, violent, aggressive or self-harm groups with different controls under very specific conditions in diverse laboratory tasks. These laboratory tasks (that tap into cognition, emotion or social domains) are not easily translated to practice. Measures with wearables from daily life instead of laboratory studies may increase ecological validity in studying the relationship between arousal and aggressive behaviour and self-harm.

An important innovative study, conducted at forensic psychiatric wards, is the work by De Looff et al. (2019) which examined physiological measures (HR, SCL) in relation to actual occurrences of aggressive behaviour in clients with MID-BIF. Results showed a significant increase in HR and SCL 20 minutes on average before aggressive behaviour occurred, showcasing the potential of the technology for the prediction of aggressive behaviour. However, very little is known about (changes in) physiological measurements/arousal *after* aggressive behaviour and self-harm. Knowledge about physiological arousal after aggressive behaviour and self-harm may increase our understanding of mechanisms of increasing and decreasing arousal. It might help in preventing recurrent incidents and could possibly guide the selection of interventions based on arousal patterns surrounding an incident.

Wearables and arousal reduction

Besides assessment of aggressive behaviour and self-harm wearables might be an effective addition to interventions aimed at (early) detection and reduction of arousal, thereby possibly preventing aggressive behaviour and self-harm. Wearables can be used to provide biofeedback, giving clients and/or their caregivers real-time insight into arousal levels. This may aid clients and caregivers to effectively detect and manage stress responses (De Looff et al., 2024; De Vries et al., 2023). To manage stress, different strategies can be used by a person experiencing the stress, such as listening to music. Music listening can be used as an accessible and low-risk strategy for the reduction of subjective and physiological stress (De Witte et al., 2020; Kume et al., 2017; Lynar et al., 2017). However, limited research is available on the effect of listening to music on physiological stress in people with MID-BIF.

Studies investigated the effect of participant selected music and music selected by the researcher on subjective and physiological stress-reduction. Alternatively, music could also be selected by innovative music technology. X-system is such technology and can be used to select songs to reach a desired state of arousal (i.e. activation or relaxation). X-system offers a web-application that predicts the effect of a song on levels of physiological arousal (Osborne et al., 2017). Nijman et al. (2023) studied the differential effect of listening to music in a preferred genre selected and ordered by X-system versus music in random order on stress reduction in clients and caregivers of a medium secure forensic psychiatric facility. This study showed that physiological indices and self-reported stress decreased after listening to music. An accelerated reduction in SCL for the X-system playlist compared to the playlist in random order was found with visual inspection of the data, but the trend was non-significant. However, this study was carried out in a controlled setting, where participants listened to music on set times in a quiet room. This might limit the representativeness and ecological validity of the findings regarding daily life effects of listening to music when feeling stressed.

Aims and outline of the thesis

Aggressive behaviour and self-harm are common in people with MID-BIF in residential (treatment) facilities and have a negative impact on both the client, other clients as well as caregivers. Often, the assessment and treatment focus mainly on psychological and social aspects. However, the crucial role of the biological component of the biopsychosocial model in the development of this behaviour has been underscored by several studies. The development of wearable technology created an opportunity to study physiological parameters in the daily life of people with MID-BIF. The aim of the present thesis is to gain more insight into the role of physiology in the assessment and (evaluation of) interventions of aggressive behaviour and self-harm in people with MID-BIF, and to gain insights in the potential, pitfalls and future directions of wearables in this field. We conducted several studies which resulted in the following chapters for the current dissertation.

In **Chapter 2** the association between aggressive behaviour and NSSI in 125 adults with MID-BIF was investigated using incident reports and standard clinical measurements. Furthermore, because knowledge on the co-occurrence does not provide sufficient information and direction for clinicians to treat and prevent self-harm and aggressive behaviour, we investigated whether heightened impulsivity and limited coping abilities (elements of personality and emotional dysregulation, which are components of the cognitive-emotional model of dual-harm) are shared risk factors of aggressive behaviour and NSSI in this sample.

Chapter 3 includes a systematic review of studies on the association between physiological parameters and self-harm. It provides an overview of the knowledge on the association between self-harm and physiological parameters by systematically summarizing a broad spectrum of studies. Research gaps and directions for future research on the relation between self-harm and physiology are described.

In **Chapter 4** changes in HR and EDA in the 30 minutes *following* aggressive behaviour (and self-harm) were studied with a large multicentre study. In this chapter, arousal, measured on the ward with wearable technology, after 165 aggressive incidents was analysed. Data were collected at eight organizations providing care to children, adolescents and adults (with MID-BIF) and comorbid psychiatric and behavioural disorders. We investigated whether physiological arousal decreased significantly within the 30 minutes following aggressive behaviour. We exploratively assessed whether we could identify unobserved (latent) classes which follow similar growth patterns to better capture the variability in physiological arousal following aggressive incidents. Finally, we explored whether type of aggressive behaviour could predict the trajectory of physiological changes following incidents.

In **Chapter 5** a clinical application of wearable technology was investigated. With a biofeedback study we explored whether listening to music is associated with reduced physiological and subjective stress in people with MID-BIF. We collected 103 music listening sessions over a period of 2-4 weeks for 11 participants throughout their daily routines. They listened to music when they received biofeedback on their increased stress level (as measured by wearable biosensor Nowatch) or when they themselves felt stressed. Participants listened either to self-selected music or to a personalised playlist compiled with X-system. This study is the first to examine changes in physiological stress during listening to music based on an innovative music selection system in people with MID-BIF in a naturalistic setting.

Finally, in **Chapter 6** the general discussion of this dissertation summarises and reflects on the main findings, and discusses the directions and implications both for research and clinical practice.

Chapter 2

The association between aggressive behaviour and non-suicidal self-injury and shared risk factors in adults with mild intellectual disability

This chapter was published as:

van Swieten, M., de Looff, P., VanDerNagel, J., & Didden, R. (2024). The association between aggressive behaviour and non-suicidal self-injury and shared risk factors in adults with mild intellectual disability. *Journal of Applied Research in Intellectual Disabilities, 37*(6), e13288. https://doi.org/10.1111/jar.13288

ABSTRACT

Background: Aggressive behaviour (AB) and non-suicidal self-injury (NSSI) are common in people with mild intellectual disability or borderline intellectual functioning, leading to adverse consequences for themselves and those around them.
Method: We investigated the relationship between AB (both total and physical in particular) and NSSI and risk factors in 125 residents in a treatment clinic using incident reports and standard clinical measurements.
Results: There was a weak correlation between AB and NSSI, as well as between impulsivity and total AB, and between coping and AB and NSSI. However, NSSI, impulsivity and coping skills did not predict AB.
Conclusion: Results do not corroborate those of other studies in this area. In future studies impulsivity, coping, aggression and NSSI may be measured using other instruments, and differences between people with and without intellectual disability regarding these variables may be explored.

INTRODUCTION

Aggressive behaviour (AB) and self-harm are common in people with mild intellectual disability or borderline intellectual functioning (IQ 50 – 85), especially in forensic care facilities (Dixon-Gordon et al., 2012; Neimeijer et al., 2021). AB often has a negative impact on both the client who displays AB, other clients and caregivers. For clients, AB may reduce possibilities for their rehabilitation, lead to seclusion or may result in transfer to another facility, or even conviction (Griffith et al., 2013; Lokman et al., 2022). Other clients or victims of aggression are at risk of negative psychological effects (e.g. anxiety and anger) and physical injury (Pelto-Piri et al., 2020). For caregivers, exposure to AB can result in: an unsafe working environment, more sick leave and an increased risk for burn-out symptoms (De Looff et al., 2018; Hensel et al., 2014).

Self-harm has many definitions in the literature. The current study differentiates between self-harm with and without suicidal consequences. Self-harm is an umbrella term that encompasses both self-injurious behaviour with intent to end one's life and self-injurious behaviour without intent to die (Shafti et al., 2021). In the Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR; American Psychiatric Association, 2022) non-suicidal self-injury (NSSI) is defined as the behaviour of intentionally inflicting damage to one's body that will likely induce bleeding, bruising or pain. Self-harm is associated with multiple adverse consequences for both the client with self-harm, other clients and caregivers. For clients, self-harm causes physical harm (Matson & Turygin, 2012) and NSSI in particular is associated with feelings of regret, embarrassment and/or shame (Muller et al., 2020). Witnessing or hearing about self-harm may trigger self-harm in other clients (James et al., 2012). For caregivers, it may affect their emotional wellbeing and may lead to feelings of anger, inadequacy, guilt (Fish, 2000), fear and powerlessness (James et al., 2012).

While AB and self-harm are different types of behaviour, research across various populations has consistently shown that these behaviours frequently co-occur, which is called 'dual-harm' (Slade et al., 2020). A systematic review including 123 studies found elevated levels of AB in people with self-harm compared with people without self-harm (O'Donnell et al., 2015). An observational study by Verstegen et al. (2020) explored which variables could predict physical violence in 614 inpatients of a forensic psychiatric hospital. They included gender, legal status, borderline personality disorder, antisocial personality disorder, schizophrenia spectrum disorder, psychopathy, self-harm, impulsivity, intellectual disability and length of stay in the analysis. Results showed a clear association between self-harm and inpatient physical AB on all outcome measures and in all analyses. Participants with self-harm had an almost nine times higher risk of exhibiting physical aggression than clients without self-harm. A review study on people with intellectual disability found a positive association of self-harm with total aggression but found contradictory results on the association between self-harm and physical AB (van den Akker et al., 2021).

However, evidence of co-occurrence does not provide sufficient information and direction for clinicians to treat and prevent self-harm and AB. Therefore, it is important to gain knowledge of the nature or source of the association between self-harm and AB. Research showed that the same set of risk factors are associated with both behaviours, which is known as multifinality (O'Donnell et al., 2015). Shared risk factors reported in literature are emotional dysregulation (Selenius & Strand, 2017) or problems regarding coping skills and impulsivity (O'Shea et al., 2014; O'Donnell et al., 2015). The review by Roberton et al. (2012) showed that a broad range of maladaptive styles of emotion regulation increase the risk of AB. Coping skills refer to one's cognitive and behavioural efforts to manage stressful situations and can be used to regulate emotional responses (Lazarus & Folkman, 1984). If coping mechanisms employed by the person are not successful, continued arousal is a likely consequence (Everly & Lating Jr, 2019). Better coping skills are associated with lower levels of self-harm (McLafferty et al., 2019). Results of a study by Weinberg and Klonsky (2012) suggest that more intense self-inflicted pain facilitates a reduction of subjective feelings of stress. This suggests that self-harm can function as a maladaptive and negatively reinforced coping strategy.

Besides emotional dysregulation or maladaptive coping skills, impulsivity has been repeatedly associated with AB and self-harm. Impulsivity was found to be a predictor of physical AB in psychiatric clients (see e.g., Verstegen et al., 2020). Recent research (Cassels et al., 2022) found that high impulsivity predicted onset of NSSI in adolescents and young adults, even when accounting for other risk factors. Furthermore, a meta-analysis revealed that individuals who engaged in NSSI self-reported greater impulsivity than individuals who did not. However, little evidence of an association between lab- based measures of impulsivity and NSSI was found (Hamza et al., 2015).

Recently, Shafti et al. (2021) presented a theoretical framework that accounts for why people may engage in dual-harm: The cognitive-emotional model of dual-harm. This model is based on models and theories including the General Aggression Model, diathesis-stress models and emotional dysregulation theories (Anderson & Bushman, 2002; Ferguson et al., 2008; Long et al., 2014). This model states that biological factors combine with adverse environmental factors (e.g. childhood trauma; Fliege et al., 2009) to develop a personality style. The personality traits predispose the person to dual-harm through their effects on cognition, arousal and affect. Dual harm can have an interpersonal function or emotion regulating function. The type of harm (i.e., aggression versus self-harm) is influenced by the social context and expectancies, which is influenced by earlier effects of the behaviour. Shafti et al. (2021) concluded that the proposed model should be tested across different populations in order to assess its generalisability. Clients with mild intellectual disability or borderline intellectual functioning are often not included in studies on self-harm (and aggression) or only form a small minority of the study's sample. Furthermore, Verstegen et al. (2020) stated that more research is needed in forensic settings, which should further explore if there are shared risk factors for self-harm and AB.

The first aim of the current study is to explore the association between AB, and in particular physical AB, and NSSI in a sample of inpatients with mild intellectual disability or borderline intellectual functioning. It was hypothesized that there is a moderate positive association between NSSI and AB in line with the dual harm hypothesis (O'Donnell et al., 2015). Furthermore, similar to Verstegen et al. (2020) we explored this association for people for whom a single AB incident was reported and for people for whom multiple incidents of AB were reported. We expected a stronger association for repeated aggression.

The second aim of the study was to investigate the role of impulsivity and coping skills (elements of personality and emotional dysregulation which are components of the cognitive-emotional model of dual-harm), in clients with mild intellectual disability or borderline intellectual functioning who are residents of a treatment facility. We hypothesised that lower impulsivity and more adaptive coping skills are associated with both less NSSI and less AB. Knowledge about the association between AB and NSSI behaviour and shared risk factors in clients with mild intellectual disability or borderline intellectual functioning could contribute to risk-assessment of these behaviours in clinical practice and may help to optimise treatments for these clients.

METHOD

Participants and setting

The sample consisted of 125 clients of whom 101 males and 24 females aged between 18 and 70 years ($M = 38.1$, $SD = 11.6$) with mild intellectual disability or borderline intellectual functioning who were admitted to Trajectum. Trajectum is a treatment facility for individuals with mild intellectual disability or borderline intellectual functioning located in the Netherlands that treats clients with various behavioural and/or mental disorders, problems in multiple areas of life and often a history of substance abuse. Clients are placed in the facility under criminal law, civic law or on a voluntary basis, often for externalizing behaviour problems (i.e. aggression or a sexual offence) and/or internalizing problems (such as NSSI and suicide attempt) (Delforterie et al., 2020). Clients are often convicted for repeated minor crimes (such as shoplifting) or a single or multiple serious crimes (such as severe physical or sexual assault/abuse). Persons without offences but who are a danger for themselves or show self-destructive behaviour are also admitted. The average treatment duration at Trajectum is two years and two months (Neimeijer et al., 2021). Treatment is divided into stages with gradually reducing security levels (with clients typically moving from high security settings to open units) while offering a combination of cognitive-behavioural programmes, art-, drama-, and psychomotor therapy, occupational therapy, skills training and pharmacotherapy. Treatment programmes are embedded in a therapeutic group climate with a daily structure, activities and care that is adapted to the emotional, adaptive and intellectual abilities of clients (Neimeijer et al., 2021).

All 125 participants had a mild intellectual disability (IQ 50–69) or borderline intellectual functioning (IQ 70–84), 60 were diagnosed with substance related disorders, 48 with personality disorder, 46 with schizophrenia or other psychotic disorders, 17 with attention-deficit/hyperactivity disorder, 15 with posttraumatic stress disorder, 15 with autism spectrum disorder, 14 with paraphilic disorder and 10 with a bipolar or depressive disorder. Several clients had multiple diagnoses. From five participants, no information about diagnosis was available. In terms of time of admission, 90 participants (72%) were administered to Trajectum for the full 2 years of the study, 16 (13%) were admitted for more than 1.5 years and 19 (15%) were admitted for less than 1.5 year of the research period (at least 3 months).

Measures and materials

Descriptive variables such as age and DSM diagnosis were collected from the clients' medical record.

Aggressive behaviour and non-suicidal self- injury

Frequency and type of AB and NSSI were retrieved from the facility's digital system for incident reporting, monitoring and analysis, called SMILE. Following each incident caregivers report the incident, describe the incident and answer several questions regarding the incident, such as: the date, time, name of the involved client, name of the victim and required care. They also report the severity and type of AB (i.e., verbal, physical or towards objects) or NSSI (small cuts, burns, welts or bruises; head banging/fist slamming against a wall; skin picking, pulling out hairs, hitting without injuries or inflicting serious injuries). Furthermore, caregivers rate the risk of recurrence and the consequences of the incident. Based on risk of recurrence and consequences, the reporting system automatically determines a risk score ranging from 1 (small risk) to 4 (extreme risk), which puts different protocols of incident assessment and prevention into effect. No psychometric information is available about the SMILE.

Impulsivity and coping skills

We used data collected with the Dynamic Risk Outcome Scales (DROS; Drieschner & Hesper, 2008) to measure impulsivity and coping skills. The DROS measures dynamic risk factors for externalising behaviour and was developed to measure treatment progress in individuals with mild intellectual disability or borderline intellectual functioning and externalising (including offending) behaviour. It is a 43-item therapist-rated instrument containing 14 subscales. All items have answer categories on a 5-point Likert-type scale, with a lower score indicating higher risk. The subscale impulsivity of the DROS consists of two items, 'lack of planning' and 'sensation seeking'. A score of 5 on the item 'lack of planning' indicates that the client has no strong impulses for problematic behaviour or considers the possible consequences and the context before engaging in behaviour. A score of 1 indicates that the client often acts in response to external triggers without thinking and does not let foreseeable serious consequences stop them from following impulses immediately. A score of 5 on the item 'sensation seeking' indicates that the client feels comfortable in an environment without a lot of stimuli and avoids high-risk situations while a score of 1 indicates that the client is, almost daily, seeking or provoking situations that generate adrenaline or give a kick and in the absence of stimuli, a state of boredom or discomfort arises. The DROS impulsivity score is the average of the two items. The DROS subscale 'impulsivity' scale has a good test–retest reliability (r = 0.8), a fair inter-rater reliability (ICC = 0.49), and a low internal consistency (α = .53) (Delforterie et al., 2020).

Coping skills were measured with the DROS subscale 'coping skills' consisting of three items: 'coping with conflictual interactions', 'coping with the urge for risky behaviour' and 'coping with other stressors'. A score of 5 on 'coping with conflictual interactions' indicates that the client is capable to prevent aggravation of the situation in most conflicts without intervention of others and usually tries to reach an adequate solution to the conflict, while a score of 1 indicates that already at minor conflicts, the client has only inadequate alternatives for action which aggravates the situation, in order to prevent this, intervention by others is regularly necessary. A score of 5 on 'coping with the urge for risky behaviour' indicates that the client has sufficient coping skills, even in difficult situations, to control and/or decrease risky urges or the client has no risky urges, while a score of 1 indicates that the client has in the most situations insufficient coping skills to control risky urges, clings to the urge and without intervention from others gives in to this urge. A score of 5 on 'coping with other stressors' indicates that the client has sufficient coping skills to cope with minor stressors independently and to cope with major stressors with support, usually without problematic consequences while a score of 1 indicates that the coping skills of the client fall short even with minor stressors and despite support which has regularly (severe or persistent) problematic consequences. A score for coping skills was calculated by the average of the three items. The DROS subscale 'coping skills' has a good test–retest reliability (r = 0.8), a fair inter-rater reliability (ICC = 0.58) and a good internal consistency (α = .82) (Delforterie et al., 2020).

Procedure

This study used measurements (medical records, incident reports and DROS) which are part of the standard care during clinical treatment within Trajectum. After admission to Trajectum all clients receive an information letter and are asked to sign informed consent if they agree that the data collected during their treatment is used anonymously for scientific research. This study only included clients who had given their written informed consent to use the data for research purposes. The study was carried out in accordance with the Declaration of Helsinki and the Guidelines for Good Clinical Practice established by the International Conference on Harmonisation. The DROS was filled in by therapists and the incident reports (SMILE) were filled in by daily caregivers. Data collected between November 2018 and November 2020 were analysed.

Statistical analyses

We used SPSS (IBM SPSS Statistics, Version 27.0) for all statistical analyses. We first calculated percentages of clients engaging in different types of AB and NSSI, frequencies of different types of AB and NSSI, number of participants engaging in 1–5 or more than five incidents and how many participants engaged in multiple types of AB and NSSI. Furthermore, we ran correlation analyses for the main study variables. The assumption of normality was violated for the variables, therefore we calculated Spearman correlations.

Second, to investigate the co-occurrence of physical AB and NSSI we adopted a similar strategy as Verstegen et al. (2020), as this is a study with a similar population (psychiatric inpatients) and setting reporting on the relation between AB and self-harm using rigorous testing methods. Three Chi-squared analyses were used to examine if the frequency of NSSI differed between: (1) clients with incidents of physical AB and clients without physical AB; (2) clients with one or more incidents; and (3) between clients with 1–5 or more than five incidents of physical AB. If sample sizes were small, we reported Fisher's exact tests. In addition, we studied the co-occurrence of total frequency of AB and NSSI with the same analyses.

Third, in line with Verstegen et al. (2020) two binary regression analyses were performed with AB and physical AB as outcome variables to analyse if NSSI could differentiate between participants with and without AB and physical AB in particular. Coping skills and impulsivity were added to the regression models as independent variables to explore whether this improved the prediction of AB. We also added other subtypes of AB to the model with physical AB as outcome variable, in line with Verstegen et al. (2020). Holm correction for multiple testing was applied.

RESULTS

Aggressive behaviour

Within the 2 years observational period, a total number of 1456 aggression incidents were recorded; 735 (50.4%) of these incidents constituted verbal aggression, 473 (32.5%) of these incidents consisted of aggression towards objects and 248 (17.0%)

of these incidents consisted of physical aggression. Of all participants, 76% (95 out of 125) exhibited some type of AB (74 males and 21 females). Of these 95 participants, 12 were involved in one incident, 41 in 1–5 aggressive incidents and 54 participants in more than five incidents. The mean number of incidents was 11.6 (SD = 18.8), with a maximum of 98. Females accounted for a relatively large proportion of the incidents: 609 incidents were exhibited by females (41.8%) and 847 by males (58.2%).

Non-suicidal self-injury

Between November 2018 and November 2020, 15.2% of all participants (19 out of 125) exhibited some type of NSSI. A total number of 164 NSSI incidents were recorded. In 44 incidents (26.8%) head banging or slamming fists against a wall was observed. In 17 incidents (10.4%) skin picking, pulling out hairs or hitting was recorded. Ninety- one incidents (56.5%) resulted in small cuts, burns, welts or bruises and 12 incidents (7.3%) inflicted serious injuries. The number of incidents per participant ranged from 0 to 37. Of the 19 participants with NSSI, 3 were only once involved, 9 were involved in 1–5 incidents and 10 were involved in more than five incidents. The proportion males and females showing NSSI was nearly equal (9 females and 10 males). In the total sample, females were a minority, so relatively more females were involved in NSSI. Females accounted for 65% of the incidents.

Table 1

Descriptive statistics and Spearman's correlations for (subtypes of) NSSI, (subtypes of) aggressive behaviour, coping skills and impulsivity

Variable	M	SD	1	1a	1b	1c	1d	2	2a	3	3a	3b	4
1. NSSI	1.31	4.67	-										
1a. NSSI small	0.73	2.84	0.95**	-									
1b. NSSI HB	0.35	1.89	0.70**	0.63**	-								
1c. NSSI SP	0.14	0.50	0.75**	0.72**	0.48**	-							
1d. NSSI serious	0.10	0.55	0.55**	0.57**	0.52**	0.34**							
2. Aggression	11.65	18.76	0.29**	0.27**	0.33**	0.29**	0.16*	-					
2a. Physical aggression	1.98	4.28	0.24**	0.27**	0.32**	0.27**	0.16*	0.72**	-				
3. Impulsivity	3.61	0.75	−0.10	−0.11	−0.15*	0.04	−0.09	−0.25**	−0.15	-			
3a. Impulsivity: Lack of planning	3.68	1.13	−0.09	−0.09	−0.15*	0.06	−0.12	−0.18*	−0.11	0.79**			
3b. Impulsivity: Sensation seeking	3.54	0.94	−0.05	−0.07	−0.06	−0.01	−0.01	−0.17*	−0.08	0.65**	0.08	-	
4. Coping	2.69	0.78	−0.15*	−0.16*	−0.14	0.05	−0.19*	−0.34**	−0.23**	0.47**	0.35**	0.30**	-

Note: NSSI small **=** *small cuts, burns, welts or bruises; NSSI HB = head banging/fist slamming against a wall; NSSI SP* **=** *skin picking, pulling out hairs, hitting without injuries; NSSI serious = inflicts serious injuries.*
**p < .05; **p < .01.*

Correlation between aggressive behaviour, non-suicidal self-injury and risk factors

Spearman's correlation showed that there was a significant but weak positive correlation between total AB and NSSI, rs = 0.29, $p < .001$ and between physical AB and NSSI, rs = 0.24, $p = .003$ (see Table 1). There was a significant but weak negative correlation (rs = -0.25) between impulsivity and total aggression, participants with higher impulsivity showed more AB. Impulsivity did not correlate significantly with physical AB and NSSI, but did correlate significantly (rs = -0.15) but very weak with the NSSI subtype headbanging/fists slamming against a wall. Because of very low internal consistency ($\alpha = 0.092$) of the DROS impulsivity subscale we also explored impulsivity items separately. It was found that 'lack of planning' correlated significantly but very weak with headbanging (rs = -0.15), while 'sensation seeking' had significant but very weak correlations with total AB (rs = -0.17). There was a significant but very weak negative correlation between coping skills and NSSI (rs = -0.15) and a significant but weak negative correlation with total aggression (rs = -0.34) and physical aggression (rs = -0.23). Participants with poorer coping skills showed more NSSI and aggression.

Co-occurrence of aggressive behaviour and non-suicidal self-injury

From the total sample of 125 clients, 17 clients engaged in NSSI and AB, 78 engaged only in AB and 2 engaged only in NSSI. Fisher's exact test comparing participants with and without AB with participants with and without NSSI showed that there was no statistically significant association between total AB and NSSI ($p = .241$). There was a significant association between physical AB and NSSI, $\chi^2(1) = 3.93$, $p = .047$[4]. However, the strength of the association between physical AB and NSSI was very small, $\phi = 0.177$, $p = .047$. When the likelihood ratio was used, which is advised for small sample sizes (Field, 2013), the association was no longer significant: $\chi2(1) = 3.81$, $p = .051$. There was no significant difference in frequency of NSSI between participants with repeated aggressive incidents and those with a single aggressive incident (total aggressive incidents: 1 versus more than one: $p = .687$, physical aggressive incidents: 1 versus more than one: $p = .687$, 1–5 versus more than 5: $p = .069$).

4 This value needs to be interpreted with caution; with correction for multiple testing this value is not falling in the significant range anymore.

Co-occurrence of aggressive behaviour and non-suicidal self-injury and possible predictor variables

A binomial logistic regression analysis was performed to test the effects of NSSI, impulsivity and coping skills on the likelihood that participants showed AB. This model was statistically significant, $\chi^2(3) = 10.68$, $p = .014$ (see Table 2) and explained 12.3% (Nagelkerke R^2) of the variance in total AB. Of the three predictor variables, none contributed significantly. When we tested the model with dichotomous predictor variables (following Verstegen et al., 2020) the model was no longer significant and none of the predictor variables were significant.

Table 2

Logistic regression model for clients with and without aggression during treatment

			95% CI		
Effect	**Estimate**	**SE**	**LL**	**LU**	**p**
Constant	4.10	1.25			.001
NSSI	0.11	0.10	0.91	1.36	.308
Impulsivity[a]	−0.39	0.33	0.35	1.29	.236
Coping[b]	−0.56	0.32	0.30	1.08	.083

Note: N = 125.

Abbreviations: CI, confidence interval; LL, lower limit; UL, upper limit. [a]DROS impulsivity, with lower score indicating stronger impulsivity. [b]DROS coping skills, with a lower score indicating poorer coping skills.

Furthermore, a binomial logistic regression analysis was performed in which we explored the effects of the four subtypes of NSSI instead of the total measure of NSSI. The outcome was comparable to the model with total NSSI as predictor. Internal consistency of the subscale impulsivity in the present dataset was very low ($\alpha = 0.092$), Cronbach's alpha of the subscale coping in this dataset was acceptable (0.711). Because of the very low inconsistency of the subscale impulsivity, the items were explored separately. However, this did not lead to different results.

The logistic regression model with physical AB as outcome variable was statistically significant, $\chi^2(3) = 12.40$, $p = .006$ (see Table 3). This model explained 12.9% (Nagelkerke R2) of the variance in physical AB. Impulsivity and coping skills did not add significantly but NSSI added significantly ($p = .034$) to the model/prediction. However, when controlled for other subtypes of AB (verbal AB and AB towards objects), NSSI was no longer significant. The analysis was also performed with dichotomous predictor variables as in Verstegen et al. (2020) but this did not result in a significant change. Finally, a binomial logistic regression was performed whereby we explored the effects of the four subtypes of NSSI (instead of a total measure of NSSI) on physical AB and comparable results were found.

Table 3

Logistic regression model for clients with and without physical aggression during treamtent

			95% CI		
Effect	Estimate	SE	LL	LU	p
Constant	0.89	0.99			
NSSI	0.15	0.07	1.01	1.33	.034
Impulsivity[a]	−0.10	0.29	0.51	1.59	.721
Coping[b]	−0.45	0.30	0.36	1.14	.128

Note: N = 125.

Abbreviations: CI, confidence interval; LL, lower limit; UL, upper limit. [a],DROS impulsivity; with lower score indicating stronger impulsivity. [b],DROS coping skills, with a lower score indicating poorer coping skills

DISCUSSION

In the present study, we explored relationships between NSSI and AB and physical AB in particular and potential risk factors (i.e., impulsivity and coping skills). We studied a sample of clients with mild intellectual disability or borderline intellectual functioning in a treatment facility. The main results can be summarised as follows. There was a significant but weak positive correlation between AB and NSSI. Impulsivity correlated weakly with total aggression and coping correlated weakly with total AB, physical AB and NSSI. NSSI, impulsivity and coping skills did not predict total AB and it did not predict physical AB when controlled for other subtypes of AB. We did not find evidence for dual-harm and the overlapping emotional regulating function of these behaviours, as stated in the cognitive-emotional model of dual-harm (Shafti et al., 2021), in our sample of individuals with mild intellectual disability or borderline intellectual functioning.

Results regarding the absence of a clear association between NSSI and physical AB are in contrast with the results of Verstegen et al. (2020) and O'Donnell et al. (2015) who found a strong association. They are partially in line with van den Akker et al. (2021) who found mixed results with their review. The difference in outcomes between our study and the study of Verstegen et al. (2020) might be attributed to the larger sample size (i.e., *N* = 614) in the Verstegen study. In addition, their sample included only a small percentage (i.e., 17.3%) of people with intellectual disability. The difference in outcomes between our study and the study of O'Donnell et al. (2015) might also be attributed to differences in intellectual functioning of the sample. The authors excluded studies if the sample included more than 10% of people with intellectual disabilities. Another possible explanation is that O'Donnell et al. (2015) utilised a different definition of self-harm. They included many studies on suicide attempts and found multiple studies on NSSI without a significant association, similar to the results of the current study.

Possibly, the association between NSSI and AB differs between people with and without intellectual disabilities. Intellectual disabilities are characterised by deficits in intellectual functioning and adaptive skills (American Psychiatric Association, 2022) which makes people more susceptible to heightened stress levels compared to the general population (e.g. Forte et al., 2011). People with intellectual disabilities are also at an increased risk of exposure to adverse life events and environmental stressors and are at high risk of developing post-traumatic stress disorder (Byrne, 2018; Didden & Mevissen, 2022). Deficits in intellectual functioning can also impair the development of adaptive strategies to cope with stress (Taylor & Novaco, 2005). Consequently, individuals with intellectual disabilities could have a more diverse range of risk factors for self-harm or aggression than people without intellectual disability, potentially resulting in a smaller overlap between self- harm and aggression in this population. Another hypothesis is related to the context, which is often more controlled and restrictive in people with intellectual disabilities (Neimeijer et al., 2021), and which might result in different pathways to AB and self-harm (and NSSI in particular). The finding that coping skills and impulsivity did not predict AB could be related to the DROS, which differed from the instrument used in the studies of O'Shea et al. (2014) and Verstegen et al. (2020).

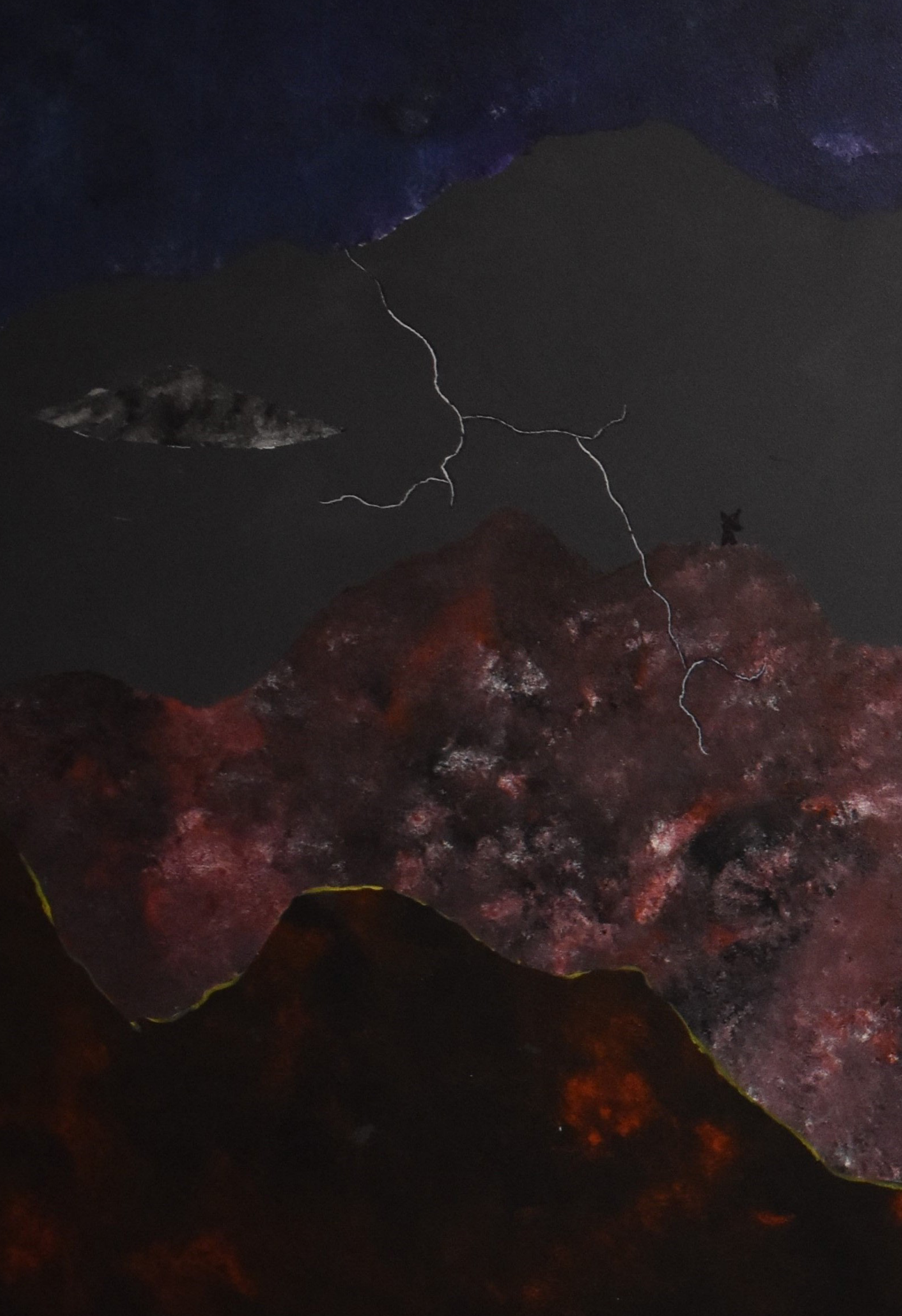
Limitations and directions for future research

The current study has several limitations. The current study was a cross-sectional observational study which limits our ability to draw causal inferences. Future research using repeated measures in a longitudinal design are needed to gain more insight in the (temporal) relation (if any) between NSSI and AB in people with mild intellectual disability or borderline intellectual functioning. With regard to the sample, the current study was conducted within one psychiatric setting for people with mild intellectual disability or borderline intellectual functioning and results cannot be generalised to other settings. Although we included a relatively large sample (i.e., 125), only 19 exhibited NSSI. Future studies using a multicentre approach could increase generalizability. Future studies with a larger sample could also explore whether gender is a moderator in the relationship between AB and NSSI and risk factors. Furthermore, the participants in the current study differed in the duration of treatment. In future studies, it would be interesting to take treatment duration into account, and to study the possible effects of treatment on the relation between NSSI and AB and its possible risk factors. With regard to the measurement instruments, it is likely that aggressive and NSSI incidents are underreported by staff, especially the milder forms of AB and NSSI, which is a problem for all studies on inpatient

AB (Arnetz et al., 2015). Moreover, not all incidents of NSSI are observed because for example clients may feel ashamed of their behaviour and hide their wounds. Also, not all participants (i.e., N = 35) were admitted the full period which probably resulted in a lower prevalence of AB and NSSI. The Covid-19 pandemic partly overlapped with the period of data collection (from February 2020 to November 2020) which might have affected the behaviour of the participants due to additional restrictions, isolation and staff shortages (Challinor et al., 2021; Theis et al., 2021). One additional limitation pertaining to measurement is our inclusion of only count data of NSSI and AB and not including the severity of incidents in the analysis. However, the categories of NSSI (i.e., from small cuts to serious injuries) were explored separately and these categories imply a degree of severity. Finally, impulsivity and coping were measured with the subscales of the DROS consisting of two and three items, respectively. Future research should use more comprehensive measurements of impulsivity and coping. This is needed to study the different components of impulsivity and coping in relation to NSSI and AB. For example, Whiteside and Lynam (2001) stated that impulsivity has a multi-faced nature. This could generate important information which can be used to adequately target impulse control problems and coping skills in treatment and improve the prevention of AB or NSSI.

Implications and conclusion

The current study contributes to the knowledge of the prevalence and relation between NSSI and AB, and physical AB in particular, in an understudied group of clients. Namely, clients with mild intellectual disability or borderline intellectual functioning in psychiatric treatment facilities. In contrast to earlier research we did not find a clear association between NSSI and aggression (dual-harm). Possibly the association between NSSI and AB differs between people with and without intellectual disabilities. With regard to clinical implications, caution is needed. Results suggest that professionals from treatment facilities for people with mild intellectual disabilities or borderline intellectual functioning could consider additional sources for measuring impulsivity and coping skills next to the DROS. Another implication might be that other risk factors could be taken into account, for example physiological stress/arousal, when preventing and treating AB and NSSI. Physiological indices have been found to be related to both AB and NSSI in previous studies. For example, De Looff et al. (2019) found a significant increase in heart rate and skin conductance in clients with mild intellectual disability or borderline intellectual functioning 20 minutes on average before AB occurred. Nock and Mendes (2008) studied the reaction of people with and without NSSI during a distressing task and found that people with NSSI had a poorer ability to tolerate the stress and higher physiological stress. Advancements in technology enable to continuously monitor physiological stress in terms of heart rate and skin-conductance, and combine these measurements with the more traditional psychosocial information. The current study provided relevant insights into the mechanisms of dual-harm in people with a mild intellectual disability or borderline intellectual functioning. The study raises several important questions that need to be taken into account when studying AB and NSSI in this vulnerable population, and replication and further research is needed.



Chapter 3

A systematic review of studies on the association between physiological parameters and self-harm

This chapter was published as:

van Swieten, M., Nijman, I., de Looff, P., VanDerNagel, J., & Didden, R. (2025). A systematic review of studies on the association between physiological parameters and self-harm. *Research in Developmental Disabilities, 162,* 105010. https://doi.org/10.1016/j.ridd.2025.105010

ABSTRACT

Background: Self-harm is common in people with intellectual disabilities and is associated with multiple adverse consequences for the client engaging in self-harm, other clients and caregivers. Self-harm is related to emotional dysregulation according to both observational and self-report data. Measures of the autonomic nervous system might provide additional insight in this relationship.
Methods: The current systematic review systematically summarised a broad spectrum of studies on the association between self-harm and physiological parameters. The search identified 2400 articles, 46 were included.
Results: In most studies, which compared electrodermal activity and heart rate in people with and without self-harm, no clear indications for a relation between physiology and self-harm was found. Studies on heart rate variability showed indications for lower heart rate variability during recovery, which could imply emotion dysregulation, findings which were supported by results from imagery studies (heart rate and skin conductance). No consistent findings were found when self-harm was studied before, during or after actual occurrences of self-harm, although this was examined by very few studies.
Conclusions: Although wearable technology has improved, the majority of studies to date are lab-studies. Future research should focus on measuring physiology in daily life before, during and after self-harm, in people with intellectual disabilities, study different types and functions of self-harm separately, and test multimodal prediction models. This knowledge could improve the understanding, prevention and assessment of this debilitating behaviour.

INTRODUCTION

Self-harm has a high prevalence and may result in multiple adverse consequences. Self-harm is an umbrella term that encompasses both self-injurious behaviour with intent to end one's life and self-injurious behaviour without intent to die (Shafti et al., 2021). Self-harm is estimated to affect 14.6 million individuals each year (Global Burden of Disease Collaborative Network, GBD, 2019). A recent meta-analysis (Lucena et al., 2022) reported a lifetime prevalence of 19% (95% CI = 17–21) in the general adolescent population. People with psychiatric disorders (e.g. affective disorders and personality disorders) and people with intellectual disabilities are found to be at higher risk of self-harm (Bøe et al., 2022; Calver et al., 2021; Flygare Wallén et al., 2023). Severe and prolonged self-harm can result in irreversible mutilation, blindness, or brain injury (Hoch et al., 2013), and it is associated with heightened suicide risk (Duarte et al., 2020). A study found that children and adolescents who engaged in self-harm were 17 times more likely to die by suicide than those not engaging in self-harm (Morgan et al., 2017). Besides medical complications, prolonged and repetitive self-harm in adolescence is associated with mental health problems in young adulthood (Daukantaitė et al., 2021).

Emotion dysregulation is one of the most often reported risk factors for self-harm (Li et al., 2023; Wolff et al., 2019). Greater emotion dysregulation is associated with higher risk of self-harm across settings, age groups, sex and sample types (i.e., clinical versus community) (Wolff et al., 2019). From a behavioural function perspective, which is often used in research and clinical practice in the field of intellectual and developmental disabilities (IDD), self-harm is proposed to be maintained via four possible reinforcement processes. These processes differ according to whether the reinforcement is positive or negative, and whether the consequent events are intrapersonal or interpersonal (Nock, 2010). Modulation of arousal is one of the most frequently reported intrapersonal functions of self-harm (Nock, 2010). Self-harm can manifest as a consequence of either under-arousal or over-arousal, as stated by the theoretical framework of homeostasis (Furniss & Biswas, 2012; Hoch et al., 2013). From this perspective, self-harm may be seen as a maladaptive coping strategy that reduces stress and re-establishes homeostasis in the short term but serves to erode health in the long term (Everly & Lating Jr, 2019; Nock, 2010).

Extensive research using behavioural observations (especially in the IDD field) and self-reports has been conducted on the relation between self-harm, emotion regulation and arousal (e.g. Armey et al., 2011; Gordon et al., 2010; Gratz et al., 2011; Knowles & Townsend, 2012; Turner et al., 2015; Vansteelandt et al., 2017). However, self-harm is often concealed (Shiner, 2008) and recall biases are common in self-report assessments of self-harm thoughts and behaviours (Klimes-Dougan et al., 2022). Measures of the autonomic nervous system (ANS) might provide additional insight in arousal and physiology underlying emotion dysregulation. The ANS can be subdivided into two branches, the sympathetic and the parasympathetic. The sympathetic branch prepares the body for action while the parasympathetic branch of the ANS regulates restorative functions and relaxation of the body (Everly & Lating Jr, 2019). Both heart rate (variability) (HR[V]) and electrodermal activity (EDA) have been used as indices of autonomic nervous system functioning research over the past decades (e.g. Boucsein, 2012; Jarczok et al., 2013; Kamath et al., 2016). Also, HRV has been repeatedly reported as a valid psychophysiological marker for emotion regulation capacity (Beauchaine & Thayer, 2015). Low resting state HRV and excessive HRV reactivity have been associated with dysfunctional emotion regulation processes (Beauchaine & Thayer, 2015). Furthermore, increased arousal can be reflected by an increase in heart rate HR and EDA. Advancements in technology, such as the development of wearable technology, have enabled continuous and non-invasive measurement of physiological parameters of the ANS (Johnson & Picard, 2020). Consequently, over the last two decades, research has increasingly focused on examining these physiological parameters in relation to self-harm.

Recently, two meta-analyses have been published on the association between self-harm and physiological parameters. Goreis et al. (2023) conducted a meta-analysis on physiological stress reactivity and self-harm, in which they examined case-control lab studies specifically reporting on stress reactivity (Goreis et al., 2023). The main findings indicated that parasympathetic ANS activity, such as HRV (Root Mean Square of Successive Differences; RMSSD and Respiratory Sinus Arrhythmia; RSA), was lower at rest and recovery in the self-harm group compared to controls. Bollato et al. (2023) conducted a review and meta-analysis on the autonomic dysregulation and self-harm thoughts and behaviours in children and young people, including case-control and cross-sectional studies. Results showed lower HRV in individuals with self-harm compared to individuals without self-harm, while overall results for EDA were not significant. They used a combined measure of baseline, activity and reactivity, and subgroup analyses found no differences in results between these different measures.

The above two meta-analyses found evidence for significantly altered stress reactivity in people who engage in self-harm and reduced parasympathetic regulation in children and young people who self-harm, respectively. However, both excluded case studies and included mainly lab studies. Much less is known on the relation between self-harm and physiological parameters in the natural environment and physiology related to actual or imaged self-harm incidents. Moreover, persons with intellectual disabilities were not represented in these studies, while this is a population at high risk of self-harm. Lastly, it should be noted that Goreis et al. included studies in which not all participants in the self-harm group had a history of self-harm.

The current systematic review aims to update and extend the aforementioned reviews on the association between self-harm and physiological parameters by systematically summarizing a broad spectrum of studies. Because of the emergence of studies on physiology in relation to self-harm and the recent developments in measuring physiology with wearables in daily life, the current systematic review had no restrictions regarding study design. We aimed to include non-lab studies and studies that examined physiology related to actual or imaged self-harm incidents. In addition, we aimed to include persons with intellectual or developmental disabilities because of their high risk of self-harm. Moreover, self-harm in individuals with IDD is likely distinct from self-harm occurring in typically developing individuals presenting with psychopathology (Hagopian & Leoni, 2017) and therefore the results of previous review studies on the association between physiology and self-harm cannot be generalized. Furthermore, studies on persons with IDD often use different methodologies (i.e., direct observation) compared to studies with typically developing individuals, using mostly self-report. Including both types of studies provides a comprehensive overview. Lastly, the present review only includes studies in which all participants in the self-harm group have a history of self-harm.

We wanted to identify research gaps and provide directions for future research on the relation between self-harm and physiology, for example with regard to physiological parameters, measurement type (baseline, activity, reactivity, recovery), sample characteristics and setting. Knowledge on the association between self-harm and physiological parameters can help understand risk factors, causes and functions of self-harm, which increases the possibilities for prevention and treatment. The clinical relevance is further amplified because new technology enables continuous and non-invasive measurement of physiology in individuals who self-harm. If more knowledge is available on the association between self-harm and physiological parameters, the technology might be used for personalized care and improvement of assessment, prevention and interventions of this debilitating behaviour.

METHOD

Literature search

The review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Page et al., 2021) guidelines. The protocol was pre-registered in PROSPERO (CRD42023401662).

The systematic review was conducted using four databases (i.e., Medline, PsycINFO, Web of Science, and Embase). As no systematic review study existed on the relation between self-harm and physiological parameters at the time of the search (February, 2023), there were no restrictions placed on publication date for inclusion of papers. A combination of search terms derived from the concepts of physiology and self-harm was used to ensure a comprehensive search of the relevant literature. The search terms associated with each of the concepts are provided in Table 1; the comprehensive search terms are available in Appendix 3A. A language filter was applied to include only articles written in English.

Table 1

Summary of Search Terms Used to Identify Studies on the Physiological Parameters of Self-harm

	Search terms
Physiology	*Physiology in general* Psychophysiolog*, physiolog*, autonomic nervous system, sympathetic, parasympathetic, arousal *HR(V)* Heart rate*, heartrate, cardiovascular reacti*, cardiac reacti*, cardiovascular respons*, cardiac respons*, electrocardiograph*, HRV, pre-ejection period*, respiratory sinus arrhythmia *SC* GSR, electrophysiolog*, galvanic, electrodermal, psychogalvanic, (skin, dermal-response*), electr*, conduct*, resist*, potential*
Self-harm	self-harm*, selfharm*, auto-mutilati*, automutilati*, self-mutilati*, selfmutilati*, auto-aggressi*, autoagressi*, self-injur*, selfinjur*, para-suicid*, parasuicid*, NSSI, self-inflict*, selfinflict*, self-damag*, selfdamag*, self-destruct*, selfdestruct, (harm*, hurt* or injur* - oneself)

Note. *HR(V) = heart rate (variability); SC = skin conductance; GSR = galvanic skin response; NSSI = non-suicidal self-injury.*

Selection process

The study selection process is outlined in Figure 1. The initial literature search yielded a total of 2724 studies. Following the removal of duplicate articles, the titles and abstracts of 2400 articles were independently screened and labelled as included or excluded by two researchers (MvS and IN). Articles were selected using the following inclusion and exclusion criteria: Studies were eligible for inclusion if they reported on the association between a measure of HR(V) or EDA and a measure of self-harm in adolescents and/or adults (mean age 12 years or older). Both experimental and observational studies reporting on measures of physiological rest, activity, reactivity, and recovery were included. Exclusion criteria were: only investigating other (endocrine) physiological measures (e.g., cortisol response), non-human subjects, no original or no peer-reviewed results (e.g., dissertations, reviews, study protocols, editorials), studies solely investigating suicide and studies with a sample with a mean age below 12 years old. To calculate interrater agreement the total number of articles with agreement were divided by the total number of screened articles (2400) and multiplied by 100. A total of 66 articles were labelled 'Maybe' because discussion was needed. The articles which were labelled 'Maybe' were treated (conservatively) as no agreement. Percent agreement was excellent (i.e., 97%). The full texts of the 85 selected studies were then divided between the two researchers based on alphabetical order and assessed to identify studies reporting on the relation between self-harm and physiology. As a result, 39 studies were deemed eligible for inclusion in this review.

Following the literature search in February 2023, the meta-analysis by Goreis et al. (2023) and the review and meta-analysis of Bellato et al. (2023) were published. To ensure the inclusion of all relevant studies, we compared the studies included in Bellato et al. (2023) and Goreis et al. (2023) with the studies we included at that moment and the inclusion and exclusion criteria, as a result of which one study was added (Bohus et al., 2000). Finally, an updated search was conducted in April 2024, which resulted in the inclusion of 6 additional studies in this review. The current review therefore included a total of 46 studies.

Data extraction

The following data were extracted from each of the included studies: first author, country, year of publication, sample description, recruitment, comparison group, design, age of sample, percentage of (fe)males, sample size, percentage of individuals with intellectual disability, severity level of intellectual disability, outcome measure, measurement type (rest, activity, reactivity or recovery), method of assessment of physiological parameter, type of experiment, method of assessment of self-harm, type and time of self-harm and findings regarding the relation between physiological parameters and self-harm.

Physiological parameters

A large variety of physiological parameters were extracted. For an overview of physiological parameters reported, their abbreviations and the branch of the ANS they reflect, see Table 2. Besides HR, HRV and EDA, ANS activation can be measured through the effect on other target organs (e.g. eyes; pupil dilation), however, these are to varying degrees more difficult to measure in daily life and therefore not included in this systematic review.

Physiological parameters can be obtained at different time points. For the current review, physiology measured at rest means that the body is in a state of minimal activity and that the participant is not engaging in self-harm or any stressful or demanding tasks. Measuring physiological parameters during activity involves assessing physiological parameters during self-harm (imagination) or tasks, for example when the participant experiences pain or another stressor. Reactivity refers to how the body responds to self-harm (imagination) or other stressful stimuli by comparing activity measures to baseline measures. Recovery refers to changes in physiology after being exposed to a stressor or self-harm.

Table 2

Overview of Physiological Parameters and their Abbreviations

Category	Sympathetic	Parasympathetic	Combination
HR			Heart rate (HR)
HRV	Pre-ejection Period (PEP)	Respiratory sinus arrhythmia (RSA) Root-mean-square of successive difference between normal heartbeats (RMSSD)	Heart rate variability (HRV)-triangular index
EDA	Electrodermal activity (EDA) Skin Conductance Level (SCL) Skin Conductance Responses (SCRs)		

Quality assessment and data synthesis

To assess the quality of case studies and case series, the 10-item Joanna Briggs Institute (JBI) critical appraisal tool for assessing case series was used (Munn et al., 2020), as suggested by Quigley et al. (2019). The JBI does not provide cut-off values for determining an overall quality score because not all questions have equal weight (Munn et al., 2020). An overall quality score of low, moderate or high was assigned to each case study by assessing the answers to each of the items. Items that exhibited minimal variability across studies were accorded less weight in the evaluation process.

To assess the risk of bias and quality of the case-control and longitudinal studies the Newcastle-Ottawa Scale (NOS; Wells et al., 2011) case-control and cohort versions of the scale were used. Cross-sectional studies were assessed with the NOS adapted for cross-sectional studies (Herzog et al., 2013). Items were customized and operationalized to the research question of this review, and several adjustments were made based on recommendations by Stang (2010). The adjusted version of the NOS can be found in Appendix 3B. A total of nine stars could be scored per study. In line with the study of Losilla et al. (2018) a score of 0-3 was considered low quality, 4-6 moderate quality and 7-9 was considered high quality. Due to different study designs and different quality assessment tools, quality scores cannot be compared between the different types of studies.

Two researchers (MvS and IN) scored all studies independently. To calibrate the scoring process, the scoring of the first five case-control studies was evaluated before proceeding to the scoring of the remaining case-control studies independently. These first five studies were not used in the calculation of the interrater agreement. Interrater agreement on the quality assessment of case-control studies was calculated for both the items and for the quality scores.

Studies were categorized into within-group and between-group comparison designs, and data were subdivided based on different types of physiological parameters (HR, HRV and EDA). To weigh the relative importance of studies, between-subjects studies were organized by the level of certainty of evidence in which study quality (based on risk of bias and quality assessment) and sample size were taken into account. Sample size was categorized into very small (0-5), small (6-29), moderate (30-49) and large ($\geq$ 50 per group). Allocation to a category was based on the smallest group in case of unequal group sizes. We used the following criteria:

Criterion 1) High quality, moderate or large sample size ($n > 30$)

Criterion 2.1) High quality, small sample size ($n < 30$)

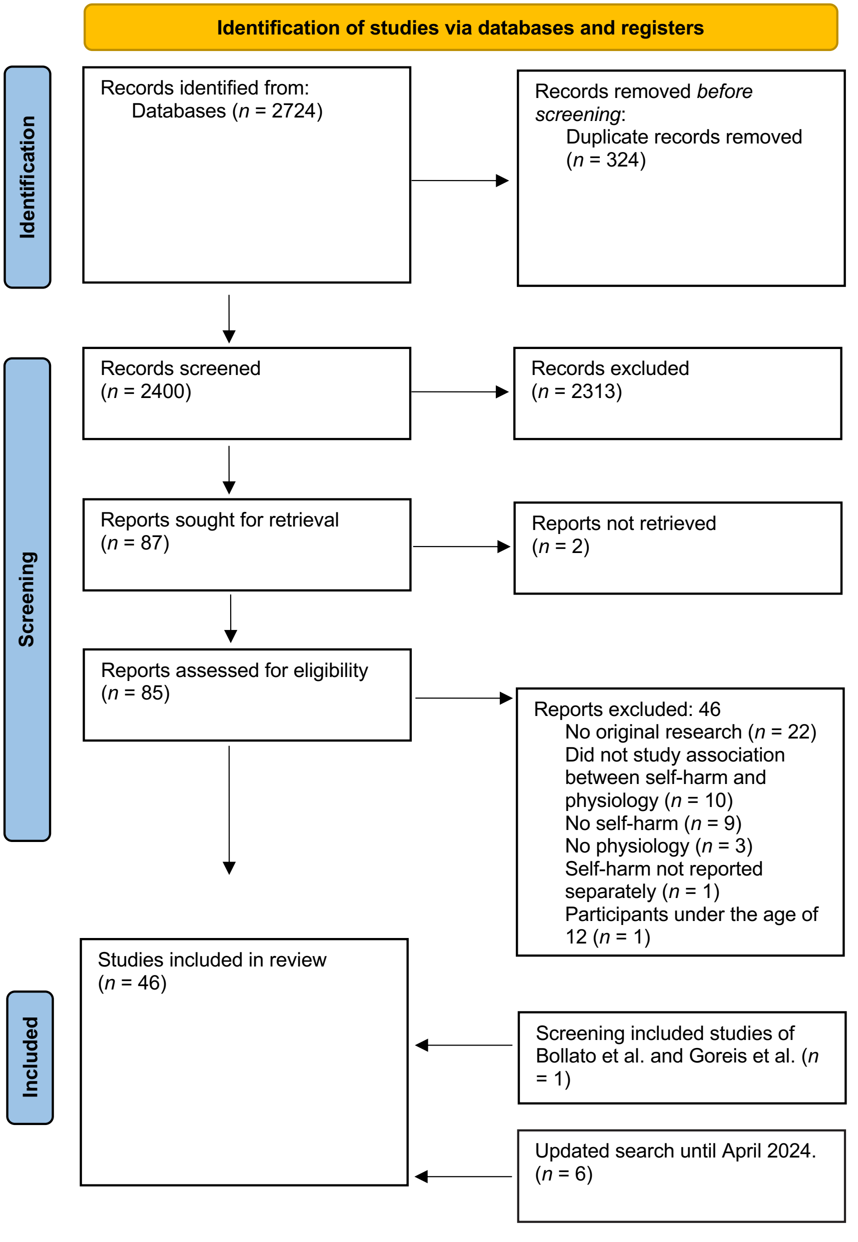
Criterion 2.2) Moderate/low quality, moderate or large sample size ($n > 30$)

Criterion 3) Moderate/low quality, small sample size ($n < 30$)

Studies that meet Criterion 1 are considered to provide the most reliable results. Criterion 2.1 and 2.2 are considered promising results, while studies that meet Criterion 3 need to be interpreted with caution.

Figure 1

Overview of the Number of Studies in the Various Stages of the Selection Process



RESULTS

General study characteristics

While 37 studies investigated the association between physiology and self-harm in a between-subject design [5](Aldrich et al., 2018; Bohus et al., 2000; Brain et al., 1998; Brain et al., 2002; Crowell et al., 2012; Crowell et al., 2005; Crowell et al., 2017; Fitzpatrick et al., 2020; Fox et al., 2018; Giner-Bartolome et al., 2017; Godfrey et al., 2022; Gratz et al., 2019; James et al., 2024; Kaess et al., 2012; Kaufman et al., 2020; Kaufman et al., 2018; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017; Lin et al., 2019; Morley, 2000; Müerner-Lavanchy et al., 2024; Naoum et al., 2019; Naoum et al., 2016; Nelson et al., 2023; Nock & Mendes, 2008; Otto et al., 2023; Reitz et al., 2015; Reitz et al., 2012; Sigrist et al., 2024; Tatnell et al., 2018; Tuna & Gencoz, 2021; van der Venne et al., 2023; Wells et al., 1999; Wielgus et al., 2016; Willis et al., 2018; Willis et al., 2017), 13 studied this association within-subject (Barrera et al., 2007; Bohus et al., 2000; Brain et al., 1998; Brain et al., 2002; Calamari et al., 1990; Ferguson et al., 2019; Freeman et al., 1999; Hall et al., 2013; Hoch et al., 2013; Koenig et al., 2018; Lydon et al., 2013; Welch et al., 2008; Wells et al., 1999)[6]. For an extensive overview of the included studies, including participant characteristics, self-harm characteristics methods and results, see table 3 (within-subject) and 4 (between-subject) in appendix 3C and 3D. Various study designs were used: case control study (*n* = 33) (Bohus et al., 2000; Brain et al., 1998; Brain et al., 2002; Crowell et al., 2012; Crowell et al., 2005; Crowell et al., 2017; Fox et al., 2018; Giner-Bartolome et al., 2017; Godfrey et al., 2022; Gratz et al., 2019; James et al., 2024; Kaess et al., 2012; Kaufman et al., 2020; Kaufman et al., 2018; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017; Lin et al., 2019; Morley, 2000; Müerner-Lavanchy et al., 2024; Naoum et al., 2019; Naoum et al., 2016; Nock & Mendes, 2008; Otto et al., 2023; Reitz et al., 2015; Reitz et al., 2012; Sigrist et al., 2024; Tatnell et al., 2018; Tuna & Gencoz, 2021; van der Venne et al., 2023; Wells et al., 1999; Willis et al., 2018; Willis et al., 2017), case study (*n* = 7) (Barrera et al., 2007; Calamari et al., 1990; Ferguson et al., 2019; Freeman et al., 1999; Hall et al., 2013; Hoch et al., 2013; Lydon et al., 2013), cohort or longitudinal study (*n* = 4) (Aldrich et al., 2018; Koenig et al., 2018; Nelson et al., 2023; Wielgus et al., 2016), and cross-sectional study (*n* = 2) (Fitzpatrick et al., 2020; Welch et al., 2008). The total number of participants ranged from 1 to 209, with six studies having a very small (0-5) number of participants (Barrera et al., 2007; Calamari et al., 1990; Freeman et al., 1999; Hall et al., 2013; Hoch et al., 2013; Lydon et al., 2013), 21

5 Due to the large number of references and in order to improve readability, the references in this chapter's results are displayed in grey.

6 Some studies used both a between-subject design and a within-subject design.

studies having a small (6-29) number of participants (Bohus et al., 2000; Brain et al., 1998; Brain et al., 2002; Crowell et al., 2012; Crowell et al., 2005; Crowell et al., 2017; Ferguson et al., 2019; Fitzpatrick et al., 2020; Giner-Bartolome et al., 2017; Gratz et al., 2019; James et al., 2024; Kaess et al., 2012; Koenig et al., 2018; Naoum et al., 2019; Naoum et al., 2016; Otto et al., 2023; Reitz et al., 2015; Reitz et al., 2012; Tatnell et al., 2018; Wells et al., 1999), 15 studies having a moderate (30-49) number of participants (Fox et al., 2018; Godfrey et al., 2022; Kaufman et al., 2020; Kaufman et al., 2018; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017; Morley, 2000; Müerner-Lavanchy et al., 2024; Nock & Mendes, 2008; Sigrist et al., 2024; Tuna & Gencoz, 2021; van der Venne et al., 2023; Welch et al., 2008; Willis et al., 2018), and four studies having a large (>50) number of participants (Aldrich et al., 2018; Lin et al., 2019; Nelson et al., 2023; Willis et al., 2017). With regard to characteristics of the participants, seven studies - all case studies - investigated the relationship between self-harm and HR in individuals with intellectual disabilities (Barrera et al., 2007; Calamari et al., 1990; Ferguson et al., 2019; Freeman et al., 1999; Hall et al., 2013; Hoch et al., 2013; Lydon et al., 2013). Eight studies reported on self-harm in persons with borderline personality disorder (BPD) (Bohus et al., 2000; Gratz et al., 2019; Naoum et al., 2019; Naoum et al., 2016; Reitz et al., 2015; Reitz et al., 2012; Willis et al., 2018; Willis et al., 2017) and one study reported on self-harm in persons with eating disorder (Giner-Bartolome et al., 2017). In addition, two studies compared persons with self-harm to persons with depression but without self-harm and healthy controls (HC) (Crowell et al., 2012; Crowell et al., 2017). Most studies (n = 28) did not focus on self-harm in participants with a specific psychiatric disorder.

Study quality

Of the studies assessed with the NOS case-control (n = 33), two had low quality (Brain et al., 2002; Reitz et al., 2015), 17 had moderate quality (Bohus et al., 2000; Brain et al., 1998; Crowell et al., 2012; Crowell et al., 2005; Crowell et al., 2017; Giner-Bartolome et al., 2017; Godfrey et al., 2022; Kaess et al., 2012; Morley, 2000; Müerner-Lavanchy et al., 2024; Naoum et al., 2019; Naoum et al., 2016; Reitz et al., 2012; Tuna & Gencoz, 2021; van der Venne et al., 2023; Wells et al., 1999; Willis et al., 2018; Willis et al., 2017) and 13 had high quality (Fox et al., 2018; Gratz et al., 2019; James et al., 2024; Kaufman et al., 2020; Kaufman et al., 2018; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017; Lin et al., 2019; Nock & Mendes, 2008; Otto et al., 2023; Sigrist et al., 2024; Tatnell et al., 2018). Percentage of interrater agreement on item-level (88.8%) and on the quality scores (low, moderate, high) was very good (i.e., 89,7%). Of the studies assessed with the NOS-cohort (n = 4), three had a moderate quality score (Koenig et al., 2018; Nelson et al., 2023; Wielgus et al., 2016) and one study had high quality (Aldrich et al., 2018). Of the studies assessed with the NOS cross sectional (*n* = 2) one study had a low quality (Welch et al., 2008) and one had a moderate quality (Fitzpatrick et al., 2020). Of the studies assessed with JBI critical appraisal tool for assessing case series (*n* = 7) two had low quality (Calamari et al., 1990; Ferguson et al., 2019), three had moderate quality (Barrera et al., 2007; Freeman et al., 1999; Lydon et al., 2013) and two had high quality (Hall et al., 2013; Hoch et al., 2013).

Between-subject studies

Five main types of experiments can be distinguished in the included studies. First, exposure to stress: social evaluative stress (for example the Trier Social Stress test, TSST), cognitive stress (for example arithmetic challenges) or a combination of these two (for example the Montreal Imaging Stress Task, MIST) (Aldrich et al., 2018; Fox et al., 2018; Gratz et al., 2019; Kaess et al., 2012; Kaufman et al., 2018; Koenig et al., 2023; Lin et al., 2019; Morley, 2000; Naoum et al., 2019; Nock & Mendes, 2008; Otto et al., 2023; Reitz et al., 2015; Reitz et al., 2012; Tatnell et al., 2018; Tuna & Gencoz, 2021; Wielgus et al., 2016; Willis et al., 2018; Willis et al., 2017). Second, sad emotion induction (watching a clip from the movie The Champ) (Crowell et al., 2012; Crowell et al., 2005). Third, physical stress/pain simulation by the Cold Pressure Task (CPT), using a hot plate or a sham and/or incision with a blade (Bohus et al., 2000; Koenig, Rinnewitz, Warth, et al., 2017; Naoum et al., 2016; Reitz et al., 2012; Tuna & Gencoz, 2021; van der Venne et al., 2023; Willis et al., 2018; Willis et al., 2017). Fourth, imagery studies, which used personalized guided imagery scripts depicting an actual episode of self-harm (Brain et al., 1998; Brain et al., 2002; Wells et al., 1999). And fifth, a discussion task between the mother and daughter to incite disagreement and emotional arousal (Crowell et al., 2017; Godfrey et al., 2022; James et al., 2024; Kaufman et al., 2020). One study used another type of experiment and measured physiology during the first session of a serious game consisting of a biofeedback intervention to teach strategies to enhance self-control and regulate negative emotion (Giner-Bartolome et al., 2017). Besides, 5 of the 37 studies did not use an experiment. Of these studies, four investigated physiology only in rest (Fitzpatrick et al., 2020; Koenig, Rinnewitz, Parzer, et al., 2017; Müerner-Lavanchy et al., 2024; Nelson et al., 2023) and one study examined diurnal variation of HRV (Sigrist et al., 2024). For an overview of the results of the between-subject studies, see Figure 2.

Heart rate

In total, 18 studies reported on the relationship between self-harm and parameters of HR (Bohus et al., 2000; Brain et al., 1998; Brain et al., 2002; Kaess et al., 2012; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017; Morley, 2000; Naoum et al., 2019; Naoum et al., 2016; Nelson et al., 2023; Reitz et al., 2015; Reitz et al., 2012; Sigrist et al., 2024; van der Venne et al., 2023; Wells et al., 1999; Willis et al., 2018; Willis et al., 2017). One study could not be allocated to rest, activity, reactivity or recovery and therefore will be discussed below (Sigrist et al., 2024). Overall, 21% of the results showed a significant relation between self-harm and HR. Sigrist et al. (2024) identified significant differences in diurnal variation of HR between a self-harm and a HC group. They found that the self-harm group showed significantly higher rhythm-adjusted mean HR (a measure of diurnal rhythmicity based on a cosine function) and lower HR amplitude during the day compared to HCs.

Rest. 11 studies reported on HR in rest (Bohus et al., 2000; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017; Morley, 2000; Naoum et al., 2019; Naoum et al., 2016; Nelson et al., 2023; Reitz et al., 2012; van der Venne et al., 2023; Willis et al., 2018). Of these, all three studies that met Criterion 1, (Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017) did not find a significant association between resting HR and self-harm. Four studies met Criterion 2.2 (Morley, 2000; Nelson et al., 2023; van der Venne et al., 2023; Willis et al., 2018), three of which did not find a significant association between resting HR and self-harm (Morley, 2000; Nelson et al., 2023; van der Venne et al., 2023), whereas one study reported that the association depended on the timeframe in which self-harm occurred (Willis et al., 2018). This study (Willis et al., 2018) noted a higher resting HR in current (BPD) patients (with self-harm at least once during 6 months prior to study participation) compared to HCs. However, this study found no significant difference in resting HR when comparing remitted BPD patients (with two or less acts of self-harm in the past year, but all self-harmed before) with HC. Lastly, four studies met Criterion 3. Among these, three studies did not find a significant association between resting HR and self-harm, while Bohus et al. (2000) found that the association depended on the level of distress experienced by the self-harm group. Specifically, this study reported significantly lower HR in rest in HCs compared to a self-harm group in self-reported distress, but no significant difference in HR in rest between HCs and a self-harm group in self-reported calmness.

Activity. Five studies reported on HR during activity (Bohus et al., 2000; Koenig et al., 2023; Morley, 2000; Naoum et al., 2019; van der Venne et al., 2023). Among these, one study met Criterion 1 (Koenig et al., 2023), and results showed that the self-harm group had a significantly lower HR during the interview and mental arithmetic tasks of the TSST compared to HCs (Koenig et al., 2023). Two studies were Criterion 2.2 studies and did not find a significant association between self-harm and HR during pain stimulation (van der Venne et al., 2023), nor during a serial subtraction math task, a mirror image star tracing task or a speech task (Morley, 2000). There were two Criterion 3 studies, which did not find a significant association between self-harm and HR during the CPT (Bohus et al., 2000) and the concentration performance test (Naoum et al., 2019).

Reactivity. Of the 12 studies which reported on HR reactivity, two met Criterion 1 (Koenig et al., 2023; Koenig, Rinnewitz, Warth, et al., 2017). One of these studies found a significant negative relationship between self-harm and HR reactivity, such that during stress induction using the TSST, HR increased more in the HC group than in the self-harm group (Koenig et al., 2023). The other Criterion 1 study did not find a significant association between self-harm and HR reactivity during the CPT (Koenig, Rinnewitz, Warth, et al., 2017). Three studies were Criterion 2.2 studies and did not find a significant association between self-harm and HR reactivity during the MIST (Willis et al., 2018), during pain stimulation (van der Venne et al., 2023), nor during a serial subtraction math task, a mirror image star tracing task or a speech task (Morley, 2000). The remaining seven studies were Criterion 3 studies (Brain et al., 1998; Brain et al., 2002; Kaess et al., 2012; Naoum et al., 2016; Reitz et al., 2015; Reitz et al., 2012; Wells et al., 1999), of which five did not find a significant association between HR reactivity and self-harm during personalized guided imagery (Brain et al., 1998; Brain et al., 2002), the TSST (Kaess et al., 2012), or the MIST (Reitz et al., 2015; Reitz et al., 2012). Two Criterion 3 studies did find a significant association between HR reactivity and self-harm (Naoum et al., 2016; Wells et al., 1999). Naoum et al. (2016) found a significant negative association between self-harm and HR reactivity, with a greater increase in HR in the HC group than in the self-harm group, while Wells et al. (1999) found a significant positive association, with greater HR reactivity across stages of a imagery script in the self-cutting and severe nail-biting group than in the mild nail-biting group.

Recovery. Of the 10 studies reporting on HR recovery after stress or pain induction, two were Criterion 1 studies. Both did not find significant associations between self-harm and HR recovery following the TSST (Koenig et al., 2023) and CPT (Koenig, Rinnewitz, Warth, et al., 2017). There were three Criterion 2.2 studies, and all three did not find significant associations between self-harm and HR recovery following the MIST (Willis et al., 2018; Willis et al., 2017) and following pain stimulation (van der Venne et al., 2023). There were five Criterion 3 studies that examined the relationship between self-harm and HR recovery, of which four did not find a significant association (Bohus et al., 2000; Naoum et al., 2019; Reitz et al., 2015; Reitz et al., 2012). However, Naoum et al. (2016) found that when receiving a pain stimulus (non-invasive blade) with artificial blood added after stress induction using the concentration performance test, HR decreased more in the HC group than in the self-harm group.

Heart rate variability

In total, 20 studies reported on HRV (Crowell et al., 2005; Crowell et al., 2017; Fitzpatrick et al., 2020; Fox et al., 2018; Giner-Bartolome et al., 2017; Godfrey et al., 2022; Gratz et al., 2019; James et al., 2024; Kaufman et al., 2020; Kaufman et al., 2018; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017; Lin et al., 2019; Müerner-Lavanchy et al., 2024; Otto et al., 2023; Reitz et al., 2015; Sigrist et al., 2024; van der Venne et al., 2023; Wielgus et al., 2016). One study could not be allocated to rest, activity, reactivity or recovery and therefore will be discussed below (Sigrist et al., 2024). Overall, 40% of the results showed a significant relation between self-harm and HRV. Sigrist et al. (2024) found significant differences in diurnal variation of HRV between a self-harm group and a HC group. Specifically, this study found that the self-harm group showed significantly lower rhythm-adjusted mean HRV and higher HRV amplitude during the day compared to HCs.

Rest. Of the 15 studies reporting on HRV in rest (Crowell et al., 2005; Fitzpatrick et al., 2020; Fox et al., 2018; Godfrey et al., 2022; Gratz et al., 2019; Kaufman et al., 2020; Kaufman et al., 2018; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017; Lin et al., 2019; Müerner-Lavanchy et al., 2024; Otto et al., 2023; van der Venne et al., 2023; Wielgus et al., 2016), seven were Criterion 1 studies (Fox et al., 2018; Kaufman et al., 2020; Kaufman et al., 2018; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017; Lin et al., 2019). Of these studies, five did not find a significant relationship between HRV in rest and self-harm (Fox et al., 2018; Kaufman et al., 2018; Koenig et al., 2023; Koenig, Rinnewitz, Parzer, et al., 2017; Koenig, Rinnewitz, Warth, et al., 2017). One of the Criterion 1 studies found that self-harm was related to lower baseline RSA (Kaufman et al., 2020). The remaining Criterion 1 study found that the relationship between self-harm and HRV depended on the timeframe in which self-harm occurred. They found that baseline HRV (RSA) was not related to lifetime self-harm. However, there was a significant positive correlation between baseline HRV (RSA) and antenatal self-harm in pregnant women (Lin et al., 2019). There was one Criterion 2.1 study (Gratz et al., 2019) and four Criterion 2.2 studies (Godfrey et al., 2022; Müerner-Lavanchy et al., 2024; van der Venne et al., 2023; Wielgus et al., 2016). Three of them did not find significant associations between HRV in rest and self-harm (Müerner-Lavanchy et al., 2024; van der Venne et al., 2023; Wielgus et al., 2016). However, Godfrey et al. (2022) found a significant positive correlation between baseline PEP and lifetime self-harm. Of the studies examining HRV in rest, there were three Criterion 3 studies (Crowell et al., 2005; Fitzpatrick et al., 2020; Quigley et al., 2019), of which Otto et al. (2023) did not find a significant difference in HRV in rest between a self-harm group and HCs. However, Crowell et al. (2005) found that the association between self-harm and HRV in rest depended on the measure of HRV used. They found baseline RSA to be negatively associated with self-harm, but no significant association between self-harm and PEP in rest was found (Crowell et al., 2005). In contrast, Fitzpatrick et al. (2020) found that higher baseline RSA predicted greater frequency of self-harm.

Activity. Of the six studies reporting on HRV during activity, one was a Criterion 1 study (Koenig et al., 2023). This study reported no significant difference in RMSSD during the TSST between a self-harm group and a HC group. Two studies met Criterion 2.2 and found no significant association between self-harm frequency and HRV during pain stimulation (van der Venne et al., 2023) and during a discussion task (Godfrey et al., 2022). The remaining three studies reporting on HRV during activity were Criterion 3 studies (Crowell et al., 2005; Giner-Bartolome et al., 2017; Otto et al., 2023). Two of these studies did not find a significant association between self-harm and HRV while playing a serious videogame (Giner-Bartolome et al., 2017) or during the MIST (Otto et al., 2023). However, Crowell et al. (2005) found a significant negative association between self-harm and RSA during negative mood induction but found no significant difference in PEP during sad emotion induction between the self-harm and HC group.

Reactivity. Of the 12 studies reporting on HRV reactivity, six were Criterion 1 studies (Fox et al., 2018; Kaufman et al., 2020; Kaufman et al., 2018; Koenig et al., 2023; Koenig, Rinnewitz, Warth, et al., 2017; Lin et al., 2019). Of these studies, three studies found a significant positive relationship between self-harm and HRV reactivity during social-evaluative stress tasks (Fox et al., 2018; Koenig et al., 2023) and social rejection (Kaufman et al., 2018), such that a self-harm group showed higher RMSSD reactivity (Koenig et al., 2023) and higher RSA reactivity (Fox et al., 2018; Kaufman et al., 2018) compared to HCs. Two studies did not find a significant association between self-harm and HRV reactivity (Kaufman et al., 2020; Lin et al., 2019). The study of Koenig, Rinnewitz, Warth, et al. (2017) also showed that the self-harm-group and non-self-harm group did not differ on HRV reactivity from baseline to painful stimulation (CPT). However, this study found the groups to differ in HRV reactivity 60 seconds before CPT onset with the self-harm group showing lower RMSSD reactivity, but these differences were no longer significant 30 seconds before CPT onset. There was one Criterion 2.1 study (Gratz et al., 2019) and three Criterion 2.2 studies (Godfrey et al., 2022; James et al., 2024; Wielgus et al., 2016) that studied the relationship between self-harm and HRV reactivity. Three studies found no relationship between self-harm and HRV reactivity in response to negative emotion induction (Gratz et al., 2019), a stress paradigm (Wielgus et al., 2016), and a discussion paradigm (James et al., 2024). One study, however, also studied mother-daughter conflict discussion and found that lifetime self-harm was related to lower HRV (PEP) reactivity during a conflict discussion task (Godfrey et al., 2022). There were two Criterion 3 studies, of which Reitz et al. (2015) also found that the self-harm group showed lower HRV reactivity from baseline to stress induction compared to the non-self-harm group. However, Crowell et al. (2017) found greater RSA reactivity in self-injuring adolescents affected by their mother's behaviour during a discussion task but not in depressed or typical adolescents.

Recovery. Of the seven studies that reported on HRV recovery, three studies were Criterion 1 (Koenig et al., 2023; Koenig, Rinnewitz, Warth, et al., 2017; Lin et al., 2019). Two of these studies found that self-harm was not related to HRV recovery (Koenig et al., 2023; Lin et al., 2019). Koenig, Rinnewitz, Warth, et al. (2017), however, found the self-harm and non-self-harm groups to significantly differ on

HRV recovery 30 seconds after pain stimulus (CPT) onset, with a lower HRV in the self-harm group, but these differences were no longer significant 60 seconds after CPT onset. There was one Criterion 2.2 study that found that lifetime self-harm and self-harm in the past six months prior to study participation was not correlated with HRV (RSA) recovery (Wielgus et al., 2016). However, they found that RSA recovery did predict self-harm in the past 6 months prior to study participation when controlling for lifetime self-harm, depressive symptoms, baseline RSA and RSA reactivity, such that lower HRV (RSA) during the recovery period was associated with increased risk of self-harm. The remaining three studies were Criterion 3 studies, of which two studies found HRV recovery to be negatively associated with self-harm (Crowell et al., 2005; Reitz et al., 2015). However, Otto et al. (2023) found no significant difference in HRV following the MIST between a self-harm and a HC group.

Electrodermal activity

In total, 14 studies reported on the relationship between self-harm and parameters of EDA. Overall, 29% of the results showed a significant relation between self-harm and EDA.

Rest. Of the eight studies which reported on EDA in rest (Aldrich et al., 2018; Bohus et al., 2000; Crowell et al., 2012; Crowell et al., 2005; Fitzpatrick et al., 2020; Morley, 2000; Tatnell et al., 2018; Tuna & Gencoz, 2021), one met Criterion 1 (Aldrich et al., 2018), which did not find a significant association between EDR in rest and lifetime self-harm, nor self-harm in the past six months. One study met Criterion 2.1 and found no differences in SCL between young adults with self-harm and HCs (Tatnell et al., 2018). There were two studies that met Criterion 2.2, which did not find significant differences in SCL between young adults that self-harmed in the past year versus HCs (Tuna & Gencoz, 2021) and between nail-biting students versus non-nail-biting students (Morley, 2000). There were four studies that met Criterion 3 (Bohus et al., 2000; Crowell et al., 2012; Crowell et al., 2005; Fitzpatrick et al., 2020). While Crowell et al. (2005) did not find a significant difference in EDR at rest between a group of parasuicidal adolescents and HCs, Crowell et al. (2012) found that self-harm was associated with lower EDR in rest when comparing a self-harm group with a depressed and HC group. In contrast, Bohus et al. (2000) found HCs to have a significantly lower EDR in rest compared to a self-harm group in self-reported distress. However, this study did not find a significant difference in EDR in rest between HC and a self-harm group in self-reported calmness. Furthermore, one cross-sectional study found that higher baseline SCR predicted greater frequency of self-harm (Fitzpatrick et al., 2020).

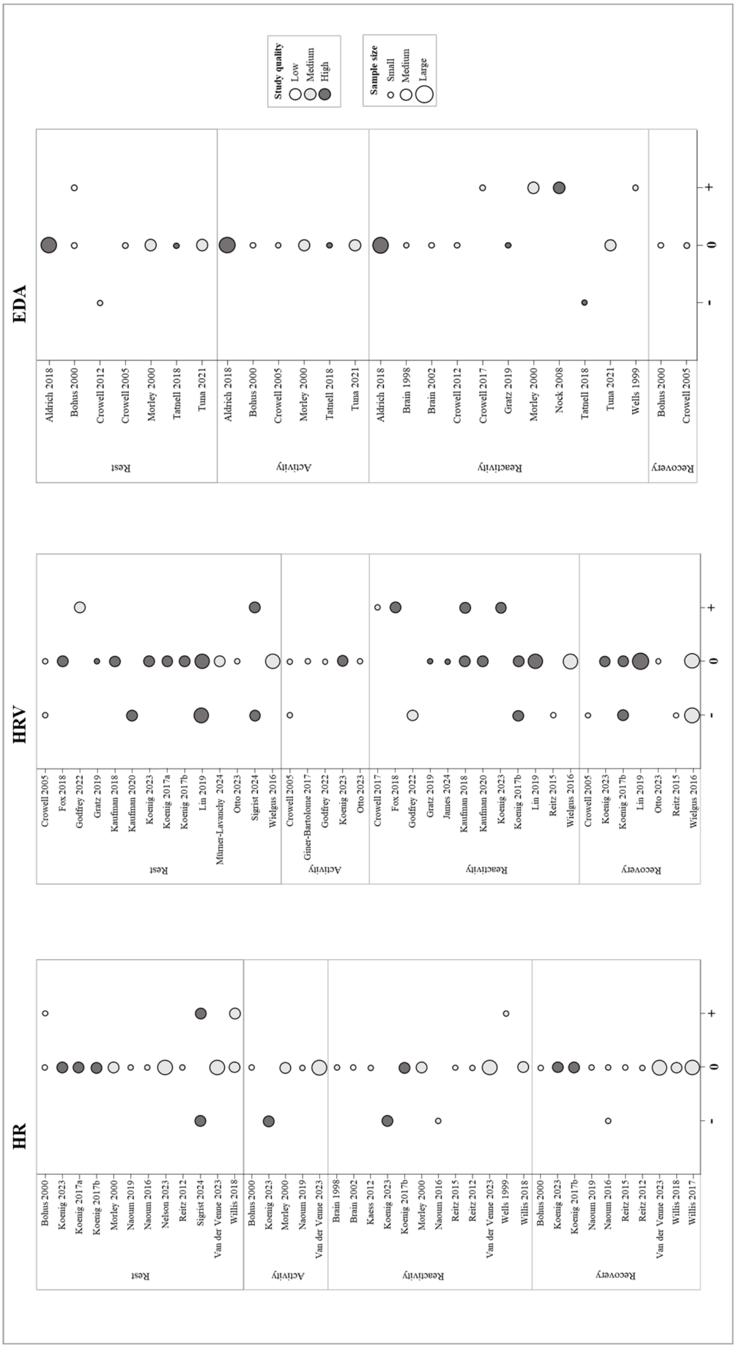
Activity. Of the six studies that examined EDA during activity (Aldrich et al., 2018; Bohus et al., 2000; Crowell et al., 2005; Morley, 2000; Tatnell et al., 2018; Tuna & Gencoz, 2021), one study met Criterion 1 and this study found that EDR during a stress task (anagram task) was not related to lifetime self-harm or self-harm in the past six months (Aldrich et al., 2018). There was one study that met Criterion 2.1 (Tatnell et al., 2018) and two studies that met Criterion 2.2 (Morley, 2000; Tuna & Gencoz, 2021) and these studies did not find a significant association between self-harm and SCL during any phase of TSST (Tatnell et al., 2018), during the CPT (Tuna & Gencoz, 2021), nor during a serial subtraction math task, a mirror image star tracing task or a speech task (Morley, 2000). However, in the study by Morley (2000) there was a significant interaction between nail-biting status and sex in a stress-inducing math task, indicating that male nail biters had higher EDA than male non-nail biters and female nail biters and non-nail biters on EDA during this task.

Reactivity. Eleven studies reported on EDA reactivity. One of these studies met Criterion 1 (Aldrich et al., 2018), which found that EDR during a stress task (solving anagrams) was not related to self-harm incidents in the past six months. There were two studies that met Criterion 2.1 (Gratz et al., 2019; Tatnell et al., 2018). Gratz et al. (2019) found no significant differences in change in SCR in response to negative emotion induction between self-harming young adults and clinical controls without self-harm history. However, Tatnell et al. (2018) did find a significant relationship between self-harm and EDR, as they found that relative to EDR before the TSST, after the TSST the self-harm group showed dampened EDR compared to the HC group. Two studies met Criterion 2.2 (Morley, 2000; Tuna & Gencoz, 2021), of which one study did not find a significant difference in SCL reactivity between the self-harm and HC group (Tuna & Gencoz, 2021). Morley (2000), however, found higher SC reactivity for nail-biters compared to non-nail-biters for a math task, a mirror image star tracing task and a speech task. Six studies met Criterion 3 (Brain et al., 1998; Brain et al., 2002; Crowell et al., 2012; Crowell et al., 2017; Nock & Mendes, 2008; Wells et al., 1999), of which three did not find significant differences in EDA reactivity between the (frequent) self-harm and non (frequent) self-harm groups across imagery (Brain et al., 1998; Brain et al., 2002), and a sad emotion induction task (Crowell et al., 2012). In contrast, the other three Criterion 3 studies found self-harm to be significantly related to higher EDA reactivity (Crowell et al., 2017; Nock & Mendes, 2008; Wells et al., 1999). Wells et al. (1999) found higher EDA reactivity across stages of imagery scripts in the self-cutting group and severe nail-biting group than in the mild nail-biting group. Crowell et al. (2017) studied mother-daughter conflict discussion and found that mother's behaviours affected EDR of self-harming adolescents, but not depressed or typical adolescents. Nock and Mendes (2008) reported that the self-harm group showed a significantly greater increase in SCL level than the non-self-harm group.

Recovery. Finally, two studies reported on EDA during recovery and both studies met Criterion 3 (Bohus et al., 2000; Crowell et al., 2005). The studies reported no significant difference in EDR between the HCs and the self-harm groups following the CPT (Bohus et al., 2000) and sad emotion induction (Crowell et al., 2005).

Figure 2.

Overview of the Results of the Studies with a Between-subject Design



Note. *Some studies reported on different sub-analysis with regards to the relationship between self-harm and physiology which can result in two different results in the figure*

0: no significant result, - : significant negative association, + : significant positive association

Within-subject studies

Of the 13 within-subject studies, ten reported on HR (Barrera et al., 2007; Brain et al., 1998; Brain et al., 2002; Calamari et al., 1990; Freeman et al., 1999; Hall et al., 2013; Hoch et al., 2013; Koenig et al., 2018; Lydon et al., 2013; Wells et al., 1999), two on HRV (Koenig et al., 2018; Welch et al., 2008), and four on EDA (Brain et al., 1998; Ferguson et al., 2019; Welch et al., 2008; Wells et al., 1999). Two types of studies characterized the majority of these studies: lab-studies, using imagery scripts (*n* = 4) (Brain et al., 1998; Brain et al., 2002; Welch et al., 2008; Wells et al., 1999), and studies focusing on the pattern of physiological parameters pre, during and/or post actual occurrences of self-harm (*n* = 6) (Barrera et al., 2007; Calamari et al., 1990; Ferguson et al., 2019; Freeman et al., 1999; Hall et al., 2013; Hoch et al., 2013; Lydon et al., 2013). Furthermore, one study examined the relation between physiology and frequency of self-harm in a longitudinal manner (Koenig et al., 2018). Due to a larger variation in study designs than between-subject studies (see above), the results regarding within-subject studies are presented by type of study.

Imagery studies

Imagery studies used personalized guided imagery scripts depicting an actual episode of self-harm. Participants had their eyes closed during imagery presentation and were instructed to concentrate on the imagery details. The scripts, which were presented in a counterbalanced order, consisted of four stages: scene setting, approach, incident, and consequence. A significant increase in psychophysiological arousal between scene setting and approach was demonstrated for HR (Brain et al., 1998; Brain et al., 2002; Wells et al., 1999)[7]. For SCL and SCR no significant change (Brain et al., 1998; Welch et al., 2008) or a significant increase was found (Wells et al., 1999) at this stage. At the incident stage when actual self-harm was imaged, a decrease was found for HR (Brain et al., 1998; Brain et al., 2002; Wells et al., 1999), SCL (Brain et al., 1998; Wells et al., 1999) and SCR (Welch et al., 2008). The reduction in physiological parameters was maintained at the final stage when the immediate consequences of the act of self-mutilation were imaged in terms of HR (Brain et al., 1998; Brain et al., 2002; Wells et al., 1999), SCL (Brain et al., 1998; Wells et al., 1999) and SCR (Welch et al., 2008). No significant changes in HRV (RSA) were found between stages (Welch et al., 2008).

Studies assessing actual occurrences of self-harm

Six studies focused on the pattern of physiological parameters prior, during and/or post actual occurrences of self-harm, which were all case-studies. One study found an increase in HR *prior* to self-harm (Barrera et al., 2007) and one study found an increase in EDA prior to self-harm (Ferguson et al., 2019). However, three

7 From the study of Wells et al. (1999), the results of the self-cutting group are described. For the results on the nail-biting groups, see Appendix 3C.

studies found no significant change in HR prior to self-harm (Freeman et al., 1999; Hoch et al., 2013; Lydon et al., 2013). One study found mean HR to be lower *during* self-harm than during non-self-harm (Barrera et al., 2007). The study of Lydon et al. (2013) reported for one participant no significant differences in HR *during* occurrences of self-harm versus non-occurrences but for the other participant HR *during* self-harm was higher than HR during non-occurrences. Also, three studies (Freeman et al., 1999; Hall et al., 2013; Lydon et al., 2013) found that self-harm resulted in a (slight) increase in HR. However, one study found a decrease in HR *after* self-harm (Barrera et al., 2007). Ferguson et al. (2019) found that EDA returned to baseline after self-harm in 56% of the cases. Moreover, one study found no significant associations between self-harm and HR post incident (Hoch et al., 2013).

Changes over time

A longitudinal study showed that changes in frequency of self-harm were not associated with changes in HR nor to changes in HRV from baseline to one-year follow-up (Koenig et al., 2018).

DISCUSSION

This systematic review aimed to provide a comprehensive overview of the current knowledge on the association between self-harm and physiology. A large variation in study design, experiments, physiological parameters and measurement instruments was found. The relation between self-harm and three physiological parameters was analysed: HR, HRV and EDA.

Main results

We included 46 studies, of which 37 used a between-subject comparison design and 13 studied the association using a within-subject comparison design. Regarding studies using a between-subject design, we found no consistent relation between EDA , HR and self-harm. However, studies on HRV showed indications for differences in physiology between people with and without self-harm, especially for recovery measures. Lower HRV in people with self-harm during recovery was found across studies that reported significant differences between people with and without self-harm.

Regarding studies using a within-subject design we found interesting differences between imagery studies, and studies examining physiology prior, during and after actual self-harm. While, studies examining physiology prior, during and after actual occurrences of self-harm found contradictory results. Imagery studies found changes in HR and SC across different stages of self-harm imagery scripts with an increased HR between scene setting and approach stages, and a decrease in HR and SC when and after actual self-harm was imaged.

Interpretation of results

Our finding that HRV showed more often significant results in relation to self-harm than EDA and HR is in line with the outcomes of the studies by Goreis et al. (2023) and Bellato et al. (2023). This results might be related to the different branches of the ANS (sympathetic, parasympathetic) the various physiological parameters are related to (Behnke et al., 2022) or to sample characteristics (studies on HRV included mainly only females whereas studies on EDA included mainly mixed samples).

With regard to HRV recovery, several studies reported poorer recovery/ preserved arousal after a painful or stressful experience in people with self-harm. This physiological response could possibly indicate that these persons have difficulties to regulate their physiological response and to regulate their emotions, which can result in turning into maladaptive strategies, such as self-harm. This supports the emotion regulating function of self-harm reported by observational and self-report-studies (Nock, 2010). The function of emotion-regulation is further supported by imagery studies that found a decrease in HR and SCL/ SCR when and after an actual self-harm incident was depicted. This decrease in arousal can be considered as a (negative intrapersonal) reinforcement for self-harm behaviour. However, it should be noted that HRV is influenced by many factors such as medication, circadian rhythms, core body temperature, metabolism, the sleep cycle and hormones and most included studies did not control for all these factors (Pham et al., 2021; Reyes del Paso et al., 2013; Shaffer et al., 2014).

Although this review found preliminary support for a relation between self-harm and ANS functioning, most results were non-significant and heterogeneous. This might be partially explained by classification bias: multiple studies included participants in the self-harm group when they only had engaged in self-harm once in their lifetime. This subgroup may differ in their physiological profile from those individuals that recently or frequently engage in self-harm, with persons engaging more frequent in self-harm having larger alterations in physiology (Willis et al., 2018). However, very few studies examined (frequent) current or recent (within last 6 months) self-harm, which might explain non-significant results.

Another factor which could contribute to the non-significant and heterogeneous results might be that various forms of self-harm were included in most studies. Self-harm has been hypothesized to serve different functions, each of which may be characterized by different physiology profiles. For instance, downregulation of high arousal and upregulation of low arousal most likely result in different physiological patterns. Similarly, self-harm as a means to regulate arousal or self-harm to regulate interpersonal situations might result in different physiological profiles (Fox et al., 2018; Nock, 2010). Bellato et al. (2023) found no association between EDA and self-harm when all types of self-harm were considered together, but specifically non-suicidal self-harm showed increased EDA, while studies that did not distinguish between different types of self-harm (suicidal ideation, suicidal self-harm, and non-suicidal self-harm) reported reduced EDA.

Future research

Several recommendations for future research can be given on the basis of the results of the current systematic review. First, studies are needed that focus on understanding the relation between physiological parameters and different types and functions of self-harm. Future studies could examine physiology separately for different types of self-harm, with different functions (as assessed with functional assessment) to explore whether distinct physiological patterns can be observed. Furthermore, studying the combination of physiological parameters with social or personal factors in relation to self-harm could enhance the understanding of risk factors and functions of self-harm, and improve the prediction of self-harm. For instance, Nelson et al. (2023) found that resting HR was not associated with self-harm in main effects analyses, but the interaction between peer conflict and resting HR at baseline significantly predicted the frequency of self-harm at follow-up, such that adolescents who experienced greater peer conflict and who exhibited a higher resting HR were at greater risk for self-harm across 1 year of early adolescence.

Second, future studies could examine the effectiveness of multimodal predictive models with a combination of physiological parameters when studying the relationship between self-harm and physiology. All studies included in this review used univariate analysis, studying physiological parameters separately. Ghiasi et al. (2020) proposed new indices that combine HRV and EDA. They demonstrated that the understanding of the balance between sympathetic and parasympathetic nervous system activity can be improved by combining certain measures derived from HRV and EDA. They measured HRV and EDA in healthy controls during pain and exposure to emotional stimuli. Results showed significant differences in the *combined value* of HRV and EDA between the periods when participants were subjected to the pain induction and when they were in a resting state beforehand. Furthermore, Imbiriba et al. (2023) studied a combination of cardiovascular activity, EDA, and motion with a wearable biosensor in youths with autism (42.8% intellectual disability). Results showed that, using machine learning with time-stamped mobile behaviour annotation and a combination of physiological signals, could predict self-harm three minutes prior to self-harm behaviour. Also research into detecting stress using machine learning with a combination of physiological parameters have achieved high accuracy in classifying stress levels (Masino et al., 2019; Patlar Akbulut et al., 2020).

Third, very few studies used daily-life assessment. Advancements in technology have improved the validity and reliability of measurements of ANS activity using wearable technology, enabling the possibilities of daily-life assessment. However, from the current study it was evident that the majority of studies included lab-based experiments. Lab-based research increases our understanding of differences in ANS reactivity to pain or environmental stressors between people who self-harm versus people without history of self-harm. However, these studies have limited ecological validity and provide no direct information of physiology before, during and after self-harm incidents. Therefore, future research should focus on measuring physiology in daily life, before, during and after self-harm to better understand the relationship between ANS activity and self-harm.

Fourth, more longitudinal research and larger samples are needed. This review included only four longitudinal studies. Longitudinal studies would improve our knowledge into the trajectories of alterations in physiology with regard to self-harm (Goreis et al., 2023). Moreover, only four of the included studies had a large sample. Also, larger samples are required to gain more insight in the role of comorbidity in altered physiology (Goreis et al., 2023), because self-harm has a high psychiatric comorbidity (Ghinea et al., 2020) but few studies yet controlled for comorbidity.

Fifth, future studies should also include people with intellectual disabilities. While, self-harm is highly prevalent among individuals with intellectual disabilities, only a few studies to date have studied the relation between self-harm and physiological parameters in this population. All of the studies included in this review that studied the relation between self-harm and ANS activity in individuals with intellectual disabilities were case studies. Even though case studies (using direct observations) offer valuable insights into the relation between self-harm and physiology, case studies are prone to bias regarding inclusion. They are often based on prevalent cases and on a sample of convenience (Porta, 2014). The risk of bias and quality assessment showed that all case-studies indeed had a degree of risk of bias regarding inclusion. To make more robust conclusions regarding the association between self-harm and physiology in this population and inform clinical practice, larger-scale studies are needed and meta-analyses can be used to better understand single case design across subjects.

Sixth, future research could compare physiological correlates of self-harm and physical aggression. Self-harm and aggression often co-occur, this co-occurrence is called dual-harm. The cognitive-emotional model of dual-harm highlights that emotion regulation is a function of both self-harm and aggression (Shafti et al., 2021). Noticeable, for aggression, heart rate in rest has repeatedly found to be the best physiological predictor (Lorber, 2004; Ortiz & Raine, 2004; Portnoy & Farrington, 2015), this was not found for self-harm in the current systematic review. Direct comparison between physiological correlates of self-harm and physical aggression has not yet been studied, future studies comparing physiological parameters in relation to aggression and self-harm could provide important knowledge on functions and markers which could be used for prevention and treatment.

Finally, an important factor to consider in future studies is the composition of gender. The majority of studies included only females, except for studies on EDA, in which the majority studied mixed samples. Studies often excluded males because of the moderating role of gender. Because of this little is known about physiology, especially HR and HRV, in males who self-harm.

Limitations

A limitation of the study is the quality and risk of bias assessment, because of different study designs we used different standardized quality assessment instruments. The majority of studies were assessed with the NOS. This scale is often used in systematic reviews. Stang (2010) wrote a critical evaluation of the NOS. Based on this criticism, we adjusted a number of items to improve the quality assessment (see Appendix 3B). Furthermore, we independently scored all studies with two authors, discussed any differences and reached good interrater-agreement. However, some items were still vulnerable to subjective interpretation or difficult to apply to all different studies and no differential weighting of items was applied. The quality assessment should therefore be interpreted with caution.

Clinical implications

The results provide support for the importance of emotion regulation therapy for persons who self-harm. Coping strategies in general and avoidance-based strategies in particular may be important targets for the treatment of self-harm (Giordano et al., 2023; Guerdjikova et al., 2014). A mixed-method study by Ataie et al. (2022) investigated the effectiveness of emotion regulation therapy in adolescents with self-harm behaviour. After intervention participants quit self-harming behaviours and stated that they had replaced avoidance with new strategies such as distraction and acceptance to regulate their emotions.

Unfortunately, current knowledge regarding physiology in relation to self-harm incidents in daily life is still insufficient to use physiological data, for example by using wearable biosensors, for assessment and prevention of self-injury in clinical practice.

What this paper adds

This systematic review is the first review that provides a comprehensive overview of studies on the relation between self-harm and physiology, with no limitations to study design and including studies with persons with intellectual disabilities. Imagery studies found changes in HR and SC across different stages of self-harm imagery scripts. Moreover, between subject studies on HRV showed preliminary indications for differences in physiology between people with and without self-harm, especially for recovery measures, possibly suggesting problems in emotion-regulation in people with self-harm. No consistent findings were found of studies in which self-harm was studied before, during or after actual occurrences of self-harm. Though advancements in technology have improved the validity and reliability of measurements of ANS activity using wearable technology, thereby enabling the possibilities of daily-life assessment, the majority of studies included lab-based experiments with limited ecological validity. Furthermore, although persons with intellectual disabilities are at increased risk of self-harm, they were only included in a minority of studies. Future research should therefore focus on measuring physiology in daily life before, during and after self-harm in persons with intellectual disabilities to better understand the relationship between ANS activity and self-harm. This could result in improved risk assessment, prevention and treatment of self-harm.

Appendix 3A

#	Query OVID (PsycINFO, EMBASE, Medline)
1	(Psychophysiolog* or physiolog* or "autonomic nervous system" or "sympathetic" or "parasympathetic" or arousal).mp.
2 HR(V)	("heart rate" or heartrate* or "cardiovascular reacti*" or "cardiac reacti*" or "cardiovascular respons*" or "cardiac respons*" or electrocardiograph* or HRV or "pre-ejection period*" or "respiratory sinus arrhythmia").mp.
3 SC	(GSR or electrophysiolog* or galvanic or electrodermal or psychogalvanic or ((skin or dermal) adj4 (response* or electr* or conduct* or resist* or potential*))).mp.
4 Self-harm	("self-harm*" or selfharm* or "auto-mutilati*" or automutilati* or "self-mutilati*" or selfmutilati* or "auto-aggressi*"or autoagressi* or "self-injur*" or selfinjur* or "para-suicid*" or parasuicid* or NSSI or "self-inflict*" or selfinflict*or "self-damag*" or selfdamag* or "self-destruct*" or selfdestruct* or ((harm* or hurt* or injur*) adj2 oneself)).mp.
5 Physiology	1 or 2 or 3
6 Physiology + self-harm	4 and 5

#	Query Web of Science
1	TS=(Psychophysiolog* OR physiolog* OR "autonomic nervous system" OR "sympathetic" OR "parasympathetic" OR arousal)
2 HR(V)	TS=("heart rate*" OR heartrate* OR "cardiovascular reacti*" OR "cardiac reacti*" OR "cardiovascular respons*" OR "cardiac respons*" OR electrocardiograph* OR HRV OR "pre-ejection period* OR "respiratory sinus arrhythmia")
3 SC	TS=(GSR OR electrophysiolog* OR galvanic OR electrodermal OR psychogalvanic OR ((skin OR dermal) NEAR/3 (response* OR electr* OR conduct* OR resist* OR potential*)))
4 Self-harm	TS=("self-harm*" OR selfharm* OR "auto-mutilati*" OR automutilati* OR "self-mutilati*" OR selfmutilati* OR "auto-aggressi*" OR autoagressi* OR "self-injur*" OR selfinjur* OR "para-suicid*" OR parasuicid* OR NSSI OR "self-inflict*" OR "selfinflict*" OR "self-damag*" OR self-damag* or "self-destruct*" OR selfdestruct* OR ((harm* OR hurt* OR injur*) NEAR/1 oneself))

Appendix 3B

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE CASE CONTROL STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate (SIB or no SIB)?

a) Yes, with independent validation *
Observation with second observer with agreement classified as very good, such as Cohen's kappa of 0.80 (Cohen, 1960; McHugh, 2012) or higher.
Or a questionnaire with, in psychometric studies established, positive evidence for reliability and validity.
b) Yes, e.g. record linkage or based on self-reports
Self-developed questionnaire, non-validated questionnaire or interview. Observation without second observer or interrater reliability below 0.80.
c) no description

2) Quality of outcome measures: physiology

a) Secure record *
HR/HRV: ECG (with two or more electrodes) or PPG
EDA: measured with electrodes placed on finger, palm of the hand or foot (Ferreira et al., 2023; Hossain et al., 2022; van Lier et al., 2020).
b) HR/HRV not measured with ECG or PPG
EDA not measured on fingers, palm of the hand or foot.
c) no description

3) Representativeness of the cases

a) Consecutive or clearly representative series of cases *
The study used a representative sampling method (such as random sampling): The occurrence of non-response or dropout was minimal, and if it did happen, it occurred randomly.
b) No representative sampling method (for example convenience sample) was used.
High percentage of dropout or non-response: which was non-random or not further examined.
Not enough information to evaluate the representativeness of cases.

4) Selection of Controls

(Comparable strategies of recruitment were used for both the case and control group.)
a) The controls consist of nearly the same population as the case population*
b) No comparable population or for example: healthy controls without further explanation
(No comparable strategies of recruitment of the SIB and control group were used.)

5) Definition of Controls

a) No history of disease *
In the controls, there is no current or recent (in the past half year) SIB
b) In the controls, there is current or recent (in the past half year) SIB

Comparability

1) Comparability of cases and controls on the basis of the design or analysis
a) Study controls for gender (in a mixed sample) *or the groups are matched or only males or only females are included.

b) Study controls for any additional factor * such as: borderline personality disorder severity (if borderline personality disorder population), age, mental disorder (PTSD, depressive disorder), etc. Or the groups are matched on one additional factor.

Exposure

1) Ascertainment of exposure

a) The effect of the exposure is measured and established (in lab studies: for example stress task is perceived as stressful, CPT is experienced as painful.) *

b) No ascertainment of exposure/ nothing reported on this matter.

2) Same method of ascertainment for cases and controls

Same method (for example same questionnaire) used to measure SIB in case and control.

a) Yes *

b) No

Note: SIB= self-injurious behaviour.

Description of adjustments:

Beside operationalization for this systematic review, we made several adjustments based on the literature:
Item 2: Deeks et al. (2003) concluded that the NOS is "suitable for use in a systematic review" and is "easy to use", but information on reliability and validity is lacking". Therefore, we added an item to evaluate the reliability and validity of the outcome measure.
Item 4 was changed based on Stang (2010), who stated "epidemiologic methods teach us that there cannot be a general preference of community controls over hospital controls because the study base principle drives the decision to sample controls from hospitals or from communities".
In the first item of Exposure we removed 'blinding' because blinding is impossible in many of the included studies, this in line with Stang (2010). In addition, the effect of a researcher on the outcome measure, physiology, is minimal and therefore blinding is less relevant.
We removed the last item of exposure (Non-Response rate) based on Stang, (2010). *'Identical response proportions of the case and control group is no safeguard against selection bias.'* Therefore we added the criteria of *random* non-response/dropout' in the third item of selection.

Appendix 3C
Table 3

Overview of included studies investigating physiological correlates of self-harm through between-subject comparisons

Author and year Country	Study design	Sample characteristics	Self-harm	Physiology	Experiment	Results	Comments
HR							
Bohus et al. (2000) Germany	Case-control study	12 patients with BPD 100% female M age (SD): 29.1 (8.4) 19 HCs 100% female M age (SD): 27.3 (7.8)	At least three incidents of self-harm in the past 2 years	ECG was measured using electrodes placed over the middle of the right collarbone and below the left lowest rib using the Kolner Vitaport I System	CPT	*Rest:* HR before the CPT significantly differed between HCs and the self-harm group in self-reported distress, with the self-harm-group having higher HR. There was no significant difference in HR before the CPT between HCs and the self-harm group in self-reported calmness. *Activity:* There were no significant differences in HR between the HCs and the self-harm group in self-reported calmness and in self-reported distress during the CPT. *Recovery:* There were no significant differences in HR between the HCs and the self-harm group in self-reported calmness and in self-reported distress after the CPT.	

Brain et al. (1998) Australia	Case-control study	Total N = 70 57% female 15 individuals with current self-harm 20 individuals with retrospective self-harm 35 controls with no history of self-harm	Multiple types A self-developed 'self-mutilative behaviors checklist'	ECG was measured using two electrodes fitted at the second rib on both sides of the torso	For individuals with self-harm imagery scripts of 4 separate events were constructed, namely: an actual episode of self-harm, accidental injury, an angry interaction and a low arousal neutral event. For control participants only scripts regarding an accidental injury, anger, and neutral event were constructed.	*Reactivity:* No significant differences between the self-harm group and the control group in psychophysiological response to control imagery were demonstrated. No significant differences between the current self-harm group and the retrospective self-harm group in psychophysiological response to self-mutilation imagery.

Brain et al. (2002) Australia	Case-control study	43 adults with a history of self-harm M age (SD): 23.5 (1.2) 58% female 29 Frequent self-harm participants (>5 life time self-harm events) 14 Infrequent self-harm group (<5 life time events)	Multiple types A self-developed 'self-mutilative behaviors checklist'	ECG was measured using two Gereonics 7 mm Ag/AgCl electrodes fitted at the second rib on both sides of the torso	Personalized guided imagery scripts depicting a neutral script and a self-harm script (actual episode of self-mutilation) were presented in 4 stages: scene setting, approach, incident, and consequence	*Reactivity:* No difference in psychophysiological response to imagery between frequent and infrequent groups.
Kaess et al. (2012) Germany	Case-control study	Total N = 28 100% female 14 patients who engaged in self-harm M age (SD): 16.6 (1.7) 14 HCs M age (SD): 16.3 (2.2)	Multiple types FASM	The HR monitor Polar RS 800 (Polar) was fixed at the chest	TSST, comprising of a mock job interview and a mental arithmetic exercise	*Reactivity:* The HR increased significantly in both groups. The difference in HR response between the self-harm and non-self-harm was not statistically significant.

Koenig et al. (2023) Germany	Case-control study	Total N = 60 adolescents 13 to 17 years old 100% female 30 adolescents with self-harm 30 HCs	Multiple types On average, patients engaged in self-harm on 89.03 (SD: 144.90; 5–720) days in the past year SITBI-G	ECG were recorded using an ekgMove 3 sensor attached to participants' chest at the base of the sternum using a flexible belt with two integrated electrodes	TSST	*Rest:* There was no significant difference between the self-harm group and HC group in HR at baseline. *Activity:* There was no significant main effect of group on HR, indicating that on average over all time points there were no differences in HR between the self-harm group and the HC group. However, there was a significant difference in HR between the groups during the interview and mental arithmetic task, with HC group having a significantly higher HR than the self-harm group during these stress tasks. *Reactivity:* There was a significant interaction between group and time, indicating that the pattern of HR change over time differed between the self-harm group and HC group. Adolescents with self-harm showed lower HR change in response to stress compared to adolescents without self-harm. *Recovery:* There was no significant difference in HR between the self-harm group and HC group in the recovery phase.

Koenig, Rinne-witz, Parzer, et al. (2017) Germany	Case-control study	Total N = 60 100% female 30 adolescents with self-harm *M* age (SD): 15.3(1.4) 30 HCs *M* age (SD): 15.3 (1.3)	Multiple types, engaging in self-harm on at least five days within the last year according to DSM-5 criteria SITBI-G	Polar® RS800CXTM	Only baseline	*Rest:* The self-harm group and HC group did not significantly differ on HR. HR was unrelated to acts of self-harm (r = -.138).	60% of the self-harm group fulfilled full diagnostic criteria for BPD.
Koenig, Rinne-witz, Warth, et al. (2017) Germany	Case-control study	Total N = 60 100% female 30 adolescents with NSSI *M* age (SD): 15.3 (1.4) 30 HCs *M* age (SD): 15.3 (1.3)	Multiple types, with skin lesion at least once in the 6 months prior to study and at least 5 incidents in the past 12 months SITBI-G	HR Polar RS800CV chest strap	CPT	*Rest:* The self-harm and HC groups did not differ on HR at baseline. *Reactivity*: The self-harm and HC groups did not differ on HR reactivity from baseline to CPT, reactivity 60 s before CPT onset, or reactivity 30 s before CPT onset. *Recovery:* Groups did not differ on HR reactivity 30 s after CPT onset, or 60 s after CPT onset.	

Morley (2000) USA	Case-control study	Total N = 83 students 42 nail biters 52% female 41 non-nail biters 54% female	Nail biting Habit questionnaire designed by the experimenter	HR PPG sensor placed on index finger of the dominant hand using Davicon MEDAC System/3	Serial subtraction math task, mirror image star tracing task, and speech task	*Rest:* There was no main effect (or interaction effect with sex) for nail-biting status in baseline level of HR, indicating that there was no group difference in HR at baseline. *Activity:* There was no main effect (or interaction effect with sex) for nail-biting status on HR activity in any of the tasks, indicating that there were no group differences in HR during the math task, star task or speech task. *Reactivity:* There was no main effect (or interaction effect with sex) for nail-biting status for HR reactivity on any of the tasks, indicating that there were no group differences in HR response during the math task, star task or speech task.	Controlled for sex

Naoum et al. (2019) Germany	Case-control study	17 patients with BPD and NSSI with skin lesion *M* age (SD): 33.0 (9.2) 0% female 20 HCs *M* age (SD): 32.1 (7.3) 0% female	Multiple types, with skin lesion at least once in the 6 months prior to the study QNSSI	HR ECG amplified with a Biosemi Active Two AD-Box	Concentration performance test followed by a non-invasive nociceptive mechanical stimulus	*Rest:* No significant difference in HR at baseline between the self-harm group and the HC group (d = 1.57). *Activity*: There was no significant difference between the groups in HR during the concentration performance test. *Recovery:* Both groups showed a decrease in HR over time, but there was no significant difference in HR decrease between the groups during the recovery phase following the concentration performance test.	Originally, 20 HCs were included. However, due to technical problems heart rate data was available for 19 HCs.

Naoum et al. (2016) Germany	Case-control study	20 patients with BPD and NSSI M age (SD): 27.1 (8.9) 100% female 20 HCs M age (SD): 27.0 (8.4) 100% female	Multiple types NSSI within 6 months prior to study participation QNSSI	HR ECG amplified with Biosemi Active Two AD-box	Modified version of the concentration performance test followed by either a pain stimulus without blood or a pain stimulus and artificial blood	*Rest:* There was no difference in HR at baseline between the self-harm and HC group (*d* = -0.12). *Reactivity:* There was an increase in HR during the concentration performance test in both the self-harm and HC group, but this increase was greater in the HC group than in the self-harm group *Recovery:* Overall, HR decreased over time, but there was no difference in HR decrease following the concentration performance test between HC and the self-harm groups. However, there was a significant condition*group*time interaction, with a greater decrease of HR in HCs in the blood condition over time than in the BPD patients .
Nelson et al. (2023) USA	Cohort study	N = 147 M age (SD): 12.3 (0.6) 51% female	Multiple types Lifetime NSSI NSSI was assessed with five questions based on past research	HR LifeSource UB-351 Wrist Digital Blood Pressure Monitor	Only baseline	*Rest:* There was no significant association between resting HR at baseline and NSSI at one year follow-up (after controlling for NSSI at baseline). However, there was a significant interaction between peer conflict and resting HR at baseline on NSSI at one year follow-up (after controlling for self-harm at baseline), such that the combination of higher peer conflict and higher resting HR was associated with greater NSSI after one year.

Reitz et al. (2015) Germany	Case-control study	17 patients with BPD with current NSSI *M* age (SD): 26.0 (6.9) 100% female 11 HCs *M* age (SD): 26.9 (8.3) 100% female	Multiple types NSSI within 6 months prior to study participation QNSSI	HR Finger mounted pulse oximetry	MIST followed by an incision or sham	*Rest and activity:* There was no main effect of group in HR over baseline and stress induction phase ($f^2 = 0.30$), indicating that there was no overall difference in HR between the self-harm and HC groups over the baseline and stress induction. *Reactivity:* Testing the effect of stress induction on HR, there was no interaction effect between time and group ($f^2 = 0.04$), indicating no difference in HR response following stress reduction between the self-harm and HC groups. *Recovery*: Testing the immediate effect of incision on HR (after stress induction to 7 minutes after incision/sham), there was no significant condition*time*group interaction. Testing the intermediate effect of incision on HR (after stress induction, and 7 min, 14 min, and 21 min after incision/sham), there was no significant interaction between condition*time*group.	Originally, 21 participants in the BPD group and 17 in the HC group were included. Due to technical problems HR data was available for 17 BPDs and 11 HCs.

Reitz et al. (2012) Germany	Case-control study	14 patients with BPD and high frequency self-harm M age (SD): 26.9 (5.9) 100% female 18 HCs without NSSI history M age (SD): 26.4 (7.4) 100% female	Multiple types High frequency self-harm (at least once per week) Borderline Symptom List 95	HR Varioport Version-B (chestbelt with ECG)	Modified MIST combined with the presentation of pictures of the International Affective Picture System followed by an incision or sham	*Rest:* HR at baseline did not differ between the self-harm and HC groups. *Rest and activity:* In the analyses to test the effect of stress induction on HR, there was no interaction effect for condition*group (f^2 = .13), nor a main effect of group (f^2 = .13) over the baseline and stress induction phase, indicating that there was no overall difference in HR between the self-harm and HC groups over the baseline and stress induction. *Reactivity:* In the analyses to test the effect of stress induction on HR, there was no interaction between condition*time*group (f^2 = .26), nor an interaction between time*group (f^2 = .18), indicating there were no significant differences between the self-harm and HC groups in HR response to stress induction. .	Because technical problems, HR data of only 12 BPD and 16 HCs was included in the analyses.

						Recovery: In the analyses to test the short-term effects of incision on HR recovery after stress induction, there was no interaction between condition*time*group (f^2 = .39), nor between time*group (f^2 = .30). For the analyses to test the long-term effects of incision on HR recovery after stress induction, there was no significant interaction between condition*time*group (f^2 = .45), or time*group (f^2 = .53). Additionally, there was no difference in time needed to return to baseline of HR between the groups.	
Sigrist et al. (2024) Germany	Case-control study	30 adolescents with NSSI M age (SD): 15.1 (1.1) 100% female 30 HCs M age (SD): 14.8 (1.5) 100% female	Multiple types Five or more incidents of NSSI during the past year, and one incident or more during the past month SITBI-G	HR ECG Move III chest belt	Diurnal variation	*Rest:* Significant differences in rhythm-adjusted mean levels of HR, as well as significant acrophase shifts (the lag from a predefined reference timepoint) were observed between the NSSI and HC groups. Additionally, significant differences in amplitudes of HR between the groups were found. Overall, the NSSI group showed a significantly higher rhythm-adjusted mean HR, and lower HR amplitude compared to HC. Furthermore, the NSSI group compared to HCs, the acrophase of HR was shifted significantly, such that peak levels in HR (morning peak) were reached at a later time point in the NSSI group.	ECG data were missing for one participant in the NSSI group, due to significant sensor non-wear time. When controlling for age, BMI, physical activity and cardiac data quality, group remained a significant predictor of rhythm-adjusted mean levels of HR, but not of amplitude or acrophase of HR.

van der Venne et al. (2023) Germany	Case-control study	164 adolescents with NSSI M age (SD): 14.8 (1.5) 100% female 45 HCs M age (SD): 14.8 (1.3) 100% female	Multiple types One lifetime incident of NSSI NSSI frequency in the past six months SITBI-G	HR ECG assessed with EcgMove 3 sensor attached to a chest belt	Pain assessment: Participants placed their hand on a plate of 32 ºC. Following a 3-min adaptation phase, temperature steadily rose to 50 ºC over 4 min. Participants were asked to keep their hand firmly on the plate until the pain became intolerable.	*Rest:* There were no significant differences in HR at baseline by NSSI frequency. *Activity:* There were no significant differences in HR during pain stimulation by NSSI frequency. *Reactivity:* The interaction between time and NSSI frequency was not significant, indicating that HR from baseline to pain stimulation did not differ by NSSI frequency. *Recovery:* There were no significant differences in HR 5 minutes after pain stimulation by NSSI frequency.	However, when controlling for ACE revealed a significant time*NSSI frequency interaction, indicating a progressive decrease in HR with increasing NSSI frequency during pain.

Wells et al. (1999) Australia	Controlled-trial	15 individuals with severe history of nail-biting 15 individuals with mild history of nail-biting 67 individuals with a history of self-cutting 43% female	Nail biting and self-cutting The severe group included a high intensity of nail-biting and a large degree of physical damage. The mild group included lower intensity of nail-biting and absence of physical damage A questionnaire was devised to record the severity and frequency of nail-biting and other self-mutilative behaviours	HR MacLab/8s data acquisition system using electrodes placed on each side of the rib cage along the lateral line with an earth on the mastoid process.	Personalized imagery scripts of an actual self-mutilative episode divided into four stages: setting the scene, approach, incident, and consequence	*Reactivity*: There was a significant stage by group interaction for HR, demonstrating that HR reactivity across stages differed across the groups. For the self-cutting group, HR increased from the first stage to the approach, then decreased through the approach stage to the actual incident and remained low for the consequence stage. For the severe nail-biting group, HR increased from the first stage to the approach stage. No across-stage differences in HR were apparent for the mild biting group.	A range of self-harming behaviours other than nail-biting were noted in all three groups.

Willis et al. (2018) Germany	Case-control study	30 patients with current BPD M age (SD): 28.0 (6.1) 100% female 30 patients with remitted BPD M age (SD): 29.0 (4.5) 100% female 30 HCs M age (SD): 28.7 (5.5) 100% female	Multiple types Current BPD group: NSSI with skin lesions at least once during 6 months prior Remitted BPD group: Two or less acts of NSSI in the last 2 years, but used NSSI before Custom made questionnaire assessing the frequencies and forms of NSSI	HR ECG with BioSemi Active Two AD-box	Modified version of the MIST followed by a small incision, a blade stimulus or a sham stimulus and a relaxation phase	*Rest:* There was no difference in HR at baseline between the remitted BPD and HC groups. There was a significant difference between the current BPD and HC groups, with current BPDs having a higher HR at baseline than HCs. *Rest and activity:* There was no main effect of group on HR during stress inducting, indicating no significant difference between any of the groups over the baseline and stress induction phases. *Reactivity:* There was no interaction between time and group for HR, indicating no difference in HR reactivity between groups from baseline to post stress induction. *Recovery:* In the analysis to test whether pain leads to a reduction of stress, there was no time*pain intensity*group interaction, indicating that the HR decrease did not differ between groups and was not related to pain perception directly after stimulus application, nor during the entire relaxation period.	As there is overlap between the current BPD group and the remitted BPD group in NSSI, only the comparisons with the HC group are reported.

Willis et al. (2017) Germany	Case-control study	57 BPD patients M age (SD): 27.9 (7.9) 100% female 60 HCs M age (SD): 27.5 (7.1) 100% female	Multiple types NSSI with skin lesions in the 6 months prior to study Custom made questionnaire assessing frequencies and forms of NSSI	HR ECG with Biosemi Active Two AD-Box	Modified version of the MIST followed by a small incision, a blade stimulus or a sham stimulus and a relaxation phase	*Recovery:* There was no difference in HR levels between the self-harm and HC groups directly after stimulus application. There were no time*group or time*group*stimulus interaction effects for HR from after stress induction to immediately after stimulus application, nor from after stress induction to 30 minutes after stimulus application.	Because of technical problems, HR data of only 55 patients in the self-harm group and 58 HCs could be analysed.
HRV							
Crowell et al. (2005) USA	Case-control study	Total N = 46 adolescents M age (SD): 15.3 (1.1) 100% female 23 self-harm adolescents 23 typical control adolescents	Multiple types LPC	RSA: Biopac MP100 system, ECG PEP: Biopac MP100 system, ECG and ICG signals Two electrodes were placed on the left ankle and the right wrist to obtain the ECG signal	Sad emotion induction using a 3 min clip from *The Champ*	*Rest:* Compared with HCs, self-harm adolescents exhibited significantly lower RSA at baseline. There was no difference between the self-harm and HC groups for PEP at baseline. *Activity:* Compared with HCs, self-harm adolescents exhibited significantly lower RSA during negative mood induction. There was no difference between the self-harm and HC groups for PEP during the sad emotion induction. *Recovery:* Compared to HCs, self-harming adolescents had significantly lower RSA during recovery.	Because of equipment failure, physiological data was missing for one self-harm and one HC participant.

Crowell et al. (2017) USA	Case-control study	Total N = 76 Female 100% 26 self-harming adolescents M age (SD): 16.3 (1.0) 24 depressed non-self-harming adolescents M age (SD): 15.6 (1.4) 24 typical adolescents M age (SD): 16.1 (1.3)	Multiple types L-SASI	RSA ECG signals were obtained using a BioPac MP100 system	10-min mother-daughter conflict discussion	*Reactivity:* The mother's behaviour affected RSA response during conflict of self-harming adolescents, but not depressed or typical adolescents.	
Fitzpatrick et al. (2020) USA	Cross-sectional study	22 self-harming individuals with BPD M age (SD): 32.7 (9.6) 63.6% female	Multiple types Engaged in self-harm ≥2 in the past five years, with at least one of those episodes occurring in the past 10 weeks. Lifetime Suicide Attempt-Self Injury Interview (LSASII)	RSA BIOPAC 5-channel acquisition system (BIOPAC Systems Inc., Model MP150, Goleta, CA). RSA was measured through the high frequency band of spectral analysis of HRV using Mindware Technologies HRV 2.33 software.	Participants were trained in two emotion regulation strategies: Distract and Notice 96 trials divided into 4 blocks Images: neutral and negative	*Baseline* There were statistically significant main effects of baseline RSA on self-harm frequency, such that higher baseline RSA predicted greater frequency of self-harm. Incidence rate ratios indicated that, for every unit increase in baseline RSA, the frequency of self-harm increased by 1.82. Correlation self-harm frequency- baseline RSA= .32	BPD severity was entered as a covariate

Fox et al. (2018) USA	Case-control study	Total N = 70 young adults M age (SD): 19.3 (0.9) 94% female 35 young adults with a recent history of self-harm 35 young adults who had never engaged in self-harm	Multiple types Modified version of the ISAS	RSA ECG was measured using Biopac MP150 Data Acquisition Unit. Electrodes were placed on chests and abdomens.	Stressor task (preparation of public speech)	*Rest:* Resting RSA was not significantly related to recent NSSI (*OR* = 0.79). *Reactivity:* RSA withdrawal demonstrated a significant relation with recent self-harm (*OR* = 3.84). Participants who did not engage in self-harm in the past year exhibited lower RSA reactivity than those who engaged in self-harm.	
Giner-Bartolome et al. (2017) Spain	Case-control study	Total N = 66 100% female 12 participants with an ED and self-harm M age (SD): 26.8 (8.8) 31 participants with an ED without self-harm M age (SD): 31.8 (9.7) 22 HC M age (SD:) 28.6 (7.9)	Self-cutting, burning, hitting, and scratching Clinical interview	HRV Two ECG electrodes placed above the ninth-tenth rib on the right and left side of the chest	A serious videogame named *Playmancer*	*Activity:* No significant differences in in-task HRV were found between ED patients with and without self-harm (*d* = 0.46) and between ED patients with self-harm and HC (*d* = 0.24).	Adjusted for age, BMI, depression score, and ED subtype.

Godfrey et al. (2022) USA	Case-control study	Total N = 60 adolescents 100% female 30 adolescents with NSSI M age (SD): 15.5 (1.5) 30 HCs M age (SD): 14.8 (1.3)	Multiple types A history of ≥3 lifetime NSSI. L-SASII	Cardiac PEP ECG and ICG (Mindware hardware)	Discussion task	*Rest:* Lifetime NSSI was significantly correlated with basal PEP (*r* = .23). *Activity:* Lifetime NSSI was not correlated with PEP during conflict (*r* = .13). *Reactivity:* Lifetime NSSI was a significant predictor of conflict PEP (when controlling for maternal support and baseline PEP). The interaction between maternal emotional support and lifetime NSSI was also significant, indicating that maternal emotional support was associated with lengthened conflict PEP among individuals with no lifetime NSSI, yet was not associated with conflict PEP among those with high lifetime NSSI. Maternal support was associated with lengthened PEP for participants up to 22 lifetime NSSI incidents, while participants with more than 22 past NSSI incidents did not demonstrate PEP lengthening in response to maternal support. The NSSI group was not significantly related to PEP reactivity (when controlling for maternal support and baseline PEP). The interaction between maternal emotional support and NSSI status did not predict PEP reactivity.	One participant from the self-harm group and two HCs were excluded from analyses due to artefacts in their physiological data. Those with few lifetime incidents of NSSI (<23) responded more similarly to those without any previous self-harm (as compared to participants who endorsed>22 NSSI.

Gratz et al. (2019) USA and Canada	Case-control study	Total N = 64 *M* age (SD): 23.9 (4.8) 72% female 21 self-harming young adults with BPD 18 self-harming young adults without BPD 25 clinical controls with no self-harm history	Multiple types Recent (i.e., past-year) and recurrent (i.e., 10 lifetime episodes) Deliberate Self-Harm Inventory	The HF component of HRV was derived using ECG acquired with the Biopac MP100/MP150 system	Neutral emotion induction: non-demanding, color-counting task. Negative emotion induction: imaginal social rejection scenario	*Rest:* No significant between-group differences in HRV during the neutral emotion induction were found (η^2_p=.02). *Reactivity:* No significant between-group differences in change in HRV were found in response to the negative emotion induction (η^2_p=.03).	These results did not change when gender was included as a covariate.
James et al. (2024) USA	Case-control study	27 adolescents with a history of NSSI *M* age (SD): 15.2 (1.4) 100% female 33 adolescents without a history of NSSI *M* age (SD): 15.2 (1.3) 100% female	Multiple types Lifetime NSSI SITBI	RSA ECG was recorded using Biopad BioNomadix	Discussion paradigm	*Rest and activity:* There was a significant main effect of NSSI history, such that adolescents with a NSSI history demonstrated higher RSA levels across all tasks (including baseline) than adolescents without this history. *Reactivity:* There was no main interaction effect between time and NSSI history, indicating no significant differences in HRV reactivity from baseline to the tasks.	

Kaufman et al. (2018) USA	Case-control study	Total N = 60 adolescents 30 suicidal adolescents 72% female *M* age (SD): 15.8 (1.2) 30 HCs *M* age (SD): 15.5 (1.0) 57% female	3+ lifetime self-harm, at least one of which needed to meet criteria for a suicide attempt L-SASII	RSA ECG electrodes were placed on the right shoulder and the left abdomen near the bottom of the rib cage	Cyberball, a mild computerized social rejection task	*Rest:* There was no difference between the groups in RSA during the baseline resting period. *Reactivity:* There was no significant interaction between phase and group in the inclusion phase (*d* = -0.17), indicating that there is no difference in RSA reactivity during social inclusion between the self-harm group and the HC group. There was a significant interaction between phase and group in the exclusion phase of the Cyberball task (*d* = -0.23). While the HC group demonstrated a nonsignificant increase in RSA during exclusion, the self-harm group demonstrated significant RSA withdrawal.
Kaufman et al. (2020) USA	Case-control study	Total N = 60 adolescents 100% female 30 adolescents with a history of ≥ 3 self-harm incidents *M* age (SD): 15.5 (1.5) 30 HCs *M* age (SD): 14.8 (1.3)	Multiple types A history of ≥ 3 self-harm incidents L-SASII	RSA ECG using Mindware hardware with a standard spot electrode configuration	1) 5m-in resting individual and dyadic baseline phases, 2) a mother–adolescent conflict discussion, 3) watching an instructional video on validation (GIVE) and 4) a validation-oriented discussion.	*Rest:* Adolescents with self-harm had significantly lower resting RSA during individual baseline (*d* = -0.60) and dyadic baseline than HCs. *Reactivity:* There was no significant interaction between group and task (*d* = -0.08), indicating that there were no significant group differences in RSA reactivity across tasks.

Koenig et al. (2023) Germany	Case-control study	Total N = 60 adolescents 13 to 17 years old 100% female 30 adolescents with self-harm 30 HCs	Multiple types On average, adolescents engaged in self-harm on 89.0 (SD: 144.9; 5–720) days in the past year SITBI-G	RMSSD ECG was recorded using an ekgMove 3 sensor attached to the chest at the base of the sternum using a flexible belt with two integrated electrodes	TSST	*Rest:* There was no significant difference between the self-harm and HC groups in HRV at baseline. *Activity:* There was no significant main effect of group on HRV, indicating that on average over all time points there were no differences in HRV between the self-harm group and the HC group. There was no significant difference in HRV between the groups in any of the stages of the TSST. *Reactivity:* There was a significant interaction between group and time, indicating that the pattern of HRV change over time differed between the self-harm and HC groups. Adolescents with self-harm showed higher HRV responses to stress compared to adolescents without self-harm. *Recovery:* There was no significant difference in HRV between the self-harm group and HC group in the recovery phase.

Koenig et al. (2017a) Germany	Case-control study	Total N = 60 adolescents 100% female 30 adolescents with self-harm M age (SD): 15.3 (1.4) 30 HCs M age (SD): 15.3 (1.3)	Multiple types, engaging in self-harm on at least five days within the last year according to DSM-5 criteria SITBI-G	vmHRV (RMSSD) A Polar® RS800CXTM	Only baseline	*Rest:* The self-harm group and HC group did not significantly differ on vmHRV. vmHRV was unrelated to acts of self-harm (*r* = .146).	60% of the self-harm group fulfilled full diagnostic criteria for BPD.
Koenig, Rinne-witz, Warth, et al. (2017) Germany	Case-control study	Total N = 60 adolescents 100% female 30 adolescents with self-harm M age (SD): 15.3 (1.4) 30 HCs M age (SD): 15.3 (1.3)	Multiple types, with skin lesion at least once in the 6 months prior to study and at least 5 incidents in the past 12 months SITBI-G	RMSSD Polar RS800CV chest strap	CPT	*Rest:* Self-harm and HC group did not differ on HRV at baseline. *Reactivity:* Groups did not differ on HRV reactivity from baseline to CPT. Groups did significantly differ in short-term HRV reactiv-ity 60 s before CPT onset, with the self-harm group showing lower HRV reactivity than HCs. However, these differences were no longer significant 30 s before CPT onset. *Recovery:* Groups significantly differed on HRV reactivity 30 s after CPT onset, with lower HRV reactivity in the NSSI group. These differences were no longer significant 60 s after CPT onset.	

Lin et al. (2019) USA	Case-control study	162 pregnant women Age: 29 (range 18-40) 100% female	Multiple types Lifetime and antenatal SITBs L-SASII	RSA ECG using wireless MindWare mobile devices	Infant cry tasks (seascape baseline, followed by infant play task, infant cry task, and a seascape recovery)	*Rest:* There was no significant correlation between baseline RSA and lifetime SITBs (r = - .07). There was a significant negative correlation (r = -.18) between baseline RSA and antenatal SITBs. *Reactivity:* There was no significant correlation between RSA reactivity during the play condition or cry condition and either lifetime (r = .04, r = .07) or antenatal SITBs (r = .10, r = .05). *Recovery:* There was no significant correlation between RSA reactivity during recovery and either lifetime (r = .04) or antenatal SITBs (r = .04).
Müerner-Lavanchy et al. (2024) Germany	Case-control study	149 patients with NSSI 40 HCs	Multiple types At least 5 episodes of NSSI in the last year SITBI-G	vmHRV (RMSSD) ECG using an ECGMove 3 Sensor	Only baseline	*Rest:* vmHRV during rest did not significantly predict who did and did not engage in NSSI (OR = 0.89).

Otto et al. (2023) Germany	Case-control study	34 participants with NSSI M age (SD): 15.9 (1.4) 94% female 28 HCs M age (SD): 16.0 (1.1) 100% female	Multiple types At least 5 episodes of NSSI in the last year SITBI	HRV (pulse-to-pulse intervals) Siemens MRI-compatible PPG on the right index finger	MIST	*Rest:* There was no significant difference in HRV between the groups during the resting state before the MIST. *Activity:* There was no significant difference in HRV between the groups during the MIST. *Recovery:* There was no significant difference in HRV between the groups during the resting state after the MIST.

Reitz et al. (2015) Germany	Case-control study	17 patients with BPD with current NSSI *M* age (SD): 26.0 (6.9) 100% female 11 HCs *M* age (SD): 26.9 (8.3) 100% female	Multiple types NSSI within 6 months prior to study participation QNSSI	HRV-triangular index Finger mounted pulse oximetry	MIST followed by an incision or sham	*Rest and activity:* There was no main effect of group on HRV over baseline and stress induction phase (f^2 = 0.30), indicating that there was no overall difference in HR between the self-harm and HC groups over the baseline and stress induction. *Reactivity:* Testing the effect of stress induction on HRV, there was an interaction between time and group (f^2 = 0.40), indicating that the change in HRV was different between groups. The self-harm group showed lower HRV reactivity from baseline to stress induction compared to the HC group. *Recovery:* Testing the immediate effect of incision on HRV (after stress induction to 7 minutes after incision/sham), there was a condition*time*group interaction (f^2 = 0.44), indicating a stronger HRV increase in the control group after sham treatment compared with the self-harm group. No group differences were found for the effects of incision. Testing the intermediate effect of incision on HRV (after stress induction, and 7 min, 14 min, and 21 min after incision/sham), there was an interaction between condition*-time*group ($f^2$0.37), in which the control group showed higher HRV in session 2 after sham compared with incision (*d*= 0.64).

Sigrist et al. (2024) Germany	Case-control study	30 adolescents with NSSI M age (SD): 15.1 (1.1) 100% female 30 HCs M age (SD): 14.8 (1.5) 100% female	Multiple types 5 or more incidents of self-harm during the past year, and 1 incident or more during the past month SITBI-G	RMSSD ECG Move III chest belt	Diurnal variation	*Rest:* Significant differences in rhythm-adjusted mean levels of HRV, as well as significant acrophase shifts (the lag from a predefined reference time-point) were observed between the self-harm and HC groups. Additionally, significant differences in amplitudes of HRV between the groups were found. Overall, the NSSI group showed a significantly lower rhythm-adjusted mean HRV, higher HRV amplitude compared to HC. Furthermore, in the NSSI group compared to HCs, the acrophase of HRV was shifted significantly, such that peak levels in HRV (peak at nighttime) were reached at a later time point in the NSSI group.	ECG data were missing for one participant in the NSSI group, due to significant sensor non-wear time. When controlling for age, BMI, physical activity and cardiac data quality, group was no longer a statistically significant predictor of rhythm-adjusted mean HRV, amplitude or acrophase of HRV.

Van der Venne (2023) Germany	Case-control study	164 adolescents with NSSI *M* age (SD): 14.8 (1.5) 100% female 45 HCs *M* age (SD): 14.8 (1.3) 100% female	Multiple types One lifetime incident of NSSI NSSI frequency in the past six months SITBI-G	RMSSD ECG assessed with EcgMove 3 sensor attached to a chest belt	Pain assessment: Participants placed their hand on a plate of 32 ºC. Following a 3-min adaptation phase, temperature steadily rose to 50 ºC over 4 min. Participants were asked to keep their hand firmly on the plate until the pain became intolerable.	*Rest:* There were no significant differences in HRV at baseline by NSSI frequency. *Activity:* There were no significant differences in HRV during pain stimulation by NSSI frequency. *Reactivity:* The overall model fit for HRV was not significant. Therefore, the interaction between time and NSSI frequency cannot be interpreted.	When controlling for depression severity revealed a significant time*NSSI frequency interaction, indicating that HRV increased stronger with greater NSSI frequency following pain induction.

Wielgus et al. (2016) USA	Prospective study	103 adolescents enrolled at middle schools M age (SD): 12.8 (0.8) 53.7% female	Suicidal ideation, suicide attempts or NSSI Lifetime history and in the past 6 months YSR, CDI-2, CBCL and K-SADS-PL	RSA ECG was measured using Biopac MP150 Data Acquisition Unit and electrodes placed on the chest and abdomen	Stress paradigm which involves solving anagrams	*Rest:* Baseline RSA was not significantly associated with lifetime SITBs, nor SITBs in the past 6 months. Baseline RSA did not predict engagement in SITBs between in the past 6 months (when controlling for lifetime SITBs, and depressive symptoms). *Reactivity:* RSA reactivity was not significantly associated with lifetime SITBs, nor SITBs in the past 6 months. RSA reactivity did not predict engagement in SITBs in the past 6 months (when controlling for lifetime SITBs, depressive symptoms, and baseline RSA). *Recovery:* RSA recovery was not significantly associated with lifetime SITBs, nor SITBs in the past 6 months. RSA recovery did predict SITBs in the past 6 months (when controlling for lifetime SITBs, depressive symptoms, baseline RSA and RSA reactivity), such that lower RSA during the recovery period was associated with increased risk for SITBs.

EDA							
Aldrich et al. (2018) USA	Cohort study	Total N = 121 M age (SD): 12.9 (0.9) 55% female Adolescents who engaged in SITBs Adolescents with no self-harm	SITBs (lifetime and in the past six months) CDI-2, YSR, K-SADS-PL, and semi-structured interviews	EDR Two electrodes on the soles of the left foot using Biopac MP150 Data Acquisition Unit	Stress task: anagram task	*Rest:* Baseline EDR was not significantly related to lifetime SITBs (r_{pb} =-0.16), nor to SITBs in the past six months (r_{pb} =-0.16). *Activity:* Stressor EDR was not significantly related to lifetime SITBs (r_{pb} = -0.14), nor to SITBs in the past six months (r_{pb} = -0.13). *Reactivity*: Stressor EDR did not predict SITBs in the past six months when controlling for age, depressive symptoms, lifetime SITBs and baseline EDR (*OR* = 1.68).	The interaction between EDR and impulsivity significantly predicted SITBs in the last six months, such that individuals with low electrodermal responding and high trait impulsivity were more likely to engage in SITBs in the last six months.

Bohus et al. (2000) Germany	Case-control study	12 patients with BPD 100% female M age (SD): 29.1 (8.4) 19 HCs 100% female M age (SD): 27.3 (7.8)	At least 3 incidents of self-harm in the past 2 years	SCRF was measured from thenar and hypothenar of the non-dominant hand using the Kolner Vitaport I System	CPT	*Rest:* SCRF before the CPT significantly differed between HCs and the self-harm group in self-reported distress, with the self-harm-group having higher SCRF. There was no significant difference in SCRF before the CPT between HCs and the self-harm group in self-reported calmness. *Activity:* There were no significant differences in SCRF between the HCs and the self-harm group in self-reported calmness and in self-reported distress during the CPT. *Recovery:* There were no significant differences in SCRF between the HCs and the self-harm group in self-reported calmness and in self-reported distress after the CPT.

Brain et al. (1998) Australia	Case-control study	Total N = 70 57% female 15 individuals with current self-harm 20 individuals with retrospective self-harm 35 controls with no history of self-harm	Multiple types A self-developed 'self-mutilative behaviors checklist'	SCL Two electrodes connected to the fingertips of the first and third fingers of the non-dominant hand	For individuals with self-harm imagery scripts of 4 separate events were constructed, namely: an actual episode of self-harm, accidental injury, an angry interaction and a low arousal neutral event. For control participants only scripts regarding an accidental injury, anger, and neutral event were constructed	*Reactivity:* No significant differences between the self-harm group and the control group in psychophysiological response to control imagery were demonstrated. No significant differences between the current self-harm group and the retrospective self-harm group in psychophysiological response to self-mutilation imagery were demonstrated.

Brain et al. (2002) Australia	Case-control study	Total N = 43 adults with a history of self-harm M age (SD): 23.5 (1.2) 58% female 29 Frequent self-harm participants (>5 life time self-harm events) 14 Infrequent self-harm group (<5 life time events)	Multiple types A self-developed 'self-mutilative behaviors checklist'	SCL Two electrodes connected to the fingertips of the first and third fingers of the non-dominant hand	Personalized guided imagery scripts depicting a neutral script and a self-harm script (actual episode of self-mutilation) were presented in 4 stages: scene setting, approach, incident, and consequence	*Reactivity:* No differences in psychophysiological response to imagery between frequent and infrequent groups were evident.

Crowell et al. (2012) USA	Case-control study	Total N = 75 adolescents 100% female 25 self-harming adolescents *M* age (SD): 16.3 (1.0) 25 depressed adolescents *M* age (SD): 15.6 (1.4) 25 typical control adolescents *M* age (SD) = 16.1 (1.3)	Multiple types L-SASI	EDR Biopac MP100 system, 2 electrodes attached to the thenar eminence of the nondominant hand	Sad emotion induction using a 3 min clip from *The Champ*	*Rest:* self-harm adolescents exhibited lower resting EDR than both the depressed and typical control groups. *Reactivity:* The different groups did not significantly predict changes in EDR across the sad emotion inducing task.
Crowell et al. (2005) USA	Case-control study	Total N = 46 adolescents *M* age (SD): 15.3 (1.1) 100% female 23 self-harming adolescents 23 typical control adolescents	Multiple types LPC	EDR Biopac MP100 system, two electrodes that were attached to the thenar eminence of the nondominant hand	Sad emotion induction using a 3 min clip from *The Champ*	*Rest:* No difference between the self-harm group and the HC group was found for EDR at baseline. *Activity:* No difference between the self-harm group and the HC group was found for EDR during the sad emotion induction. *Recovery:* No difference between the self-harm group and the HC group was found for EDR during recovery.

Crowell et al. (2017) USA	Case-control study	Total N = 76 Female 100% 26 self-harming adolescents M age (SD): 16.3 (1.0) 24 depressed non-self-harming adolescents M age (SD): 15.6 (1.4) 24 typical adolescents M age (SD): 16.1 (1.3)	Multiple types L-SASI	EDR Biopac MP100 system, two electrodes attached to the thenar eminence of the nondominant hand	A 10-min mother-daughter conflict discussion	*Reactivity:* The mothers' behaviours affected EDA response during conflict of self-harming adolescents, but not depressed or typical adolescents.

Fitzpatrick et al. (2020) USA	Cross-sectional study	22 self-harming individuals with BPD M age (SD): 32.7 (9.6) 64% female	Participants were trained in two emotion regulation strategies: Distract and Notice 96 trials divided into 4 blocks Images: neutral and negative	Multiple types Engaged in self-harm ≥2 in the past 5 years, with at least 1 of those episodes occurring in the past 10 weeks. Lifetime Suicide Attempt-Self Injury Interview (LSASII)	Ns-SCR and SCL Two electrodes (BIOPAC Systems Inc., Model EL507) were placed on the medial phalanges of the index and middle fingers of the nondominant hand.	*Baseline* There were statistically significant main effects of baseline SCR on self-harm frequency, such that higher baseline SCR predicted greater frequency of self-harm. Incidence rate ratios indicated that, for every unit increase in baseline and baseline SCR, the frequency of self-harm increased by 1.23. A repeated measures ANOVA examining differences in tonic SCL changes conditions was not statistically significant. Correlations with self-harm frequency: Baseline SCR: .56*; Tonic SCL negative react trials: -.51*; Tonic SCL distract trials: -.31; Tonic SCL notice trials: -.34*.	BPD severity was entered as a covariate

Gratz et al. (2019) USA and Canada	Case-control study	Total N = 64 *M* age (SD): 23.9 (4.8) 72% female 21 self-harming young adults with BPD 18 self-harming young adults without BPD 25 clinical controls with no self-harm history	Multiple types Recent (i.e., past-year) and recurrent (i.e., 10 lifetime episodes) Deliberate Self-Harm Inventory	SCR Electrodes on two fingers on the nondominant hand, via Biopac Systems, MP100/ MP150 systems	Neutral emotion induction: non-demanding, color-counting task. Negative emotion induction: A social rejection stressor: audio-recorded, imaginal social rejection scenario.	*Reactivity:* No significant between-group differences in change in SCR were found in response to negative emotion induction (relative to neutral emotion induction).	Results did not change when gender was included as a covariate.

Morley (2000) USA	Case-control study	Total N = 83 students 42 nail biters 52% female 41 non-nail biters 54% female	Nail biting Habit questionnaire designed by the experimenter	SCL Two sensors to the fingerprint areas of the middle and ring fingers of the dominant hand using Davicon MEDAC System/3	Serial subtraction math task, mirror image star tracing task, and speech task	*Rest:* There was no significant main effect (or interaction effect with sex) for nail-biting status in baseline level of SC, indicating that there was no group difference in SCL at baseline. *Activity:* There was no main effect for nail-biting status on SC activity in any of the tasks, indicating that there were no group differences in SCL during the math task, star task or speech task. However, there was a significant interaction effect with sex in the math task, indicating that male nail biters differed from the other three groups. *Reactivity:* There was a significant main effect for nail-biting status for all tasks (math, star and speech tasks), reflecting higher SC reactivity for nail-biters compared to non-nail biters. There were also significant interaction effects with sex, demonstrating that male nail-biters differed from the other three groups.	Controlled for sex

Nock and Mendes (2008) USA	Case-control study	62 adolescents and young adults with a lifetime history of NSSI M age (SD): 17.4 (1.8) 80% female 30 HCs M age (SD): 16.7 (2.0) 73% female	Multiple types Lifetime history of NSSI SITBI	SCL Biopac TSD203 transducers placed on distal phalanges of the middle and ring fingers of the nondominant hand	DTT and social problem solving skills test	*Reactivity:* Self-harmers exhibited a significantly greater increase in SCL than non-self-harmers over time (*d*=0.57), and this difference became especially pronounced in the later minutes of the DTT, when participants are informed that their answers were consistently incorrect.	The analyses were repeated with other diagnoses that might correlate with SC as covariates and results remained after accounting for these diagnoses.
Tatnell et al. (2018) Australia	Case-control study	Total N = 78 young adults M age (SD): 20.1 (1.7) 67% female 25 young adults with a history of NSSI 53 HCs	Multiple types Lifetime NSSI ISAS	SCL and SCR Biopac Systems-MP150, electrodes on the non-dominant hand.	TSST (preparation, speech task and math task) and images from the International Affective Picture system	*Rest:* There were no differences in SC for participants with and without self-harm during baseline. *Activity:* There were no differences in SC between NSSI groups during any phase of the TSST (preparation, speech and math). *Reactivity:* There was an interaction between image valence, self-harm group and time, indicating that relative to physiological responses before the stressor, after the stressor the self-harm group recorded a dampened physiological response to positive images compared to HCs, who responded more to positive images after the stressor.	Controlled for respiration induced artefacts, diagnosis of mental illness and recent depressive symptoms.

Tuna and Gencoz (2021) Turkey	Case-control study	Total N = 70 young adults *M* age (SD): 21.1 (1.58) 60% female 34 students who have reported NSSI in the past year 36 students who have never engaged in NSSI (HCs)	Multiple types In the past year ISAS	SCL ProComp Infinitiy sensors (Thought Technology, Canada) placed on the distal phalanx of the middle and ring fingers of the left hand	DTT and CPT	*Rest:* There was no difference in SCL between the NSSI and HC groups at baseline. *Activity:* There was no difference in SCL between the groups at the CPT before the DTT, nor at the CPT after the DTT. There was no main effect of group during the DTT. *Reactivity:* There was no interaction effect between group and time, indicating no difference in SCL reactivity between the groups from before the DTT to the 20th card of DTT.

Wells et al. (1999) Australia	Controlled trial	15 individuals with severe history of nail-biting 15 individuals with mild history of nail-biting 67 individuals with a history of self-cutting 43% female	Nail biting and self-cutting The severe group included a high intensity of nail-biting and a large degree of physical damage. The mild group included lower intensity of nail-biting and absence of physical damage A questionnaire was devised to record the severity and frequency of nail-biting and other self-mutilative behaviours	SCL MacLab/8s data acquisition system using electrodes on the first and third fingers of the non-dominant hand	Personalized imagery scripts of an actual self-mutilative episode divided into four stages: setting the scene, approach, incident, and consequence.	*Reactivity:* There was a significant stage by group interaction for SCL, demonstrating that SCL reactivity across stages differed across the groups. For the self-cutting group, there was a significant pattern of arousal change across the stages, with an increase from the scene stage to the approach stage, followed by a decrease from the approach stage to the incident stage when self-cutting was described. For the severe nail-biting group there was a decrease from the setting the scene stage to the consequence stage, and from the incident stage to the consequence stage. For the mild nail-biting group no across-stage differences were apparent.	A range of self-harming behaviours other than nail-biting were noted in all three groups.

Note. NSSI = non-suicidal self-injury; SITBs = self-injurious thoughts and behaviours; Multiple types = the study included participants with a history of various types of self-harm, including both mild and severe forms of self-harm; HR = heart rate; ECG = electrocardiograph; M = mean; SD = standard deviation; HC = healthy control; FASM = Functional Assessment of Self-Mutilation; TPT = Tourniquet Pain Test; TSST = Trier Social Stress Test; SITBI = Self-Injurious Thoughts and Behavior Interview; SITBI-G = The German version of the Self-Injurious; Thoughts and Behavior Interview BPD = borderline personality disorder; CPT = Cold pressor test; PPG = photoplethysmography; QNSSI = Questionnaire for Non-Suicidal Self-injury; MIST = Montreal Imaging Stress Task; HRV = heart rate variability; LPC = Lifetime Parasuicide count; RSA = respiratory sinus arrhythmia; PEP = pre-ejection period; ICG = impedance cardiography; L-SASI = Lifetime-Suicide Attempt Self-Injury; ISAS = Inventory for Statements about Self-Injury; L-SASII = Lifetime-Suicide Attempt Self-Injury Interview; ED = eating disorder; BMI = body mass index; = HF = high frequency; RMSSD = root mean square of successive differences; vmHRV = vagally mediated heart rate variability; DSM-5 = Diagnostic and Statistical Manual of Mental Disorders, Fifth edition; SC = skin conductance; SCRF = skin conductance response fluctuations; CDI-2 = Children's Depression Inventory; YSR = Youth Self-Report; K-SADS-PL = Kiddie Schedule of Affective Disorders and Schizophrenia for School-Age children Present and Lifetime version; CBCL= Child Behavior Checklist'; EDR = electrodermal response; SCL = skin conductance level; SCR = skin conductance response; DTT = Distress Tolerance Test.

Appendix 3D
Table 4

Overview of included studies investigating physiological correlates of self-harm through within-subject comparisons

Author and year Country	Study design	Sample characteristics	Comparison	Self-harm	Physiology	Results	Comments
HR							
Barrera et al. (2007) Canada	Case study	N = 3 1. Age 27 Female ID in the upper profound range 2. Age 42 Female Severe ID 3. Age 33 Male Profound ID	Self-harm present vs. absent Pre-during-post 4 conditions: demand, attention, alone and control	1. Hand-to-head punches and head-to-object hits 2. Hand to head punches and hair pulling 3. Hand-to-head: Punches and slaps, ear pulls, hair pulls. Hand-to-torso: chest slaps. Head-to-object: hits on doors or walls. Observation using video recordings	HR LifeShirt vest and three-lead ECG electrodes attached to the upper chest and left abdomen, with a fourth electrode (temperature sensor) attached to the shoulder blade.	1. HR activity appeared to consistently increase just prior to each self-harm occurrence, and decrease immediately after. On the other hand, random non-occurrences showed no similar or consistent pre-post pattern and tended to be flatter. Mean HR levels during self-harm occurrences were lower than during random non-occurrences (111.1 vs. 132.2 bpm) but less stable.	1. The self-harm HR pattern was not influenced by topographical motion differences and posture. 2. For more information on HR during different conditions (demand, attention, alone and control), see Barrera et al. (2007)

2. HR activity was seen to increase in all cases within 1–2 s just prior to self-harm occurrence, and to decrease during the self-harm burst itself or 2–3 s immediately after or both. Random non-occurrences did not show any characteristic HR or similar pre-post patterns across conditions.
Mean HR levels during self-harm occurrences were lower than during random non-occurrences (87.37 vs. 101.26 bpm).

3. Mean HR levels per condition were positively correlated with higher self-harm frequencies in the experimental conditions, but not in the control conditions.
Self-harm occurrences were all accompanied by a clear reduction of HR activity levels. This drop was very pronounced in three out of conditions. However, the reductive effect was relatively short lasting and did not carry over into the post-self-harm intervals.
In comparison to the pp 1 and 2 findings, pp3's self-harm HR waveform appeared to be a briefer but stronger effect (10–15% variance in most conditions). HR activity during random non-occurrences showed no discernable pre-post pattern and was generally flat.

Bohus et al. (2000) Germany	Experimental study	12 patients with BPD 100% female M age (SD): 29.1 (8.4) 19 HCs 100% female M age (SD): 27.3 (7.8)	CPT during conditions of self-reported calmness (BPD-C) and during self-reported distress (BPD-D).	At least 3 incidents of self-harm in the past 2 years	HR ECG was measured using electrodes placed over the middle of the right collarbone and below the left lowest rib using the Kolner Vitaport I System.	Before the CPT (base rates), HR did not differ between the two conditions. During and after the CPT there were no significant differences.
Brain et al. (1998) Australia	Experimental study	N = 70 57.14 % female N = 15 currently self-harm group N = 20 retrospective self-harm group N = 35 control group	Imagery scripts	Multiple types A self-developed 'self-mutilative behaviors checklist' Cutting was the most frequently reported method of self-harm (97%)	HR ECG was measured using two electrodes fitted at the second rib on both sides of the torso.	Significant between stage differences were demonstrated for the self-harm script HR, significant increase in psychophysiological arousal between stage 1 and stage 2 was demonstrated for HR, Arousal decreased significantly at stage 3 when actual self-harm was imaged HR, This reduction in psychophysiological arousal was maintained at stage 4 when the immediate consequences of the act of self-harm were imaged.

However, the reductive effect was relatively short lasting and did not carry over into the post-self-harm intervals. In comparison to the pp 1 and 2 findings, pp3's self-harm HR waveform appeared to be a briefer but stronger effect (10–15% variance in most conditions). HR activity during random non-occurrences showed no discernable pre-post pattern and was generally flat.

Brain et al. (2002) Australia	Experimental study	Adults with a history of self-harm N = 43 M age (SD): 23.5 (1.2) 58% female Frequent self-harm participants (>5 life time self-harm events, N = 29)	Personalized guided imagery scripts depicting a neutral script vs. a self-harm script: an actual episode of self-harm was presented in 4 stages: scene setting, approach, incident, and consequence.	Multiple types A self-developed 'self-mutilative behaviors checklist'	HR ECG was measured using two Gereonics 7 mm Ag/AgCl electrodes fitted at the second rib on both sides of the torso.	Psychophysiological arousal increased significantly from stage 1 to stage 2 for HR. Arousal decreased significantly from stage 2 to stage 3, when actual self-harm was depicted for HR. Arousal remained at this lower level from stage 3 to stage 4. No significant variation in psychophysiological arousal was evident across the 4 stages of the neutral script.
Calamari et al. (1990) USA	Case study	N = 1 23 years Female Severe ID	Before and after propranolol	Self-biting, scratching, or head banging Recording by direct care staff as part of the facility's standard treatment protocol.	HR Measurement method is not reported.	Increases in resting HR correlated with development of increased self-harm, and correspondingly subsided as improvements in behaviour were achieved.

Freeman et al. (1999) USA	Case study	N = 2 Age 33, 38 Male Severe ID	Before and after self-harm	1. Biting his own wrist 2. Hitting, banging or slapping head, chest, legs or walls with his hand, biting his finger Observation using video recordings	HR Polar Vantage XL heart rate monitor, electrocardiogram. Chest belt.	1. *Preceding* Results show a low likelihood of HR increase in the 15 seconds preceding self-bite, comparable to no-self-harm. *Recovery* The probability of HR increase 15 sec. after self-bite was significantly higher than after no-self-harm. *Absolute HR* On days when his average HR was low- there were higher levels of problem behaviour and the probability of HR increase *following* self-bite became consistently higher. 2. *Preceding* Results show a low likelihood of HR increase in the 15 seconds preceding self-harm, comparable to no-self-harm. *Recovery* The probability of HR increase 15 seconds following finger bite and bang/slap was significantly higher compared to no self-harm.	2. The behaviour requiring the least amount of physical exertion had the highest probability of HR increase.

Hall et al. (2013) USA	Case study	N = 1 Boy diagnosed with PraderWilli syndrome Age: 12 Mild ID	Duration of self-harm Trend following the onset of self-harm Measurements during 5 types of sessions: ignore, alone, attention, play, and demand	Skin picking Self-Injury Trauma Scale (SIT) Observation using video recordings	HR Polar XR belt around his chest	There was a significant positive association between skin picking and HR levels (τ= .48). Out of a total of 15 bouts of skin picking recorded across ignore and alone sessions, 11 bouts resulted in a significant increase in HR. None of the bouts resulted in a decrease in HR.	There was no association between HR and activity levels (τ=.02) indicating that the association between skin picking and HR was not confounded by physical activity. Interestingly, skin picking which produced tissue damage and bleeding were reliably associated with increased HR's whereas skin picking which did not result in tissue damage was not associated with an increase in HR.

Hoch et al. (2013) USA	Case study	N = 1* Age: 17 Female Severe or profound ID		Skin Picking Observation using video recordings	HR Polar chest belt	No significant associations between self-harm and HR indices were found	* N=3, but only one was >12 years old and included
Koenig et al. (2018) Germany	Longitu-dinal	18 adolescents with NSSI M age (SD): 15.2 (1.4) 100% female	Number of NSSI acts at baseline and follow-up after intervention	Multiple types At least 5 incidents of NSSI during the past 12 months Self-injurious thoughts and behavior interview (SITBI-G)	HR Polar RS800CXTM Monitor	Change in mean HR was not correlated with change in NSSI acts.	
Lydon et al. (2013) Ireland	Case study	2 males with ASD and frequent challenging behaviour Participant 1: 12-year old boy with mod-erate ID Participant 2: 16-year old boy with cog-nitive delay	Participant 1: self-harm and tantrum behaviour vs. non-occur-rence Participant 2: self-harm vs. non-occur-rence	Participant 1:Self-biting, body-hit-ting Participant 2: Self-bit-ing Observer rated	HR Polar© RS800cx HR chest strap and wrist watch	Participant 1: There was no HR pattern that consistently preceded self-harm, while post-self-harm increases were most common, accompanying 4 of the 7 occurrences. Participant 2: There was no HR pattern that consistently preceded self-harm. Increases were most common, but did not deviate from chance levels. Following self-harm, increases were most common, although most were slight.	

3

Wells et al. (1999) Australia	Experimental study	Experiment 1: 15 individuals with severe history of nail-biting 15 individuals with mild history of nail-biting 15 individuals with no history of nail-biting Experiment 2: 15 individuals with severe history of nail-biting 15 individuals with mild history of nail-biting 67 individuals with a history of self-cutting 43% female	Experiment 1: Imagery of a nail-related event vs. a neutral event Experiment 2: The 4 stages of an imagery of an actual self-harm episode (setting the scene, approach, incident, and consequence)	Experiment 1: Nail-biting Experiment 2: Nail-biting and self-cutting The severe group included a high intensity of nail-biting and a large degree of physical damage. The mild group included lower intensity of nail-biting and absence of physical damage A questionnaire was devised to record the severity and frequency of nail-biting and other self-harming behaviours	HR MacLab/8s data acquisition system using electrodes placed on each side of the rib cage along the lateral line with an earth on the mastoid process.	Experiment 1: There was a trend toward a script by stage interaction for HR response. A difference for the severe group on the nail-related script between the scene and approach stages was evident. No such difference across stages was observed for the mild and control groups. There was a difference between the overall response to the nail-related script and the neutral scripts at the approach stage. Experiment 2: In the self-cutting group, HR increased from the first stage to the approach, then decreased through the approach stage to the actual incident of self-cutting. HR remained low for the consequence stage but did not decrease further. In the severe nail-biting group, there was a significant increase from the first stage to the approach stage prior to biting beginning. Although a reduction in HR was noted after this stage, this decrease was not significant. There were no differences between stages for the mild biting group.	A range of self-harm behaviours other than nail-biting were noted in all three groups

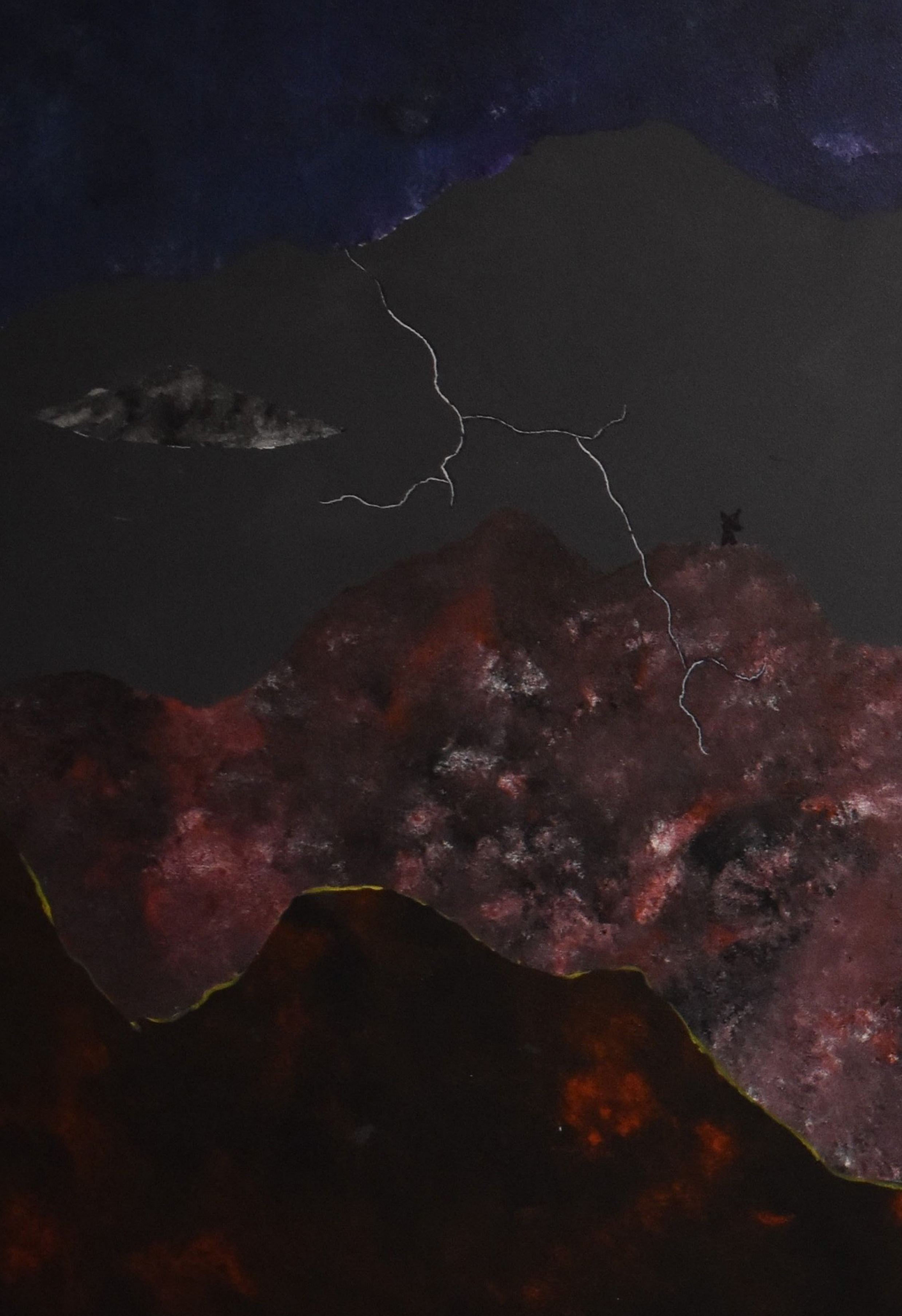
HRV						
Koenig et al. (2018) Germany	Longitudinal	18 adolescents with NSSI M age (SD): 15.2 (1.4) 100% female	Number of NSSI acts at baseline and follow-up after intervention	Multiple types At least 5 incidents of NSSI during the past 12 months Self-injurious thoughts and behavior interview (SITBI-G)	HRV (vmHRV) Polar RS800CXTM Monitor	Change in vmHRV was not correlated with change in NSSI acts.
Welch et al. (2008) USA	Experimental study	42 individuals with BPD and NSSI or suicide attempt (SA) M age (SD): 31 (9.5) 91% female	Imagery of an episode of NSSI vs. imagery of an accidental injury, an emotionally neutral event, an SA, and an accidental death Approach vs incident stage Incident vs consequence stage	Multiple types Lifetime history of suicide attempt or NSSI, with at least one of those acts in the past year Lifetime Parasuicide Count	HRV (RSA)	No significant differences in RSA were found between stages.

SC						
Brain et al (1998) Australia	Experi-mental study	N = 70 57.14 % female N = 15 current self-harm group N = 20 retro-spective self-harm group N = 35 control group	Imagery scripts	Multiple types A self-developed 'self-mutilative behaviors checklist' Cutting was the most frequently reported method of self-harm The total sample of self-harming participants had deliberately injured themselves a median of 30.5 times (range = 3–350) over a mean period of 51.8 months (SD = 59.2, range = 2.5 months to 30 years).	SCL was measured via two electrodes connected to the fingertips of the first and third fingers of the non-dominant hand.	Significant between stage differences were demonstrated for the self-harm script. A significant increase in psychophysiological arousal between stage 1 and stage 2 was not demonstrated for SCL. Arousal decreased significantly at stage 3 when actual self-harm was imaged. This reduction in psychophysiological arousal was maintained at stage 4 when the immediate consequences of the act of self-harm were imaged.
Bohus et al. (2000) Germany	Experi-mental study	12 patients with BPD 100% female *M* age (SD): 29.1 (8.4) 19 HCs 100% female *M* age (SD): 27.3 (7.8)	CPT during conditions of self-reported calmness (BPD-C) and during self-reported dis-tress (BPD-D).	At least 3 incidents of self-harm in the past 2 years	SCRF Skin conductance was recorded from thenar and hypothenar of the non-dominant hand.	Before the CPT (base rates), SCRF in patients during distress differed significantly from SCRF in BPD patients when calm, with lower SCRF before CPT during calmness. During and after the CPT there were no significant differences.

Ferguson et al. (2019) USA	Case study	8 individuals with severe ASD and moderate ID *M* age (SD): 15.9 (2.5) 0% female	EDA before, during and after challenging behaviour(among which self-harm)	Multiple types Observations over 1 year.	EDA Q-Sensor pod wristband on wrist or ankle	*Preceding* Individuals displayed an anticipatory rise in EDA prior to engaging in self-harm: percent of times anticipatory rise prior to behaviour: 89%. *Recovery* Percent of times EDA returned to baseline after self-harm: 56%.	See Ferguson et al. (2019) for mean duration of return to baseline and mean increase prior to self-harm.
Welch et al. (2008) USA	Experimental study	42 individuals with BPD and NSSI or suicide attempt (SA) *M* age (SD): 31 (9.5) 91% female	Imagery of an episode of NSSI vs. imagery of an accidental injury (AI), an emotionally neutral event, an SA, and an accidental death Approach vs incident stage Incident vs consequence stage	Multiple types Lifetime history of suicide attempt or NSSI, with at least one of those acts in the past year Lifetime Parasuicide Count	SCR	From approach to incident stage no significant difference in SCR was found. A significant decrease in SCR was found from incident to consequent stage.	A contrasting pattern was found in the control AI script in which SCR increased from approach to incident. Decreases in the AI script after injury imagery were not statistically significant. There were no significant changes in SCR between the approach/ incident and incident/consequence stages within the neutral imagery.

3

Wells et al. (1999) Australia	Experi-mental study	Experiment 1: 15 individuals with severe history of nail-biting 15 individuals with mild history of nail-biting 15 individuals with no history of nail-biting Experiment 2: 15 individuals with severe history of nail-biting 15 individuals with mild history of nail-biting 67 individuals with a history of self-cutting 43% female	Experiment 1: Imagery of a nail-related event vs. a neutral event (in 4 stages: setting the scene, approach, incident and consequence) Experiment 2: The 4 stages of an imagery of an actual self-harm episode (setting the scene, approach, incident, and consequence)	Experiment 1: Nail-biting Experiment 2: Nail-biting and self-cutting The severe group included a high intensity of nail-biting and a large degree of physical damage. The mild group included lower intensity of nail-biting and absence of physical damage A questionnaire was devised to record the severity and frequency of nail-biting and other self-harming behaviours	SCL MacLab/8s data acquisition system using electrodes on the first and third fingers of the non-dominant hand	Experiment 1: A significant script by stage interaction was observed for SCL, with a difference in response to the nail-related script between the scene and the consequence stages, the approach and consequences stages, and the incident and consequences stages. Experiment 2: In the self-cutting group, there was an increase from the scene stage to the approach stage, followed by a decrease from the approach stage to the incident stage. Arousal remained low for the consequence stage. In the severe nail-biting group significant decreases were apparent from the scene stage to the consequence stage and from the incident stage to the consequence stage. In the mild nail-biting group there were no differences between stages.	A range of self-harming behaviours other than nail-biting were noted in all three groups



Chapter 4

Changes in heart rate and electrodermal activity after aggression in residential treatment facilities: A multicentre study

This chapter is submitted as: Nijman, I.*, van Swieten, M.*, Bogaerts, S., Didden, R., Embregts, P., Hagoort, K., Hasselman, F., Koldijk, S., Konijn, C., Masthoff, E., VanDerNagel, J., Nijhof, K., Noordzij, N., Popma, A., Scheepers, F., de Schepper, F., Smeets, K., Strijbosch, E., & de Looff, P. (2025). *Changes in heart rate and electrodermal activity after aggression in residential treatment facilities: A multicentre study.*

* Shared first author, these authors contributed equally to this work

ABSTRACT

Background: Aggressive behaviour frequently occurs in residential treatment settings and can have severe negative psychological and physical consequences on clients and staff members. Previous studies found that physiological arousal, measured using parameters that reflect autonomic nervous system functioning such as heart rate and electrodermal activity, increases during the period leading up to an aggressive incident. Less is known about the course of physiological arousal after aggression. Understanding physiological arousal post-aggression is important, as lingering physiological activation may increase the risk of subsequent aggression through excitation transfer.

Method: The current multi-centre study examined the trajectory of physiological arousal after an aggressive incident occurred among children, adolescents and adults in clinical treatment centres. To this end, data from three studies on physiological arousal and aggression were combined for secondary analyses which involved 60 clients who showed 165 aggressive incidents. Care staff monitored the aggressive behaviour of the clients. Using the Empatica E4, heart rate (HR) and electrodermal activity (skin conductance level [SCL] and skin conductance response [SCR]) were measured up to 30 minutes after the aggressive incident.

Results: Longitudinal multilevel models showed that, on average, heart rate (b = -0.60, p = .008), SCL (b = -0.10, p = .002), and SCR (b = -0.11, p = .002) decreased over the course of 30 minutes after aggressive incidents. However, despite an overall decrease in physiological parameters, latent class growth curve analyses revealed that subgroups of clients exist that follow different trajectories in heart rate and electrodermal activity after the aggressive incidents, showing decreasing, stable, or increasing patterns of physiological arousal.

Conclusion: The findings suggest heterogeneity in post-aggression physiological responses which highlights the need for individualized approaches in managing aggression.

INTRODUCTION

Aggressive behaviour from clients is common in residential treatment facilities (Bowring et al., 2017; Frowijn et al., 2024; van Swieten et al., 2024; Verstegen et al., 2020; Weltens et al., 2021). In the current study, we defined aggression as behavioural acts, either verbal or physical, intended to harm another person, oneself or an object (Bjorkqvist & Niemela, 1992). A systematic review found that 65 to 99% of staff members in residential treatment settings experienced some form of aggression during their careers, involving physical aggression in 38 to 82% of cases (Weltens et al., 2021). The percentages of clients who exhibited aggression ranged from 7.5 to 75.9%, with a weighted mean of 23%. An increased risk for aggression was found in clients with mild intellectual disability or borderline intellectual functioning (MID-BIF; Nieuwenhuis et al., 2022; Weltens et al., 2021).

Aggression can have severe negative psychological and physical consequences for both the client showing the behaviour, group members and support staff. For the client, it may hinder rehabilitation by affecting the therapeutic alliance (Stevenson & Taylor, 2020) and requiring safety measures, such as seclusion, to protect all involved (Cowman et al., 2017; Vruwink et al., 2022). For group members, it can negatively impact the perceived safety, thereby influencing the overall group climate. The group climate, in turn, can affect the likelihood of further aggressive incidents (van den Tillaart et al., 2018). For support staff, clients' aggression can create an unsafe working environment, causing feelings of anger, fear, and physical injury (Frowijn et al., 2024; Verstegen et al., 2024). Staff exposed to frequent or severe aggression report increased sick leave and burnout symptoms compared to staff members who do not experience aggression (De Looff et al., 2018; Hensel et al., 2014; Kind et al., 2018). Therefore, aggression not only contributes to staff shortage and higher workloads, but also leads to overtime and reduced continuity in care. Ultimately, this may undermine a facility's ability to consistently deliver safe, high-quality care (Brown & Kalaitzidis, 2013; Fanneran et al., 2015; Luther et al., 2017; Robson & Attard, 2019).

Aggression results from a complex interplay of biological, psychological and social factors, the so called biopsychosocial model (Steinert & Whittington, 2013). Physiology as a component of the biopsychosocial model has gained interest in relation to aggression (De Looff, Cornet, et al., 2022b; Lorber, 2004; Popma & Raine, 2006). Physiology was difficult to study in real-life settings until recent technological advancements (Johnson & Picard, 2020). Studies using these technologies show that increased physiological arousal is linked to an increased likelihood of aggression, irrespective of its cause (Warburton & Anderson, 2018). Physiological arousal can be measured with parameters reflecting autonomic nervous system (ANS) functioning, such as heart rate (HR), and electrodermal activity (EDA). EDA includes a tonic component, the skin conductance level (SCL), which varies gradually, and a phasic component, the skin conductance response (SCR), which reflects rapid responses to stimuli (Boucsein, 2012). A meta-analysis on aggression and physiology even reported that some physiological indicators had larger effect sizes than several psychosocial risk factors (De Looff, Cornet, et al., 2022b). This might suggest that combining physiological measures with psychosocial information may provide deeper insights into the assessment and treatment of aggression (De Looff et al., 2022). For example, Goodwin et al. (2019) suggested that aggression toward others could be predicted one minute in advance using three minutes of prior biosensor data (HRV and EDA). Similarly, De Looff et al. (2019) found significant increases in HR, SCL and SCR in clients with MID-BIF 20 minutes on average before aggression occurred.

While few studies have investigated physiological arousal preceding aggression, even less is known about physiological arousal *after* aggression. Understanding post-incident physiological arousal is important due to the phenomenon of excitation transfer (Zillmann, 1971). This phenomenon occurs when increased physiological arousal from a previous incident has decreased but not fully dissipated. Often, physiological arousal takes longer to dissipate than individuals realize (Groves & Anderson, 2016). For example, Nakajima et al. (2017) found that college students had elevated HR after exercising vigorously for one minute, despite believing their physiological arousal had returned to baseline. If provoked before the initial physiological arousal has fully dissipated, a person may misattribute the lingering physiological arousal to the new provocation. This heightened state of physiological arousal may predispose a person to respond with aggression in response to relatively minor provocations.

Most studies that investigated physiological arousal *after* aggression used self-reporting rather than physiological measures. Ferguson et al. (2018), for example, found reductions in *subjective* arousal (self-reported stress) after participants engaged in aggression. Also, Daffern and Howells (2009) found that aggression in clients in a secure setting often had the function of tension reduction. Verona and Sullivan (2008) used physiological measures and observed reductions in HR after aggressive, but not non-aggressive responses to interpersonal stressors. However, this study was limited to a laboratory setting and involved only mild forms of aggression. Based on aforementioned studies we hypothesize that some form of downregulation into homeostasis on a physiological level should be apparent after aggression. To our knowledge, no studies have been conducted that measure physiological parameters after aggression in naturalistic settings. Knowledge about physiological arousal *after* aggression as it occurs in naturalistic settings may inform the development of strategies for understanding, preventing, and managing recurrent aggression.

Advancements in wearable technology have enabled continuous, non-invasive measurement of ANS activity in real-life settings (Johnson & Picard, 2020). These technologies make it possible to study the relationship between aggression and physiological parameters not only in lab settings but also in real-life contexts. The present study aimed to investigate the level and course of physiological arousal *after* aggression in children, adolescents, and adults using repeated physiological measurements in daily life. It was hypothesized that HR, SCL and

SCR would decrease in the 30 minutes after aggressive incidents. Additionally, the study explored whether distinct subgroups with different post-incident physiological arousal trajectories could be distinguished, and whether these patterns relate to specific subtypes of aggression (verbal aggression, aggression towards objects, auto-aggression and physical aggression towards others).

METHOD

Setting

For this study, data from three studies on physiological arousal and aggression were combined for secondary analyses (De Looff et al., 2019; Hagoort et al., 2025; Nijhof et al., 2023). Hagoort et al. (2025) studied the feasibility of using wearables in clinical care, while Nijhof et al. (2023) and De Looff et al. (2019) studied physiological changes *preceding* aggression in residential youth care and forensic psychiatry, respectively. Hagoort et al. (2025) collected data at the University Medical Centre Utrecht (UMCU) child psychiatric inpatient treatment facility in the Netherlands. This centre offers treatment to children aged 6 to 14 years with severe developmental problems and includes an inpatient ward for weekday stays, including evenings and nights, and a daycare unit operating from 8:00 a.m. to 3:00 p.m. on weekdays. Most children attend a nearby special education school for part of the day. Data from the study by Nijhof et al. (2023) were collected from three Dutch residential facilities (Pluryn, Juzt, and Levvel) specialised in treating youth aged 10 to 18 years with externalizing behavioural problems. Data from the study by De Looff et al. (2019) were collected from four (forensic) inpatient treatment facilities in the Netherlands. These facilities provide treatment to adults with MID-BIF (IQ 50 - 85) and comorbid severe psychiatric and/or behavioural disorders.

Participants

All participants (N = 60), and their legal guardians when applicable, provided written informed consent before study participation for the use of their anonymous data in research. Some participants dropped out and were therefore excluded from this study, because they refused to wear the watch measuring physiological data, or withdrew consent (De Borg N = 4; Pluryn, Juzt and Levvel N = 6; UMC N = 0). Furthermore, participants who did not engage in aggressive incidents during the study period (as scored with the Modified Overt Aggression Scale plus [MOAS+], see Procedure) were also excluded.

The initial sample in the study by Hagoort et al. (2025) included 30 children, without applying specific inclusion or exclusion criteria for participation. From this sample, the current study included[8] 14 boys and two girls with a mean age

8 Based on the inclusion criteria described above in combination with exclusion based on artefact correction, described in *Statistical analysis*.

of 8.94 years (*SD* = 1.81, range: 6-12). The most prevalent primary DSM-5 (American Psychiatric Association, 2013) diagnoses were attention deficit hyperactivity disorder (ADHD; *n* = 10), autism spectrum disorder (ASD; *n* = 7), and oppositional defiant disorder (ODD; *n* = 3). Participants had an IQ above 80.

Nijhof et al. (2023) initially selected 25 adolescents aged 12-18 years with frequent rates of reactive aggression (i.e., arising from frustration or emotions) and an IQ above 70. All participants had an expected residential stay longer than 3 months, access to a mobile phone, and staff deemed the risk of damaging or selling the Empatica (see below) as low. Their mean age was 15.73 years (*SD* = 1.66). Common primary DSM-5 diagnoses were posttraumatic stress disorder (PTSD; *n* = 4), reactive attachment disorder (*n* = 4), ADHD (*n* = 3) and ASD (*n* = 3). For one participant, no psychiatric evaluation was performed. From this sample, the current study included[8] 17 adolescents, including five males and 12 females.

De Looff et al. (2019) initially selected 104 adults with MID-BIF without inclusion or exclusion criteria for participation. Support staff reported one or more aggressive incidents for 36 participants. Of the participants that showed aggression during the study period, the current study included[8] 27 adults including 17 males and 10 females, with a mean age of 29.00 years (*SD* = 7.66; range: 18-46 years). DSM-5 diagnoses included: substance use disorders (*n* =17), PTSD (*n* = 2), schizophrenia spectrum and other psychotic disorders (*n* = 6), ASD (*n* = 6), unspecified personality disorder (*n* = 4), ADHD (*n* = 3), disruptive impulse control and conduct disorders (*n* = 2), and depressive disorder (*n* = 1).

The combined sample of the current study included 60 participants: 16 children, 17 adolescents and 27 adults.

Measures and materials

Heart rate and electrodermal activity

HR and EDA were measured with the Empatica E4 wristband (Garbarino et al., 2014). The Empatica E4 measures EDA, peripheral skin temperature, blood volume pulse (BVP) from which HR and inter-beat-interval are derived, and motion with a 3-axis accelerometer. Participants of the studies by Hagoort et al. (2025) and De Looff et al. (2019) wore the Empatica E4 on their non-dominant hand and participants of the study by Nijhof et al. (2023) wore the Empatica E4 on their preferred hand. HR was expressed in beats per minute (bpm), SCL was measured in Micro siemens (μS), and SCR was expressed in Peaks Per Minute (PPM). Both SCL and SCR are parameters of EDA. HR and EDA measurements on the wrists are prone to artifacts, for example because of movement or pressure on the electrodes and devices (Boucsein, 2012; De Looff et al., 2022). To address this issue, the current study used several packages that are available to detect artifacts, pre-process the signals, and extract relevant features (Kleckner et al., 2017; Martínez et al., 2017;

Wickham et al., 2019). The physiological signals were automatically checked by automated recognition software (Kleckner et al., 2017; Taylor et al., 2015) with the use of the E4 dashboard, the new version of the Wearalize dashboard, and the wearables package in R which incorporates several algorithms for pre-preprocessing and batch analyses (De Looff, Duursma, et al., 2022).

Aggressive incidents

For the data collected for the studies by Nijhof et al. (2023) and De Looff et al. (2019), staff members observed aggression and assessed the type of aggression using the MOAS+ (Crocker et al., 2006; Drieschner et al., 2013; Oliver et al., 2007). The MOAS+ categorizes aggression into verbal aggression, physical aggression, aggression against objects, sexual aggression, and auto-aggression (self-harm). The MOAS+ demonstrated good reliability, with Cohen's kappa values ranging from .65 to .90 (Oliver et al., 2007).

Hagoort et al. (2025) collected behavioural data in two complementary ways. First, staff members recorded child observations for diagnostic purposes in the electronic client file, according to the facility's standard procedures. Second, all agitated and potentially aggressive behaviours were recorded by independent observers (research interns) and/or staff members or teachers. These recordings included the timeframe and a description of the behaviour. For the current study, two researchers (MvS and IN) independently coded these incident reports using the MOAS+ to classify the behaviour as aggressive or non-aggressive. Incidents that did not fall into any of the MOAS+ categories or lacked sufficient observations were excluded. Additionally, the type of aggression was independently coded by both researchers. If multiple types of aggression occurred during a single incident, the most severe form was selected for statistical analyses (see Statistical analysis). According to MOAS+ scoring, the severity order of aggression types was: physical aggression, auto-aggression, aggression against property and verbal aggression. Cohen's kappa indicated substantial inter-rater agreement[9] (κ = .72). Discrepancies were resolved through discussion between the two researchers who coded the incident reports to reach consensus.

Procedure

All three studies that provided the data for the current study were conducted in accordance with the Declaration of Helsinki's ethical principles for research involving human participants.

The ethics approval of the study by Hagoort et al. (2025) from which the UMCU data was obtained was reviewed by the Medical Research Ethics Committee at UMCU, the Netherlands (18-886). The study was evaluated as 'no research related

9 Based on the most severe subtype of aggression if multiple types of aggression were observed during one incident.

medical ethical approval needed' because it imposed minimal additional burden on the client without interrupting or changing their regular clinical program. Participants received an information letter available in three versions: one for participants under the age of 12, one for participants aged 12 and older, and one for parents. Parents were contacted by phone to provide information about the study. Following this, detailed study information was sent to them. Subsequently, parents and participants were given time to consider participation. Children received additional explanations to ensure clarity and address any questions. If participation was agreed upon, the parents and/or the child signed the informed consent form. For participants aged 12 and older, both the child and the parents signed the consent form. Data were collected at UMCU between April 2019 and January 2021. Participants wore the Empatica E4 five days during one week. Children admitted for clinical treatment wore the Empatica E4 from 8 a.m. to 8 p.m., while children in day treatment wore the Empatica E4 from 8 p.m. to 3 p.m.

The study protocol by Nijhof et al. (2023), from which the data of Pluryn, Juzt, and Levvel was obtained was reviewed by the Medical Ethical Committee of Arnhem-Nijmegen and was evaluated as compliant with the Dutch Law on Medical Research in Humans, requiring no further ethical approval (ref.no. NL63138.091.17). Clinicians identified potential participants. These adolescents and their legal guardians were informed through an information letter. Nijhof et al. (2023) collected data between April 2018 and October 2018. Participants wore the Empatica E4 for a total of three weeks during the daytime. To register the timing of incidents, staff members could use a watch to timestamp the exact moment they observed aggression.

Ethical approval for the study of De Looff et al. (2019), from which the data of De Borg was obtained, was granted by the Ethics Committee of the Faculty of Social Sciences of Radboud University Nijmegen (ECSW2015-1901-282). The participants were invited and informed about the study through email, flyers and posters. De Looff and colleagues (2019) collected data between May 2015 and August 2017. Data were collected throughout five consecutive days during the daytime. To register the timing of incidents, staff could use a watch to timestamp the exact moment they observed aggression.

Statistical analysis

For this study, physiological data from the 30 minutes *after* an aggression incident were extracted and divided into six 5-minute epochs. For each epoch, the average HR, SCL, SCR, movement and temperature were calculated. Similarly to the study by De Looff et al. (2019), only epochs with less than 25% artefacts in the EDA signal were included. Incidents occurring within 30 minutes of another aggressive incident by the same client were considered overlapping and not treated as independent. In such cases, the last recorded registration of aggression within the timeframe was used for analysis.

To determine whether physiological arousal significantly decreased within the 30 minutes after aggression, separate multilevel analyses were conducted for each physiological parameter (HR and EDA-parameters SCL and SCR). The data structure included three levels: six 5-minute epochs (Level 1) nested within incidents (Level 2), nested within individuals (Level 3). De Looff et al. (2019) and Nijhof et al. (2023), who analysed the 30 minutes preceding aggression, found that modelling the third level did not substantially improve explained variance or model fit. To assess this for the current study, null-models were first fitted for both the two-level model (with incidents as the highest level) and the three-level model (with epochs nested in incidents nested within individuals). Deviance tests determined whether adding the third individual level improved model fit. The three-level model significantly outperformed the two-level model for all physiological parameters, as indicated by lower AIC values (HR: ΔAIC = 75.43, SCL: ΔAIC = 12.85, SCR: ΔAIC = 12.12) and significantly lower model deviance (HR: $\chi^2(1) = 77.77$, $p < .001$, SCL: $\chi^2(1) = 14.85$, $p < .001$, SCR: $\chi^2(1) = 14.12$, $p < .001$). Therefore, the three-level structure was used in subsequent analyses.

For each outcome measure, the same model building strategy was used. First, the fixed part of the model was constructed, followed by the random part. The first model included only a linear effect of time. The starting time of the aggressive incident that had been registered was coded as 0. Second, a quadratic effect of time was added. Third, time varying covariates for movement and temperature were then added. These variables were grand mean-centred to ensure interpretability by centering around the overall sample mean while preserving both within- and between- group variation. Fourth, time-invariant covariates facility (reference category: De Borg) and type of aggression (reference category: physical aggression) were added. Aggressive incidents were grouped into two categories: verbal and physical aggression (the latter category consisted of aggression towards objects, auto-aggression and physical aggression towards others). This categorization was chosen due to the relatively small number of incidents involving auto-aggression and physical aggression towards others, which could reduce statistical power when analysing all four types of aggression separately in the multilevel analyses. Finally, a random effect of time was added. An autoregressive covariance structure (AR1) was tested, as it was expected that time points closer together would be more correlated. Model improvement was assessed using the Akaike Information Criteria (AIC) and deviance tests. When the inclusion of additional variables resulted in lower AIC and significantly reduced deviance, the more complex model was preferred (Heck & Thomas, 2020). In the case of improved model fit, individual regression coefficients were examined. Significant predictors were retained, and non-significant predictors were excluded in subsequent models. The final model is presented in the results section, while detailed information of each model step is provided in Appendix 4A.

Following the multilevel modelling, we exploratively assessed whether unobserved (latent) classes could be identified which follow similar growth patterns to better capture variability in physiological arousal patterns after aggressive incidents. To examine whether there are distinct growth curves of the outcome measures over time, latent class growth curve analyses (LCGA) were performed for HR, SCL and SCR across six 5-minute epochs after the aggressive incidents. To identify the optimal number of classes, models with one to eight classes were estimated to capture the variability among incidents. The best-fitting class solution was determined by the lowest values of the Bayesian Information Criterion (BIC), AIC and AIC3 (Frankfurt et al., 2016). When BIC, AIC, and AIC3 suggested different number of classes, a bootstrapped likelihood ratio test (BLRT) was performed between the indicated models to determine the best fitting class solution, as BLRT is considered more reliable for identifying the correct number of classes (Nylund et al., 2007). The final class solution is reported in the results section, with detailed model fit statistics for all tested models (one to eight classes) provided in Appendix 4A. Labels were assigned to the classes based on variation in intercepts and slopes. Intercepts were categorized using tertiles of the variation of HR, SCL, and SCR at the moment of the incident (T0), categorized as low, moderate and high. Slope labels reflected whether trajectories showed a significant increase, decrease, or non-significant change over time. After identifying the best-fitting class solution, we examined whether the type of aggression and incident setting predicted latent class membership using the bias-adjusted three-step approach (Vermunt, 2010). Results regarding type of aggression as a predictor of latent class membership are reported in the results section. Information on the organisation as a predictor of latent classes is provided in Appendix 4A.

RESULTS

Descriptives

Initially, 222 aggressive incidents were included, consisting of 1250 5-minute epochs. Due to an artefact percentage exceeding 25 percent, 454 epochs were excluded. Also, four incidents were excluded due to missing data on the type of aggression. This resulted in a total of 783 five-minute epochs from 165 incidents recorded across 60 individuals included in the data analysis. Among these individuals, 27 (45.0%) were included from De Borg, 17 (28.3%) from Pluryn, Juzt, Levvel and 16 (26.7%) from UMCU. The number of recorded incidents per person ranged from 1 to 9, with a median of 2 (interquartile range [IQR] = 3.00). Of the incidents, 61.2% was an incident of verbal aggression, and 38.8% of physical aggression (31.2% physical aggression towards others, 54.1% physical aggression towards objects, 15.6% physical aggression towards oneself/auto-aggression).

Physiological parameters after aggression

Heart rate (HR)

Table 1 shows the results of the best-fitting multilevel model for HR. The findings showed that time has a significant negative effect on HR (b = -0.60, p = .008), indicating that HR decreased, on average, over the 30 minutes after an aggressive incident. Movement had a significant positive effect on HR (b = 71.46, p < .001). Figure 1 shows the regression line from the best fitting model (Model 7) and the observed means across six 5-minute epochs after the aggressive incidents.

Table 1

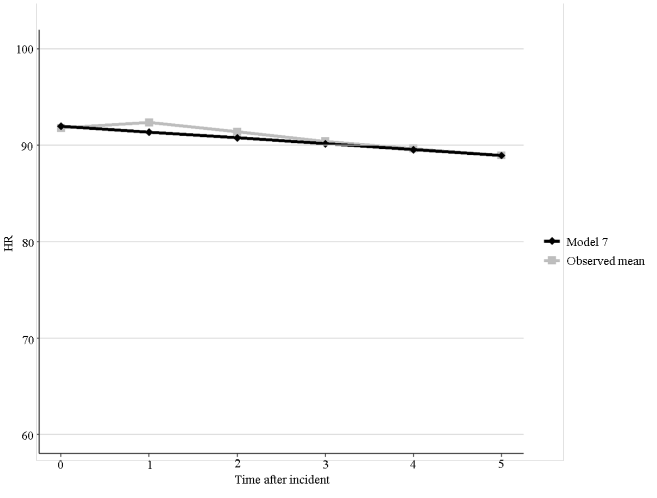
Summary of best-fitting multilevel models for HR, SCL and SCR

Effect	HR	SCL	SCR
	Fixed effects		
Intercept	91.82** (1.14)	1.73** (0.27)	3.86** (0.30)
Time	-0.61** (0.23)	-0.10** (0.03)	-0.11* (0.05)
Movement	71.17** (11.35)		10.98** (2.00)
	Random effects		
	Unstructured		
Level 1 – σ^2_e	86.39**	[a]	[a]
Level 2 – σ^2_1	20.43	5.61**	4.43**
Level 2 – σ_{01}	0.36	-0.71**	-0.484**
Level 2 – σ^2_1	0.48	0.13**	0.19**
Level 3 – σ^2_1	31.25**	1.74**	2.23*

*Notes. Standard errors are in parentheses. *p < .05. **p < .01.* [a] *Variance differs across timepoints*

Figure 1

Observed mean HR and regression line of best fitting model (Model 7) across six 5-minute epochs following aggressive incidents

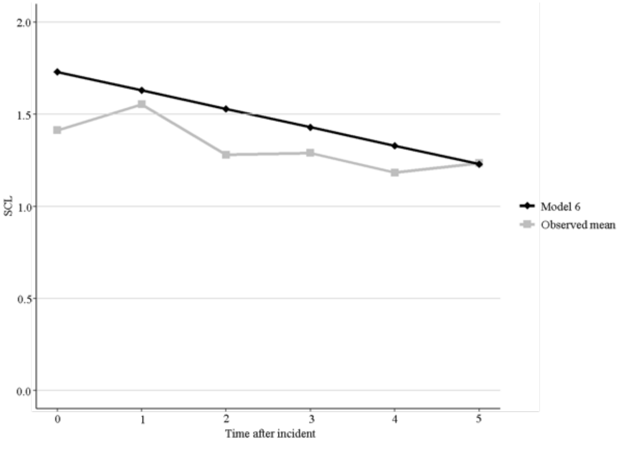


Skin conductance level (SCL)

Table 1 shows the results of the best-fitting multilevel model for SCL. The final model showed that time had a significant negative effect on SCL (b = -0.10, p = .002), indicating that SCL decreased over the 30 minutes after an aggressive incident. Figure 2 shows the regression line from the best-fitting model (Model 6) and the observed means across six 5-minute epochs after the aggressive incidents.

Figure 2

Observed mean SCL and regression line of best fitting model (Model 6) across six 5-minute epochs following aggressive incidents

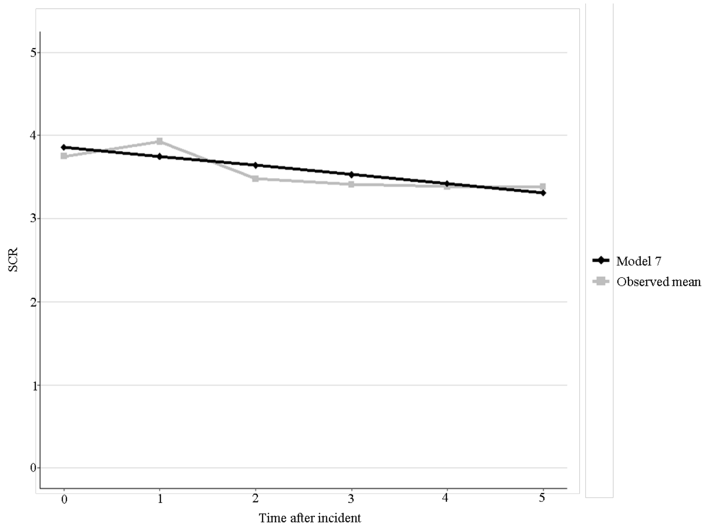


Skin conductance response (SCR)

Table 1 shows the results of the best-fitting multilevel model for SCR. The final model showed a significant negative effect of time on SCR (b = -0.11, p = .002), suggesting that SCR decreases, on average, during the 30 minutes after an aggressive incident. The covariate movement was a significant positive predictor of SCR (b = 10.98, p < .001). Figure 3 shows the regression line from the best fitting model (Model 7) and the observed means across six 5-minute epochs after the aggressive incidents.

Figure 3

The observed mean SCR and regression line of best fitting model (Model 7) across six 5-minute epochs following aggressive incidents



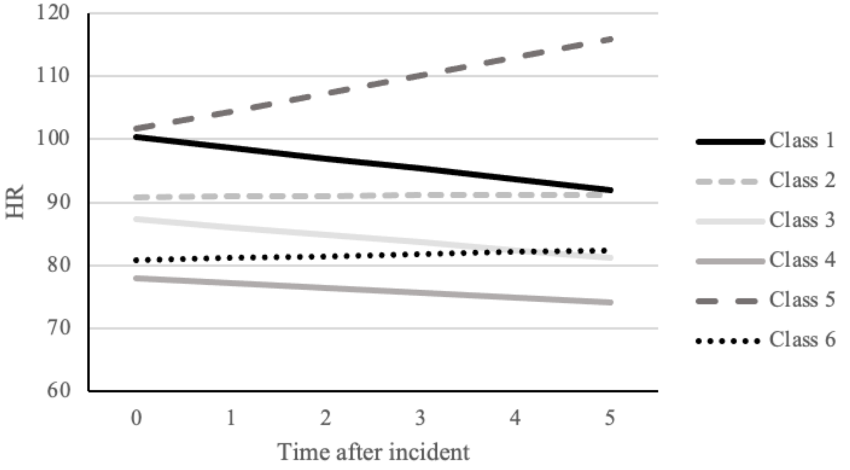
Distinct trajectories in physiological parameters over time

Heart rate (HR)

The LCGA results identifying distinct HR trajectories over time after aggressive incidents showed that a six-class model fitted the data the best. The estimated mean HR trajectories of the six-class model are shown in Figure 4. Each class was labelled based on variations in HR intercepts and slopes. The largest class, consisted of 31.22% of incidents, was characterized by a high HR at the time of the incident (M = 100.39, tertile 3) and a significant decrease over time (b = -1.70, 95% CI [-3.03; -0.38]), labelled as 'High start value with decreasing HR'. The second class included 29.09% of incidents, characterized by a moderate HR at the time of the incident (M = 90.84, tertile 2) and no significant change over time (b = 0.07, 95% CI [-0.77; 0.92]), labelled as 'Moderate start value with stable HR'. The third class consisted of 20.67% of incidents and was characterized by a moderate HR at the time of the incident (*M* = 87.30, tertile 2) and a significant decrease in HR over time (*b* = -1.23, 95% CI [-2.19; -0.27]), labelled as 'Moderate start value with decreasing HR'. The fourth class consisted of 10.21% of incidents, characterized by a low HR at the time of the incident (*M* = 77.97, tertile 1) and significant decrease in HR over time (*b* = -0.78, 95% CI [-1.60; -0.04]), labelled as 'Low start value with decreasing HR'. The fifth class comprised 6.78% of incidents, characterized by a high HR at the time of the incident (*M* = 101.60, tertile 3) and a significant increase over time (*b* = 2.86, 95% CI [0.11; 5.61]), labelled as 'High start value with increasing HR'. The sixth and smallest class consisted of 2.02% of incidents, was characterized by a low HR at the time of the incident (*M* = 80.89, tertile 1) and no significant change over time (*b* = 0.29, 95% CI [-0.26; 0.85]), labelled as 'Low start value with stable HR'. The type of aggression significantly predicted the latent class an incident belonged to (Wald = 586.91, $p < .001$). Verbal aggression was associated with the 'Low start value with stable HR' trajectory, while aggression against objects was associated with 'High start value with increasing HR' and 'Low start value with decreasing HR' trajectories. Auto-aggression was more common in 'High start value with decreasing HR', 'Moderate start value with stable HR' and 'Low start value with stable HR' trajectories. Physical aggression was associated to 'Moderate start value with decreasing HR', 'Low start value with decreasing HR', and 'Moderate start value with stable HR' trajectories.

Figure 4

Estimated mean trajectories of HR following an aggressive incident for each class



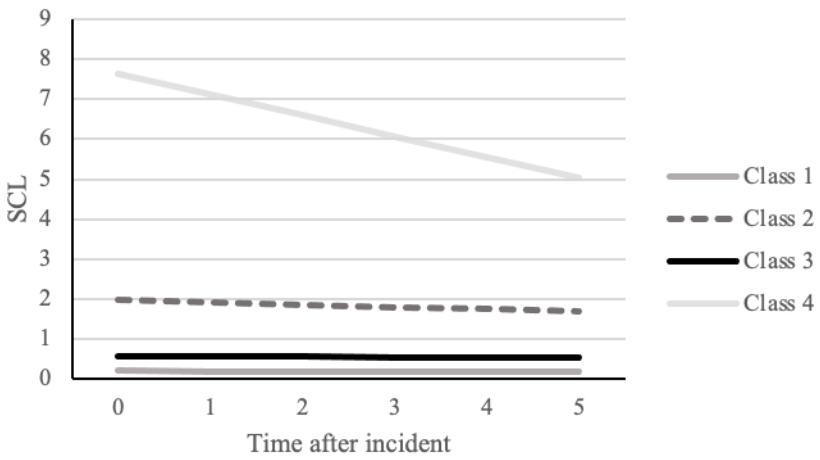
Skin conductance level (SCL)

Four distinct SCL trajectory classes were identified. The estimated mean trajectories of SCL for each class are shown in Figure 5. The largest class consisted of 37.48% of incidents, and was characterized by a low SCL at the time of the incident (*M* = 0.21, tertile 1) and no significant change over time (*b* = -0.01, 95% CI [-0.01; 0.00]). This class was labelled 'Low start value with stable SCL'. The second class included 29.12% of incidents and was characterized by a high SCL at the time of the incident (*M* = 1.97, tertile 3) and no significant change over time (*b* = -0.06, 95% CI [-0.14; 0.03]). This class was labelled 'High start value with stable SCL'. The third class, consisting of 22.89% of incidents, was characterized by a moderate SCL at the time of the incident (*M* = 0.58, tertile 2) and no significant change over time (*b* = -0.01, 95% CI [-0.04; 0.02]), and was labelled as 'Moderate start value with stable SCL'. The fourth and smallest class, comprising 10.51% of incidents, was characterized by a high SCL at the time of the incident (*M* = 7.63, tertile 3) and a significant decrease in SCL over time (*b* = -0.52, 95% CI [-0.94; -0.09]) and was labelled as 'High start value with decreasing SCL'.

Type of aggression was found to significantly predict latent class membership (Wald = 222.42, $p < .001$). Verbal aggression, aggression against objects, and physical aggression were associated with a higher likelihood of following the 'High start value with decreasing SCL' trajectory. Auto-aggression was more likely to align with the 'High start value with stable SCL' and 'Moderate start value with stable SCL' trajectories.

Figure 5

Estimated mean trajectories of SCL following an aggressive incident for each class



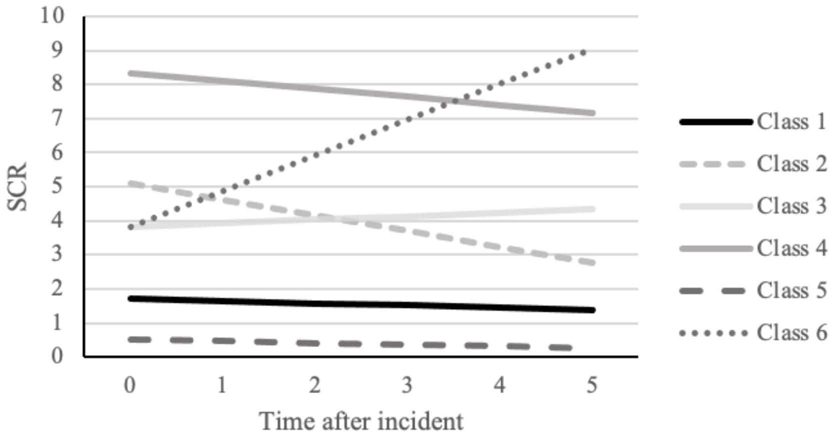
Skin conductance response (SCR)

Six classes of incidents with distinct trajectories of SCR over time were identified. The estimated mean trajectories of SCR after the incidents for each class are shown in Figure 6. The largest class, representing 24.16% of incidents, was characterized by a low SCR at the time of the incident (M = 1.73, tertile 1) and no significant change over time (b = -0.07, 95% CI [-0.16; 0.01]). This class was labelled 'Low start value with stable SCR'. The second class consisted of 24.03% of incidents and was characterized by a high SCR at the time of the incident (M = 5.09, tertile 3) and a significant decrease over time (b = -0.46, 95% CI [-0.70; -0.23]), and was labelled as 'High start value with decreasing SCR'. The third class, comprising 22.21% of incidents, was characterized by a moderate SCR at the time of the incident (M = 3.83, tertile 2) and no significant change over time (b = 0.10, 95% CI [-0.02; 0.22]), and was labelled as 'Moderate start value with stable SCR'. The fourth class consisted of 13.66% of the incidents and was characterized by a high SCR at the time of the incident (M = 8.32, tertile 3) and no significant change in SCR over time (b = -0.23, 95% CI [-0.47; 0.02]). This class was labelled 'High start value with stable SCR'. The fifth class included 11.02% of incidents and was characterized by a low SCR at the time of the incident (M = 0.53, tertile 1) and a significant decrease over time (b = -0.05, 95% CI [-0.11; -0.01]). This class was labelled 'Low start value with decreasing SCR'. The sixth and smallest class, comprising 4.93% of incidents, was characterized by a moderate SCR at the time of the incident (M = 3.82, tertile 2) and a significant increase in SCR over time (b = 1.05, 95% CI [0.67; 1.42]), and was labelled as 'Moderate start value with increasing SCR'.

The type of aggression significantly predicted latent class membership (Wald = 248.31, $p < .001$). Verbal aggression, aggression against objects, and physical aggression were linked to the 'Low start value with decreasing SCR' trajectory, whereas auto-aggression was associated with the 'Low start value with stable SCR', 'Moderate start value with stable SCR' and 'Moderate start value with increasing SCR' trajectories.

Figure 6

Estimated mean trajectories of SCR following an aggressive incident for each class



DISCUSSION

Building on prior laboratory studies examining physiological responses after aggression, this study investigated changes in HR, SCL and SCR during the 30 minutes after aggressive incidents in a naturalistic setting. On average, all three parameters declined over time. However, distinct patterns of change emerged. LCGA revealed subgroups with different trajectories: some showed the expected decreases, while others had stable or increasing patterns. This highlights the heterogeneity of post-aggression physiological responses. Although the type of aggression did not predict the overall decrease across parameters, it was significantly associated with the trajectory subgroup the incident belonged to, suggesting behaviour type influences the course of physiological recovery.

Main Results

The observed decrease in physiological arousal after aggressive incidents aligns with earlier findings, such as those by Verona and Sullivan (2008) and Hokanson and Burgess (1962), who reported HR reductions after aggression in laboratory settings. Our study expands on this by including additional physiological parameters – SCL and SCR – which also showed decreases after aggression. Importantly, by examining aggression in a naturalistic rather than a controlled environment, our findings provide greater ecological validity. It has been suggested that the overall decrease in physiological arousal may reflect a regulatory function of aggression in response to heightened emotional states (Roberton et al., 2012). This interpretation fits with the frustration-aggression hypothesis, which proposes that negative emotions following aversive events, such as frustration, lead to physiological over-arousal that motivates aggression (Berkowitz, 1993). In this framework, overall, aggression may serve as a mechanism through which individuals reduce arousal.

Heterogeneity in physiological patterns

Despite the overall decrease in HR, SCL and SCR over time, considerable heterogeneity was observed in the trajectories of participants' physiological responses after the aggressive incidents. The majority (62%) of incidents fell into a subgroup where participants' HR decreased within 30 minutes after the incident. In contrast, only about one-third (35%) of incidents showed a decrease in SCR, and a mere 11% showed decreasing SCL trajectories. These differences between physiological indices may be partly explained by the different ANS mechanisms underlying each measure (Behnke et al., 2022). HR is influenced by both the sympathetic and parasympathetic branches of the ANS (Jarczok et al., 2013), while SCL and SCR are primarily regulated by the sympathetic nervous system (SNS) (Boucsein, 2012). SNS activation drives the body's "fight or flight" response, while parasympathetic nervous system (PNS) activity promotes a "rest and digest" state (Berntson et al., 2008; Porges, 2007). After a stressor ends, PNS dominance typically resumes to restore homeostasis (Kahle et al., 2016). Because SCR and SCL are not affected by parasympathetic activation, their return to baseline may be slower than HR. Furthermore, while both SCL and SCR reflect SNS activity, they capture different aspects of arousal (Zhang et al., 2014). SCL is thought to represent a more general level of activation, whereas SCR is more closely associated with immediate emotional reactions to specific stimuli or events (Boucsein, 2012; Christopoulos et al., 2019).

In addition to decreasing trajectories, stable (31% HR, 60% SCR, 89% SCL) and even increasing trajectories (7% HR, 11% SCR) were observed. This suggests that not all aggressive incidents result in decreased physiological arousal and suggest that various moderating factors may influence these outcomes. Measures of ANS activity exhibit circadian rhythms, with fluctuations throughout the day (Jarczok et al., 2019), which may partly explain the variability in physiological patterns. Individual differences in stress reactivity and recovery also likely contribute to the variability in trajectories. According to Wilder's Law of Initial Values, the physiological response is partly dependent on the baseline level prior to the event (Everly & Lating Jr, 2019; Wilder, 1950). Additionally, cognitive expectancies of the aggressor may modulate the physiological response. Bushman (2002) found that aggression functioned as a form of emotion regulation only for individuals who believed that aggression would reduce their negative emotions. This suggests that biological processes are influenced by cognitive factors, such as expectancies, which is consistent with the Cognitive Activation Theory of Stress (CATS;

Ursin & Erikson, 2010). CATS states that individuals develop expectancies based on prior experiences with stress, which can either attenuate or amplify the physiological stress response. This aligns with the findings that individuals with positive expectancies about coping showed lower cortisol responses to stress compared to individuals with negative expectancies (Pulopulos et al., 2020).

Additionally, characteristics of the incident could play a role in the physiological response after aggression. Aggression against someone unrelated to the cause of negative emotions may not produce the same physiological effects. Verona and Sullivan (2008) found that a decrease in physiological arousal after aggression did not occur when a target of aggression differed from the source of frustration. This suggests that venting through actions not directed at the source of frustration may be ineffective in reducing physiological arousal. However, a meta-analysis suggests that anger from provocation can be reduced through non-targeted activities, such as mindfulness and meditation (Kjærvik & Bushman, 2024). Similarly, Bushman (2002) found that venting through aggressive activities, like striking a punching bag, actually increased anger and aggression by keeping aggressive thoughts active in memory. This aligns with findings that rumination on angry thoughts prolongs and intensifies anger (Wilkowski & Robinson, 2010). Supporting this, studies found that engaging in activities that induce emotions incompatible with anger, such as watching something funny, reduces the likelihood of aggression, both in naturalistic and laboratory settings (Baron, 1983; Lutz & Krahé, 2018). These findings indicate that displaced aggression may not reduce arousal effectively and could reinforce aggressive cognitions. Conversely, changing anger into non-aggressive and emotionally incompatible activities may reduce arousal in a healthier, non-destructive manner.

Type of aggression

Our finding that verbal and physical aggression did not differ in the average decrease of physiological parameters is consistent with the findings of Hokanson and Burgess (1962), who found a decrease in HR following both verbal and physical laboratory aggression, with no significant difference between the two types of aggression. This may suggest a more general recovery process after aggression, independent of whether the aggression is verbal or physical. Both verbal and physical aggression are typically preceded by heightened physiological arousal (De Looff et al., 2019), which could be followed by a compensatory decrease as the body attempts to restore homeostasis.

Although the type of aggression did not predict the overall change in physiological parameters over time, it was associated with the specific post-aggression trajectories. Each of the four types of aggression was associated with multiple physiological response patterns. Verbal aggression was associated with stable (HR) and decreasing trajectories (SCL and SCR). Aggression against objects showed both increasing (HR) and decreasing (HR, SCL, and SCR) trajectories. Auto-aggression was linked to decreasing (HR), stable (HR, SCL, and SCR) and increasing (SCR) trajectories. Physical aggression was associated with decreasing (HR, SCL, and SCR) and stable (HR) trajectories. The fact that each type of aggression was associated with multiple trajectories may reflect the diverse functions and motivations behind similar behaviours. For instance, auto-aggression may serve as a form of emotion regulation (Hooley & Franklin, 2018), but also as a way to express distress or seek help (Klonsky et al., 2013). The physiological response may vary based on the specific behavioural function underlying the aggressive act. For example, Lydon et al. (2013) suggested that socially motivated auto-aggression (e.g., to gain attention or avoid a social task demand), may not elicit physiological arousal, whereas behaviour aimed at escaping aversive states might be accompanied by physiological change. Furthermore, because auto-aggression also involves physical pain, and pain is known to elevate physiological arousal (van der Venne et al., 2023), this could explain its association with increasing trajectories.

Strengths and limitations

To our knowledge, this is the first study to investigate physiological responses (in terms of HR, SCL and SCR) after aggression in a naturalistic setting. In addition to its high ecological validity, another strength is the multicentre approach, with data collected from eight different facilities. However, several limitations must be acknowledged. First, despite training and clear instructions, and the use of independent observers or watches to timestamp the moment of incidents, there were indications that incidents were not always registered accurately. For example, the number of incidents logged at rounded times was higher than would be expected by chance. It is likely that during aggressive incidents, staff members were focused on de-escalating and ensuring safety, and only recorded the time afterwards, based on estimation, once the immediate threat had passed and they had the capacity to reflect on the timing of the event. The timing of incidents may have been recorded somewhat earlier or later than they actually occurred. This could have had an effect on the temporal alignment between the recorded incident and the physiological data, limiting our ability to precisely capture physiological changes after the incidents.

Second, this study lacked information on how other clients or staff members responded to the aggression. Such incidents disrupt the ward environment, potentially increasing stress among clients and triggering countertransference reactions in staff members, which may escalate the situation further. Staff interventions may have influenced physiological arousal after the incident. De-escalation attempts could have been successful, or may have led to more stress as a result of coercive measures or seclusion, as these actions are often experienced as stressful (Chieze et al., 2019; Hallett et al.). We had no information on what occurred after the aggression, leaving crucial contextual factors unexplored. Moreover, we did not have information on what specifically provoked the aggression. Aggression often occurs in an interpersonal context (Richter & Whittington, 2006), and interactions with staff or other clients may contribute to or provoke the escalation of such behaviour (Haugvaldstad & Husum, 2016; Papadopoulos et al., 2012). This complicates our ability to disentangle the association between the function of the behaviour and the corresponding physiological response.

Furthermore, issues with signal quality may have influenced the results. We experienced data loss, with 34% of epochs excluded due to exceeding the 25% artefact criterium. Despite technological advances in wearable technology, physiological measurement in daily life remains vulnerable to artefacts (Hu et al., 2015). For example, low and stable EDA values may be attributed to insufficient electrode coupling of the wearable device. One latent class had a mean SCL under 0.5 micro Siemens, which may suggest that the recorded signal largely consisted of noise, a problem also reported in previous studies (Milstein & Gordon, 2020). This shows the trade-off between the feasibility of collecting ecologically valid data and ensuring signal quality. Advancing wearable technology is therefore important to improve the validity and reliability of assessment of peripheral physiological biomarker assessment in daily life .

Future directions

Although the current study identified distinct trajectories of physiological arousal after aggressive incidents, future research should investigate whether specific trajectories are associated with a lower or higher risk of subsequent aggression. According to the theory of excitation transfer (Zillmann et al., 1972) residual arousal after an aggressive incident may increase the likelihood of subsequent aggression, especially if provocation occurs before arousal has returned to baseline. This suggests that individuals who do not show a decrease in physiological arousal may be at heightened risk of subsequent aggression. Conversely, operant learning theory posits that a reduction in arousal after aggression may function as negative reinforcement, thereby increasing the chance of future aggression (Franz et al., 2019). If aggression leads to relief, individuals may associate it with emotional regulation. Supporting this, Verona and Sullivan (2008) found that individuals who experienced the greatest reductions in arousal after aggression were at increased risk for subsequent aggression. However, these findings do not necessarily imply that arousal reduction causes future aggression. Rather, it may reflect that individuals with heightened physiological reactivity are also those who tend to engage in the most aggressive behaviour. In other words, the reduction in arousal may be a consequence, not a cause. Future research could also explore whether individuals consistently show similar physiological recovery patterns across incidents, suggesting person-dependence, or whether patterns vary based on situational factors, suggesting situation-dependence. The current study found differences in physiological parameters during aggressive incidents. However, since baseline levels were not accounted for, it remains unclear whether individuals returned to their personal baseline within the 30 minutes after aggression. Although De Looff et al. (2019) found an average increase in physiological parameters in the 30 minutes *preceding* aggression, individual trajectories may vary. The extent of decrease required for homeostatic regulation likely depends on the prior arousal increase. Future research should incorporate baseline measurements, consider diurnal variation, and examine pre-incident trajectories to better understand post-aggression recovery.

Clinical implications

Identifying clinical implications from the results of this study is not straightforward. The identification of distinct trajectories of physiological arousal after aggression suggests a potential need for individualized approaches to managing aggressive incidents. For instance, individuals who continue to experience heightened arousal after aggression may benefit from targeted interventions, such as relaxation exercises or biofeedback. Biofeedback has been shown to enhance individuals' awareness of their stress responses and support more effective regulation of physiological arousal (Yu et al., 2018). On the other hand, if an individual's arousal has already decreased and they have returned to a calmer state, introducing additional interventions may be unnecessary or even counterproductive. However, intervention decisions should not rely solely on physiological data but must complement the clinical judgment of support staff members. By combining physiological measurements with clinical observations and structured risk assessments, professionals can make more informed decisions about the need of additional interventions and support. Furthermore, insights into physiological arousal patterns, both prior to and following aggression, may help to clarify the function of aggression in clients. This understanding could be used to inform personalized treatment strategies aimed at addressing the underlying causes of aggression, and preventing the use of restrictive measures.

Appendix 4A

Results multilevel models HR

The results of the multilevel analyses for HR are shown in Table 2. First, the null-model with HR as outcome was fitted. To check whether adding a third level resulted in a better model fit, we compared the deviance and AIC of the two-level model (epochs nested in incidents) to the model fit of the three-level model (epochs nested in incidents nested within individuals). The three-level model resulted in a lower AIC compared to the two-level model (ΔAIC = 75.43) and a significantly lower model deviance, $\chi 2(1) = 77.77$, *p* <.001, indicating that the three-level model provided a significantly better model fit compared to the two-level model. Therefore, the individual level is added as a third level in subsequent models. The results of the null model (Model 0) indicate that there is significant variability at all three levels of the model, within incidents (*p* <.001), between incidents (*p* <.001), and between individuals (*p* = .006). To quantify the proportion of variance that can be attributed to each level, the ICC was calculated for both the individual and incident level. The ICC shows that 22.90% of the variance in HR is attributable to variation between incidents within individuals (Level 2; ICC = 0.23) and that 22.32% of the variance in HR is attributable to differences between individuals (Level 3; ICC = 0.22). 54.78% of the variance in HR represents residual variability within incidents. Both levels contribute a substantial proportion of the variance, suggesting it was appropriate to use a multilevel structure including both levels in the model. In Model 1 time was added as a linear effect. This resulted in significantly reduced model deviance, $\chi 2(1) = 8.48$, *p* = .004. Time was negatively associated with HR (b = -0.63, *p* = .002), indicating that HR decreased over time following the aggressive incident. In Model 2, time was added as a quadratic term, which resulted in a higher -2LL compared to the previous model (Δ Deviance = -77.72), indicating that the addition of a quadratic term for time did not result in a significant improvement of the model. Therefore, the quadratic effect of time was removed in subsequent models. In Model 3, the time-varying covariates of movement and temperature were added to the model. Compared to Model 1, model deviance significantly decreased, $\chi 2(2) = 41.88$,*p* < .001. There was still a significant negative association between time and HR (b = -0.62, *p* = .001) when controlling for movement and temperature. Movement significantly predicted HR (b = 67.68, *p* <.001), indicating that more movement is associated with higher HR. Temperature did not significantly predict HR (b = -0.03, *p* = .406). Therefore, in Model 4, temperature was removed from the model, while movement was retained, which improved the -2LL and the AIC while also including less parameters (Δ Deviance = -3.77). This significantly improved model fit compared to Model 1, $\chi 2(1) = 45.65$, *p* < .001. In Model 5, the time-invarying predictors of organisation and type of aggression were added. The addition of organisation and type of aggression resulted in a significantly better model fit, $\chi2(3) = 9.88$, $p < .020$. However, there were no significant effects of organisation or type of aggression on HR. In Model 6 the interaction terms between time and organisation and time and type of aggression were added. This did not result in significantly improved model fit, $\chi2(3) = 2.36$, $p = .501$. As there were no significant main effects or interaction effects of organisation or type of aggression, these were removed from the final model. In the final model (M7) a random effect of time was added and an autoregressive covariance structure was used. This resulted in a significantly better model fit compared to Model 4, $\chi2(3) = 22.48$, $p < .001$. The fixed effects showed that time has a significant negative effect on HR (b = -0.60, $p = .008$), indicating that on average, the HR decreases in the 30 minutes following an aggressive incident. Of the covariates, movement had a significant positive effect on HR (b = 71.46, $p < .001$). Figure 1 shows the regression line of the best fitting model (Model 7) and the observed means across six 5-minute epochs following the aggressive incidents.

Table 2

Fitted multilevel models for HR across the six 5-minute epochs following aggressive incidents

Effect	Model 0	Model 1	Model 3	Model 4	Model 5	Model 6	Model 7
Fixed effects							
Intercept	90.21** (1.01)	91.85** (1.13)	91.84** (1.13)	91.84** (1.12)	92.78** (2.21)	91.54** (2.46)	91.82** (1.14)
Time		-0.63** (0.20)	-0.62** (0.19)	-0.62** (0.19)	-0.62** (0.19)	-0.14 (0.46)	-0.61** (0.23)
Movement			-67.68** (10.69)	67.70** (10.69)	67.69** (10.71)	67.42** (10.73)	71.17** (11.35)
Temperature			-0.04 (0.04)				
Organisation (ref = UMCU)							
De Borg					-1.35 (2.63)	-0.36 (2.97)	
Pluryn					-1.27 (2.78)	-0.79 (3.10)	
Verbal aggression (ref = physical aggression)					0.07 (1.37)	1.15 (1.73)	
Interactions							
De Borg*Time						-0.38 (0.52)	
Pluryn*Time						-0.19 (0.52)	
Verbal aggression*Time						-0.42 (0.41)	
Random effects							
	Variance components						Unstructured
Level 1	84.85**	83.71**	79.21**	79.21**	79.23**	79.40**	86.39**
Level 2 - intercept	33.81**	34.53**	33.22**	33.11**	33.44**	33.25**	20.43
Level 3 - intercept	32.19**	31.53**	32.22**	32.28**	34.11**	34.37**	31.25**
Level 2 - slope							0.48
Goodness of fit							
-2LL	5848.96	5840.48	5798.60	5794.83	5784.95	5782.59	5772.35
AIC	5854.96	5846.48	5804.60	5800.83	5790.95	5788.59	5784.35
$\Delta\chi^2$		8.48**	41.88**	45.65**[1]	9.88*	2.36	22.48**[2]
Δdf		1	2	1	3	3	3

Notes. Standard errors are in parentheses. * $p < .05$. ** $p < .01$. Model 2 is omitted from the table. [1]compared to Model 1. [2]compared to Model 4.

Results multilevel models SCL

The results of the multilevel analyses for SCL are shown in Table 3. The three-level model provides a significantly better fit compared to the two-level model, ΔAIC = 12.85; $\chi 2(1) = 14.85$, $p < .001$. Therefore, the individual level was added as a third level in subsequent models. The ICCs indicated that 51.85% of the variance in SCL is attributable to variation between incidents within individuals (Level 2; ICC = 0.52) and that 34.13% of the variance in SCL is attributable to differences between individuals (Level 3; ICC = 0.34). 14.02% of the variance in SCL represents residual variability within incidents. In Model 1, time was added as a linear effect, resulting in significantly reduced model deviance, $\chi 2(1) = 5.07$, $p = .024$. Time was negatively associated with SCL (b = -0.06, $p < .001$), indicating that SCL decreased over time following the aggressive incident. In Model 2, time was added as a quadratic term, which resulted in a higher -2LL compared to the previous model (Δ Deviance = -38.60). Therefore, the quadratic effect of time was removed in subsequent models. In Model 3, the time varying covariates of movement and temperature were added to the model. Compared to Model 1, model deviance did not significantly decrease, $\chi 2(2) = 5.23$, $p = .073$, indicating that the addition of movement and temperature did not improve model fit. Therefore, movement and temperature were removed from subsequent models. In Model 4 the time-invariant predictors of organisation and type of aggression were added to the model. Compared to Model 1, the addition of organisation and type of aggression did not result in a significant decrease of model deviance, $\chi 2(3) = 0.94$, $p = .816$. In Model 5 the interaction terms between time and organisation and time and type of aggression were added, which resulted in a higher -2LL compared to the previous model (Δ Deviance = -1.31). Therefore, organisation and type of aggression, and the interaction terms with time were not included in the final model. In the final model (M6) a random effect of time was added. Using an autoregressive covariance structure resulted in non-convergence. Therefore, a diagonal covariance structure was used. This resulted in a significantly better model fit compared to Model 1, $\chi 2(7) = 341.53$, $p < .001$. Based on the final model, the fixed effect showed that time has a significant negative effect on SCL (b = -0.10, $p = .002$), which means that SCL decreases during the 30 minutes after an aggressive incident has occurred.

Table 3

Fitted multilevel models for SCL across the six 5-minute epochs following aggressive incidents

Effect	Model 0	Model 1	Model 3	Model 4	Model 5	Model 6
Fixed effects						
Intercept	1.45** (0.23)	1.62** (0.24)	1.62** (0.24)	1.55** (0.49)	1.47** (0.50)	1.73** (0.27)
Time		-0.06** (0.02)	-0.06** (0.02)	-0.06** (0.02)	-0.03 (0.04)	-0.10** (0.03)
Movement			0.27 (1.08)			
Temperature			0.00 (0.01)			
Organisation (ref = UMCU)						
De Borg				0.03 (0.60)	0.35 (0.61)	
Pluryn				-0.08 (0.64)	-0.12 (0.65)	
Verbal aggression (ref = physical aggression)				0.12 (0.32)	0.08 (0.33)	
Interactions						
De Borg*Time					-0.12 (0.05)*	
Pluryn*Time					0.02 (0.05)	
Verbal aggression *Time					0.01 (0.04)	
Random effects						
	Variance components					Unstructured
Level 1	0.70**	0.69**	0.69**	0.69**	0.68**	a
Level 2 - intercept	2.60**	2.60**	2.61**	2.61**	2.63**	5.61**
Level 3 - intercept	1.71*	1.71*	1.70*	1.82*	1.79*	1.74**
Level 2 - slope						0.13**
Goodness of fit						
-2LL	2434.58	2429.51	2434.74	2428.57	2429.88	2087.98
AIC	2440.58	2435.51	2440.74	2434.57	2435.88	2107.98
$\Delta\chi^2$		5.07*	5.23	0.94[1]		341.53**[1]
Δdf		1	2	3		7

*Notes. Standard errors are in parentheses. * p < .05. **p < .01. Model 2 is omitted. [1] compared to Model 1. [a] Variance differs across timepoints*

4

Results multilevel models SCR

The results of the multilevel analyses for SCR are shown in Table 4. A three-level model resulted in a lower AIC (ΔAIC = 12.12) and significantly lower model deviance, $\chi^2(1) = 14.12, p = <.001$, compared to the two-level model. Therefore, a three-level multilevel model was used. Of the variance in SCR, 41.01% was attributable to variation between incidents within individuals (Level 2; ICC = 0.41), 26.42% of the variance in SCR was attributable to differences between individuals (Level 3; ICC = 0.26), and 32.57% of the variance in SCR represents residual variability within incidents. After adding time to the model (M1) a significantly reduced model deviance, $\chi^2(1) = 4.11, p = .043$ was found. Time was significantly negatively associated with SCR (b = -0.10, $p = .003$), indicating that SCR decreased over time following the aggressive incident. Adding time as a quadratic term resulted in a higher -2LL compared to the previous model (Δ Deviance = -5.53) and for this reason the quadratic effect of time was not included in subsequent models. The time-varying covariates of movement and temperature (M3) significantly improved the model fit, $\chi^2(2) = 33.13, p < .001$ compared to M1. In model 3 there was still a significant negative association between time and SCR (b = -0.11, $p = .002$) when controlling for movement and temperature. Movement significantly predicted SCR (b = 12.04, $p <.001$), indicating that higher movement is associated with higher SCR. Temperature did not significantly predict SCR (b = 0.00, $p = .781$). Therefore, in Model 4, temperature was removed from the model, while movement was retained, which improved the -2LL and the AIC while also including less parameters (Δ Deviance = -7.07). This significantly improved model fit compared to Model 1 without the inclusion of movement, $\chi^2(1) = 40.20$, $p < .001$. In Model 5 the time-invariant predictors of organisation and type of aggression were added to the model. Compared to Model 4, the addition of organisation and type of aggression did not result in a significant decrease of model deviance, $\chi^2(3) = 4.66, p = .199$. In Model 6 the interaction terms between time and organisation and time and type of aggression were added, which resulted in a higher -2LL compared to the previous model (Δ Deviance = -4.48). Therefore, organisation and type of aggression, and the interaction terms with time were not included in the final model. In the final model (M7) a random effect of time was added. Using an autoregressive covariance structure resulted in non-convergence. Therefore, a diagonal covariance structure was used. This resulted in a significantly better model fit compared to Model 4, $\chi^2(7) = 72.90, p < .001$. The final model showed significant negative effect of time on SCR (b = -0.11, $p = .002$), which means that over the course of the 30 minutes after an aggressive incident the SCR decreases. The covariate movement had a significant positive effect on SCR (b = 10.98, $p < .001$).

Table 4

Fitted multilevel models for SCR across the six 5-minute epochs following aggressive incidents

Effect	Model 0	Model 1	Model 3	Model 4	Model 5	Model 6	Model 7
Fixed effects							
Intercept	3.57** (0.26)	3.84** (0.28)	3.84** (0.28)	3.84** (0.28)	4.56** (0.57)	4.73** (0.60)	3.86** (0.30)
Time		-0.10** (0.04)	-0.11** (0.03)	-0.11** (0.03)	-0.11** (0.03)	-0.17* (0.08)	-0.11* (0.05)
Movement			12.04** (1.96)	12.04** (1.95)	12.12** (1.95)	11.89** (1.95)	10.98** (2.00)
Temperature			0.00 (0.01)				
Organisation (ref = UMCU)							
De Borg					-0.76 (0.69)	-0.87 (0.73)	
Pluryn					-0.39 (0.74)	-0.87 (0.77)	
Verbal aggression (ref = physical aggression)					-0.42 (0.36)	-0.32 (0.40)	
Interactions							
De Borg*Time						0.04 (0.09)	
Pluryn*Time						0.18* (0.09)	
Verbal aggression *Time						-0.04 (0.07)	
Random effects							
	Variance components						Unstructured
Level 1	2.54**	2.51**	2.39**	2.39**	2.39**	2.38**	[a]
Level 2 - intercept	3.20**	3.20**	3.05**	3.01**	2.90**	2.92**	4.43**
Level 3 - intercept	2.06*	2.06*	2.21*	2.24*	2.50**	2.47**	2.23*
Level 2 - slope							0.19**
Goodness of fit							
-2LL	3270.51	3266.40	3233.27	3226.20	3221.54	3226.02	3153.30
AIC	3276.51	3272.40	3239.27	3232.20	3227.54	3232.02	3173.30
$\Delta\chi^2$		4.11*	33.13**	40.20**[1]	4.66		72.90**[2]
Δdf		1	2	1	3		7

*Notes. Standard errors are in parentheses. * $p < .05$. ** $p < .01$. Model 2 is omitted from the table. [1] compared to Model . [2]compared to Model 4. [a] Variance differs across timepoints*

Selecting classes in Latent Class Growth Curve Analysis

HR

We performed a LCGA to identify distinct trajectories over time in HR following the aggressive incidents. The fit indices of the LCGA models with one to eight classes used to determine the best fitting class solution are shown in Table 5. The 3-class solution showed the lowest BIC value, while the 7-class model has the lowest AIC, and the 7-class model was the preferred model based on the AIC3. We therefore performed two BLRTs to determine the best fitting class solution. Based on the results of the BLRT, the 6-class model had a significantly better model fit compared to the 3-class model ($p < .001$). However, the 7-class model did not provide an additional improved model fit over the 6-class model ($p = .198$). Therefore, the 6-class model was retained as the best fitting solution.

SCL

The fit indices of the LCGA models with one to eight classes used to determine the best fitting class solution for SCL are shown in Table 5. The BIC, AIC and AIC3 tended to decrease with the addition of additional classes. Therefore, strictly adhering to these fit indices, we would conclude that there are eight or more classes in the data. In Figure 7 we plotted the values of the model fit statistics from the models with one to eight classes. The rate with which each additional class decreases the model fit statistics tends to be smaller after the addition of more classes. And there is an elbow at the 4-class solution. The addition of more classes beyond the 4-class model improves the model only slightly. Therefore, we focus on the more parsimonious 4-class model.

SCR

The fit indices of the LCGA models with 1 to 8 classes used to determine the best fitting class solution for SCR are shown in Table 5. The 6-class solution showed the lowest BIC value, while the 8-class model has the lowest AIC, and the 7-class model was the preferred model based on the AIC3. We therefore performed two bootstrapped likelihood ratio tests (BLRT) to determine the best fitting class solution. Based on the results of the BLRT, the 7-class model did not have a significantly better model fit compared to the 6-class model ($p = .158$). Similarly, the 8-class model did not provide significantly better model fit over the 6-class model ($p = .082$). Therefore, the 6-class model was retained as the best fitting solution.

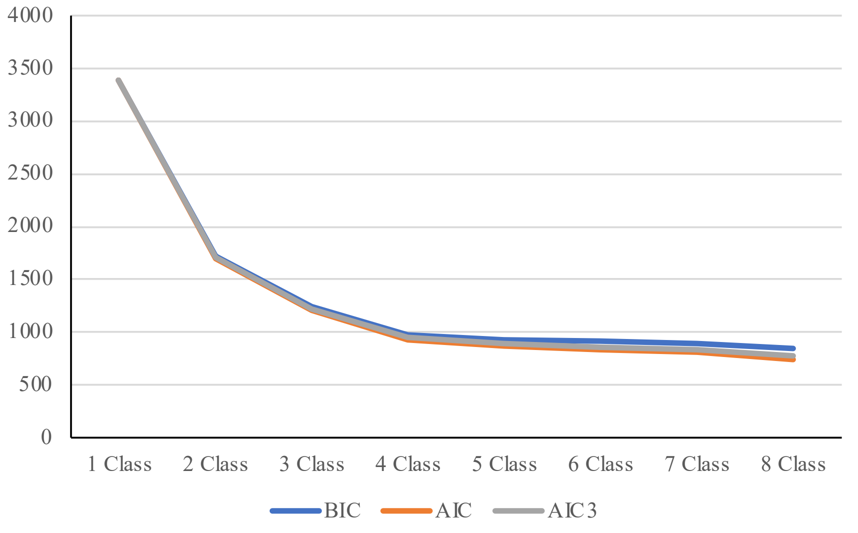
Table 5
Comparison of model fit for Latent Class Growth models of HR, SCL and SCR

	HR				SCL				SCR			
Classes	BIC	AIC	AIC3	BLRT	BIC	AIC	AIC3	BLRT	BIC	AIC	AIC3	BLRT
1-Class model	6149.07	6139.75	6142.75		3389.70	3380.38	3383.38		3810.83	3801.51	3804.51	
2-Class model	5999.51	5977.77	5984.77		1718.80	1697.06	1704.06		3329.39	3307.65	3314.65	
3-Class model	5917.24	5883.08	5894.08		1244.54	1210.37	1221.37		3197.56	3163.39	3174.39	
4-Class model	**5917.49**	**5870.90**	**5885.90**		**981.76**	**935.17**	**950.17**		**3114.88**	**3068.29**	**3083.29**	
5-Class model	5924.24	5865.23	5884.23		934.63	875.62	894.62		3093.56	3034.55	3053.55	
6-Class model	5926.64	5855.21	5878.21	<.001[1]	912.08	840.64	863.64		3088.38	3016.94	3039.94	
7-Class model	5936.86	5852.99	5879.99	.198[2]	892.09	808.23	835.23		3097.85	3013.99	3040.99	.158[2]
8-Class model	5950.73	5854.44	5885.44		843.76	747.48	778.48		3108.96	3012.68	3043.68	.082[2]

Note. *The selected models are in bold. Abbreviations: BIC, Bayesian information criterion; AIC, Akaike information criterion; BLRT, bootstrapped likelihood ratio test.*
[1] compared to the 3-class model. [2] compared to the 6-class model.

Figure 7

Model fit statistics of latent class growth curve models of SCL with increasing number of classes



Predictors of latent classes

HR

Using a bias-adjusted three-step approach, we assessed whether the type of aggression and the organisation in which the incident occurred predicted the latent class membership of incidents. The type of aggression was found to significantly predict latent class membership (Wald = 586.91, $p < .001$). Incidents of verbal aggression were found to be more likely to belong to class 6 (low intercept, stable slope) and less likely to belong to class 1 (high intercept, decreasing slope) and class 2 (moderate intercept, stable slope). Aggression against objects was associated with a higher likelihood of belonging to class 4 (low intercept, decreasing slope) and class 5 (high intercept, increasing slope), and a lower likelihood of belonging to class 1 (high intercept, decreasing slope), class 2 (moderate intercept, stable slope) and class 6 (low intercept, stable slope). Incidents of auto-aggression were more likely to belong to class 1 (high intercept, decreasing slope), class 2 (moderate intercept, stable slope) and class 6 (low intercept, stable slope), while being less likely to belong to class 3 (moderate intercept, decreasing slope), class 4 (low intercept, decreasing slope), and class 5 (high intercept, increasing slope). Physical aggression was more likely to belong to class 2 (moderate intercept, stable slope), class 3 (moderate intercept, decreasing slope), and class 4 (low intercept, decreasing slope) and less likely to belong to class 1 (high intercept, decreasing slope), class 5 (high intercept, increasing slope), and class 6 (low intercept, stable slope). The organisation in which the incidents occurred significantly predicted latent class membership (Wald = 64.42, $p < .001$). Incidents collected at de Borg were more likely to belong to class 4 (low intercept, decreasing slope) and less likely to belong to class 1 (high intercept, decreasing slope) and class 2 (moderate intercept, stable slope). Incidents from Pluryn were more likely to belong to class 3 (moderate intercept, decreasing slope) and less likely to belong to class 4 (low intercept, decreasing slope). Incidents from UMCU were more likely to belong to class 1 (high intercept, decreasing slope) and class 2 (moderate intercept, stable slope) and less likely to belong to class 3 (moderate intercept, decreasing slope).

Figure 8

Type of aggression as predictor of latent classes of HR

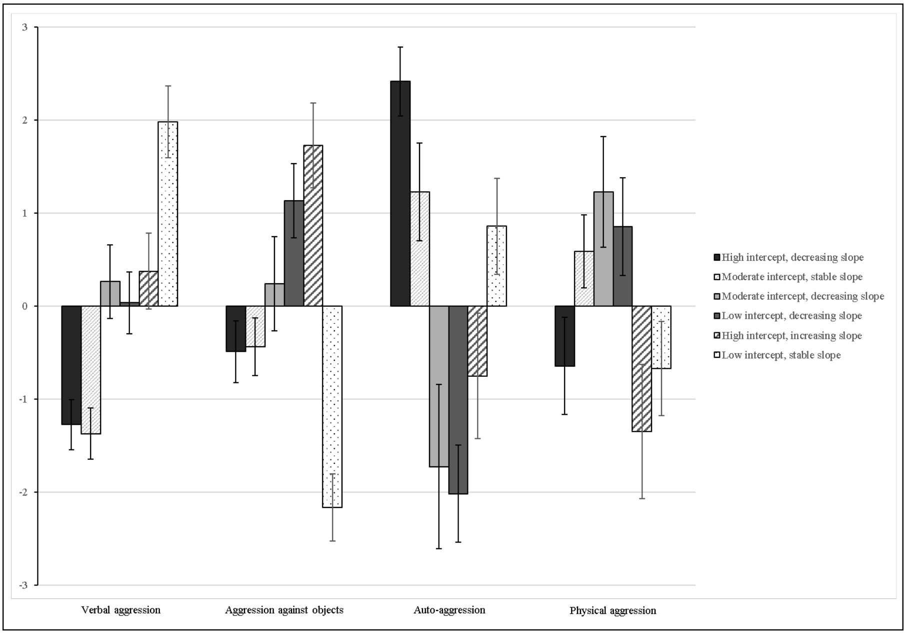
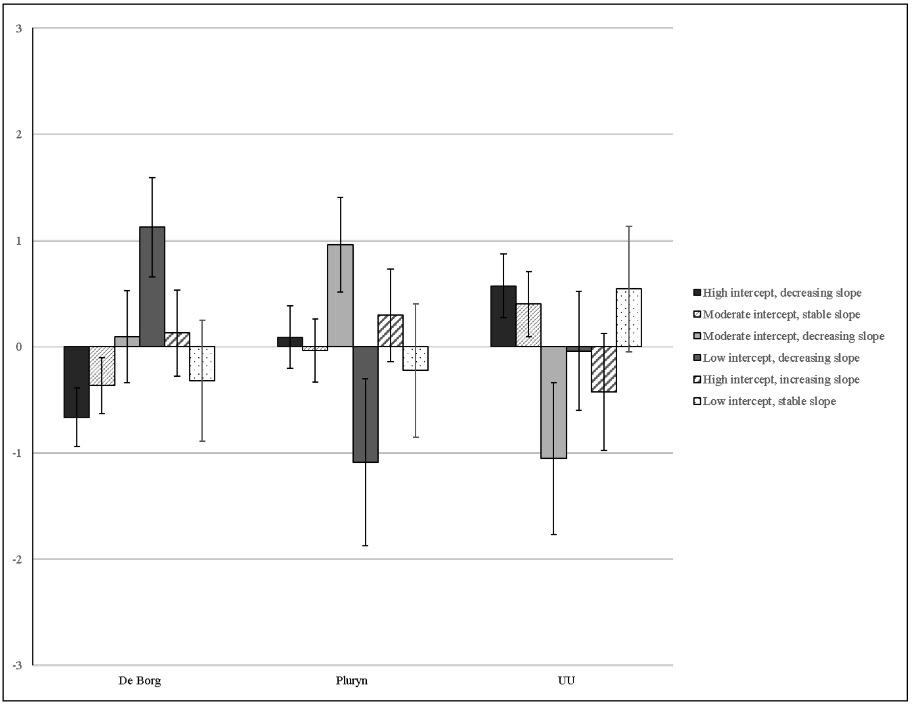


Figure 9

Organisation as predictor of latent classes of HR



SCL

Using a bias-adjusted three-step approach, we assessed whether the type of aggression and the organisation in which the incident occurred predicted the latent class membership of incidents. The type of aggression was found to significantly predict latent class membership (Wald = 222.42, p < .001). Incidents of verbal aggression were found to be more likely to belong to class 4 (high intercept, decreasing slope) and less likely to belong to class 2 (high intercept, stable slope) and class 3 (moderate intercept, stable slope). Aggression against objects was associated with a higher likelihood of belonging to class 4 (high intercept, decreasing slope), and a lower likelihood of belonging to class 2 (high intercept, stable slope), and class 3 (moderate intercept, stable slope). Incidents of auto-aggression were more likely to belong to class 2 (high intercept, stable slope) and class 3 (moderate intercept, stable slope), while being less likely to belong to class 4 (high intercept, decreasing slope). Physical aggression was more likely to belong to class 4 (high intercept, decreasing slope) and less likely to belong to class 2 (high intercept, stable slope), class 3 (moderate intercept, stable slope). The organisation in which the incidents occurred significantly predicted latent class membership (Wald = 47.73, p < .001). Incidents collected at de Borg were more likely to belong to class 1 (low intercept, stable slope) and less likely to belong to class 3 (moderate intercept, stable slope). Incidents from Pluryn were more likely to belong to class 2 (high intercept, stable slope) and class 3 (moderate intercept, stable slope) and less likely to belong to class 1 (low intercept, stable slope). Incidents from UMCU were more likely to belong to class 1 (low intercept, stable slope) and less likely to belong to class 2 (high intercept, stable slope).

Figure 10

Type of aggression as predictor of latent classes of SCL

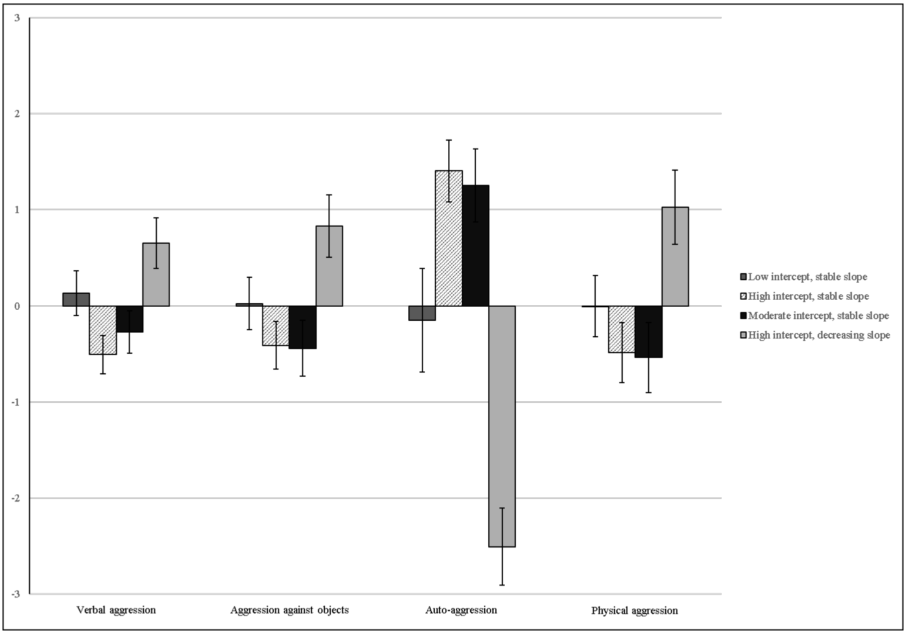
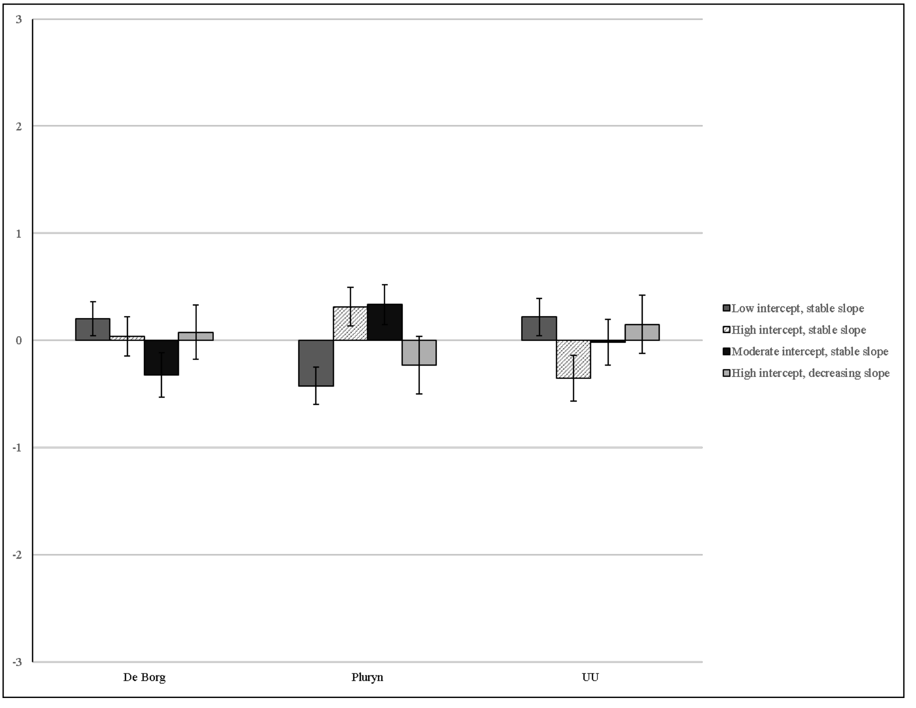


Figure 11

Organisation as predictor of latent classes of SCL



SCR

The type of aggression was found to significantly predict latent class membership (Wald = 248.31, $p < .001$). Incidents of verbal aggression were found to be more likely to belong to class 5 (low intercept, decreasing slope) and less likely to belong to class 2 (high intercept, decreasing slope), class 3 (moderate intercept, stable slope), and class 6 (moderate intercept, increasing slope). Aggression against objects was associated with a higher likelihood of belonging to class 5 (low intercept, decreasing slope), and a lower likelihood of belonging to class 4 (high intercept, stable slope). Incidents of auto-aggression were more likely to belong to class 1 (low intercept, stable slope), class 3 (moderate intercept, stable slope) and class 6 (moderate intercept, increasing slope), while being less likely to belong to class 5 (low intercept, decreasing slope). Physical aggression was more likely to belong to class 5 (low intercept, decreasing slope) and less likely to belong to class 1 (low intercept, stable slope). The organisation in which the incidents occurred significantly predicted latent class membership (Wald = 53.48, $p < .001$). Incidents collected at de Borg were more likely to belong to class 5 (low intercept, decreasing slope). Incidents from Pluryn were more likely to belong to class 6 (moderate intercept, increasing slope) and less likely to belong to class 2 (high intercept, decreasing slope), class 3 (moderate intercept, stable slope) and class 5 (low intercept, decreasing slope). Incidents from UMCU were more likely to belong to class 3 (moderate intercept, stable slope).

Figure 12

Type of aggression as predictor of latent classes of SCR

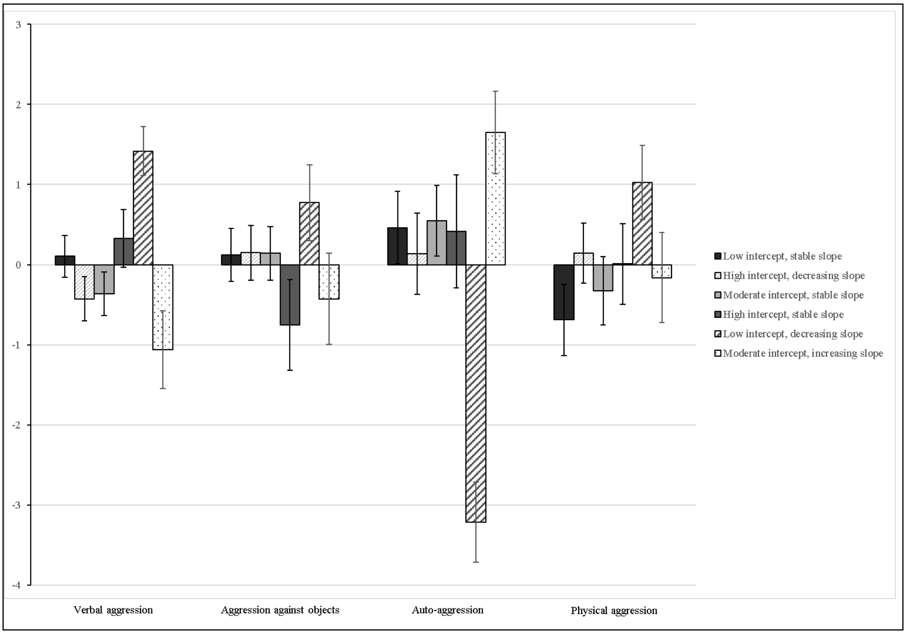
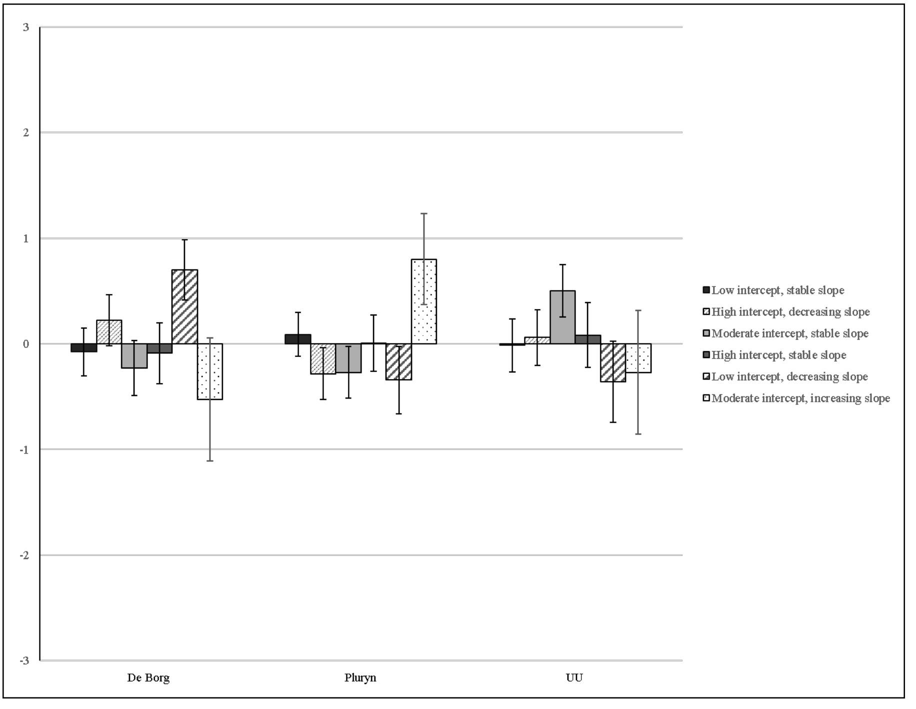
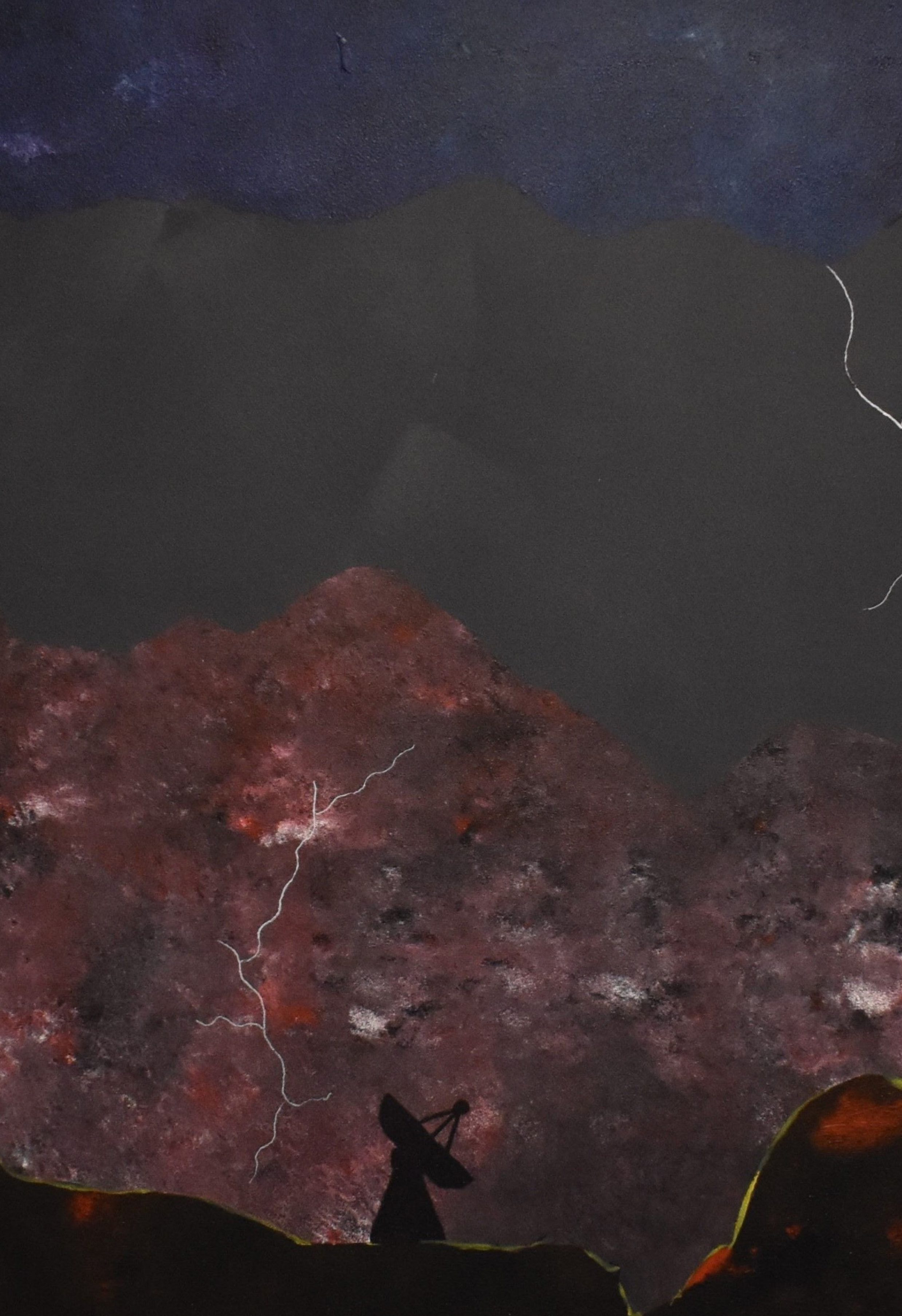
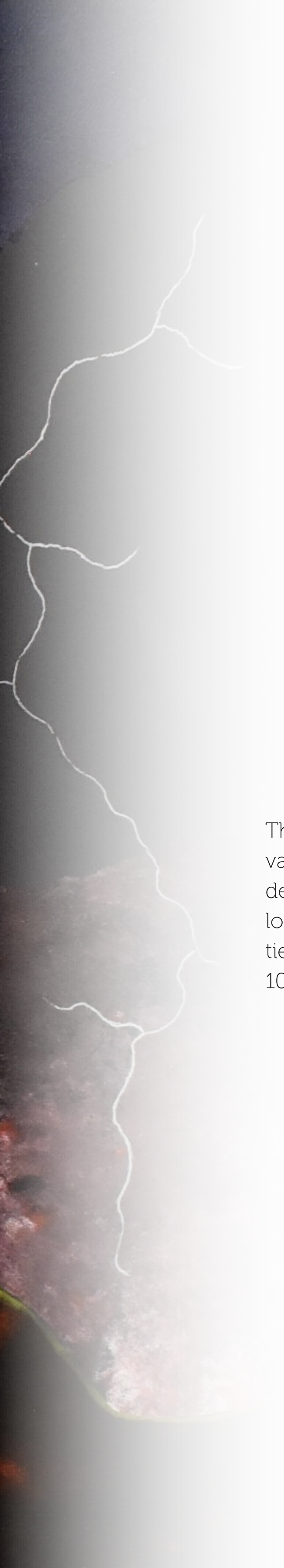


Figure 13

Organisation as predictor of latent classes of SCR







Chapter 5

Listening to music is associated with reduced physiological and subjective stress in people with mild intellectual disabilities: A biofeedback study

This chapter was published as:
van Swieten, M., de Looff, P., VanDerNagel, J., Bouwmeester, S., & Didden, R. (2025). Listening to music is associated with reduced physiological and subjective stress in people with mild intellectual disabilities: A biofeedback study. *Research in Developmental Disabilities, 161*, 104976. https://doi.org/10.1016/j.ridd.2025.104976

ABSTRACT

Background: Many people with mild intellectual disabilities are at increased risk to experience stress. Reducing stress is important because experiencing prolonged and elevated stress can have detrimental effects on mental and physical health and it is associated with aggressive behaviour and self-harm.
Aims: This preliminary study investigated whether an intervention combining biofeedback with listening to music is effective and whether a personalized music playlist is more effective than self-selected music in reducing physiological and subjective stress in participants with mild intellectual disabilities.
Methods: We collected 103 music listening sessions over a period of 2-4 weeks for 11 participants throughout their daily routines. They listened to music when they received biofeedback on their increased stress level (as measured by wearable biosensor Nowatch) or when they themselves felt stressed. Participants listened either to self-selected music or to a personalised playlist compiled with X-system, music technology that predicts the effect of a song on levels of autonomic arousal. Pulse rate (PR) and skin conductance level (SCL) were measured with the EmbracePlus and subjective feelings of stress and mood were measured with two scale questions. After the intervention phase, participants and their caregivers completed a short questionnaire to evaluate their experiences with using the X-system playlist.
Results: Mixed regression analyses showed reductions in PR and SCL during listening to music, and indications were found for reductions in subjective stress and improvement of mood after intervention. Listening to music compiled with X-system was not more effective than listening to self-selected music. However, lower combined arousal values (a feature of X-system) from self-selected and X-system music predicted lower PR and SCL, indicating that these indices can be used to select songs that have a relaxing or energizing effect.
Conclusions and implications: The present study suggests that music listening is associated with both subjective and physiological stress reduction. Listening to music might be an accessible, inexpensive and empowering strategy for stress reduction and improving emotion regulation, which could also benefit mental and physical health. Several challenges were encountered while implementing the intervention and suggestions for future research are given.

INTRODUCTION

Regulating stress levels is of great importance as experiencing prolonged and elevated stress can have detrimental effects on mental and physical health. Stress can be defined as (an actual or anticipated) state of disrupted homeostasis (Ulrich-Lai & Herman, 2009), by a wide range of intrinsic or extrinsic, real or perceived challenges or stimuli, defined as stressors (Bali & Jaggi, 2015). Inadequate regulation of acute stress is associated with chronic stress, which can lead to mental health problems such as anxiety and depression (Everly & Lating Jr, 2019; Konstantopoulou et al., 2020; Turner et al., 2020) and physical health problems, including cardiovascular disease, metabolic diseases and a weakened immune system (Agorastos & Chrousos, 2022; Everly & Lating Jr, 2019; Kivimäki et al., 2023). In addition, inadequate coping and increased stress levels are associated with problems in emotion regulation and impulse control, which are risk factors of destructive behaviours such as aggressive behaviour and self-harm (Hooley & Franklin, 2018). People with mild intellectual disabilities or borderline intellectual functioning (MID-BIF; IQ 50-85) are at increased risk to experience (high levels of) stress compared to the general population (e.g. Forte et al., 2011; Griffith et al., 2013) because of deficits in intellectual functioning and adaptive skills (American Psychiatric Association, 2022), combined with an increased risk of exposure to adverse life events or even post-traumatic stress disorder (Mevissen et al., 2016; Wigham & Emerson, 2015). Moreover, living in a residential setting is associated with multiple stressors such as experiencing a lack of control over the environment and choice of activities, aggressive behaviour from other clients, and stressful auditory stimulation (e.g. noisy environment) (Griffith et al., 2013; Neimeijer et al., 2021).

Supporting clients with MID-BIF in stress regulation can be challenging for caregivers, as they may not be aware of the various stressors affecting their clients or their current level of stress (Lunsky & Bramston, 2006). In addition, individuals who experience significant amounts of stress can find it difficult to communicate their stress as their interoceptive awareness of stressful bodily states might be impaired (Bellemans et al., 2022; Schulz & Vögele, 2015). Furthermore, caregivers often report that self-harm or aggressive behaviour was not preceded by any 'warning signs' or a visible build-up in tension. Wearable biosensors that measure physiological arousal related to stress, such as increased heart rate and electrodermal activity (EDA) (Boucsein, 2012; Everly & Lating Jr, 2019) and provide real-time insight in arousal (i.e. biofeedback) may aid clients and caregivers to effectively detect and manage stress (De Vries et al., 2023). For instance, Ter Harmsel et al. (2021) concluded that ambulatory biofeedback interventions seem to aid emotion and stress regulation, both on a psychological and physiological level, though these physiological measures were used less frequently. Similar results were reported by the systematic review by (De Witte et al., 2019) who found preliminary evidence that the use of biofeedback can improve both physiological and psychological indicators of stress. They concluded that biofeedback could provide an accessible and low-cost addition to stress interventions, but further research into the effectiveness of different components of biofeedback interventions is needed.

Following stress detection, adequate coping strategies and interventions are needed to reduce stress to an optimal level (Everly & Lating Jr, 2019, Figure 1.3). Deficits in intellectual functioning can impair the development of such adaptive strategies to cope with stress (Taylor & Novaco, 2005). Indeed, people with MID-BIF more often use maladaptive coping strategies than people without MID-BIF (Hartley & MacLean Jr, 2008).

Listening to music can be an adequate coping strategy; it has repeatedly shown to reduce subjective stress and affect the physiological response to stress, for example in terms of decreased heart rate (De Witte et al., 2020; Lynar et al., 2017). Furthermore, research showed that SCL and mood can be directed toward an energized or calm state by listening to music and that SCL remains in these states for at least 30 min after listening to music (Van der Zwaag et al., 2012). Listening to music could be an accessible intervention for reducing stress in people with MID-BIF because it hardly appeals to their cognitive skills, is low-cost, low-risk, popular as a daily activity and is widely applicable because it can be accessed at many locations and moments. However, limited research is available on the effect of listening to music on physiological stress in people with MID-BIF.

Caregivers have expressed concerns about whether clients are selecting appropriate music for relaxation, as they have regularly observed instances where clients, for example, choose to listen to hard rock when feeling stressed. Alternatively, music could also be selected and ordered by innovative music technology. X-system is such technology and can be used to select songs to reach a desired state of arousal (i.e. activation or relaxation). X-System is designed to predict the innate neurophysiological response to songs and offers a web- application that predicts the effect of a song on levels of autonomic arousal (Osborne et al., 2017). Nijman et al. (2023) studied the differential effect of listening to music in a preferred genre selected and ordered by X-system versus music in random order on stress reduction in clients and caregivers of a medium secure forensic psychiatric facility. This study showed that physiological indices and self-reported stress decreased significantly after listening to music. An accelerated reduction in SCL for the X-system playlist compared to the playlist in random order was found with visual inspection of the data, but the trend was non-significant. A limitation of this study was its controlled setting: participants listened to music on appointment times in a quiet room, which might limit the representativeness of daily life effects. Furthermore, this study selected playlists from each participant's preferred music genre, but these playlists might not have contained the specific song that participants regularly listen to or particularly like. Listening to specific preferred songs might improve the effectiveness of music listening on stress-reduction (Jiang et al., 2013).

The aim of the current preliminary study was to investigate whether an intervention combining biofeedback with listening to music (selected with X-system or self-selected) is associated with a reduction of (increased) physiological and subjective stress in people with MID-BIF. The current naturalistic study contributes to the literature by studying this in the ward at moments of 'real' stress instead of appointed pre-defined times and with specific preferred songs instead of a preferred genre. First, it was examined whether listening to music was associated with a reduction in pulse rate (PR, the number of heart beats per minute, measured on the wrist), SCL and subjective stress. Second, it was examined whether PR and SCL over time differed between the X-system condition and self-selected music condition. Furthermore, the relationship between combined arousal values of all songs, from both the X-system and self-selected music condition, and PR and SCL was explored. It was expected that while listening to music in general predicts a decrease in PR, SCL and subjective stress, we hypothesized that X-system predicts a greater decrease in PR, SCL and subjective stress than self-selected music. Finally, lower combined arousal values of listened songs were hypothesized to be associated with lower PR and SCL. Besides preliminary results on the association between music listening (with X-system) and stress reduction, this paper provides suggestions for future research.

METHOD

Participants

This study was conducted at 8 locations of 4 treatment facilities for adults with MID-BIF, severe behavioural and/or mental disorders, problems in multiple areas of life and often a history of substance abuse. Clients are admitted to the facilities under criminal law, civic law or on a voluntary basis, often for externalizing behaviour problems (i.e. aggression or a sexual offence) and/or internalizing problems (such as self-harm and suicide attempts) (Delforterie et al., 2020). The participants were recruited by caregivers, trainers and therapists. The following inclusion criteria were used: (1) the client had MID or BIF, (2) the client was allowed to use a phone for participating in the study; (3) the client had a basic understanding of using mobile applications; (4) informed consent was given by the client and their legal representative; (5) the client had no severe hearing loss; and (6) the client's therapist evaluated the participant as eligible, considering the risks and benefits of participating for the client. This resulted in an initial sample of 19 participants, of whom two participants dropped out on the first day of the music intervention phase. Reportedly, this was due to stress increase related to participation in the study and loss of motivation due to unknown reasons, respectively. In addition, five participants were excluded from the analyses because they did not listen to music longer than 10 consecutive minutes and one participant was excluded from the analyses because of missing data on music listening.

The final sample consisted of 11 participants of whom six were females and five were males. The age ranged from 22 to 57 years old (*M* = 31.6; *SD* = 11.1). Six participants had a mild intellectual disability, four had borderline intellectual functioning, and one participant did not want to share whether s/he had MID or BIF. The following psychiatric disorders were reported in the participants' medical record: substance use disorder (*n* = 6), posttraumatic stress disorder (PTSD; *n* = 5), personality disorder (*n* = 5), attention deficit hyperactivity disorder (ADHD; *n* = 2), autism spectrum disorder (*n* = 2), depressive disorder (*n* = 1) and social anxiety disorder (*n* = 1). Six participants were prescribed psychotropic drugs at the time of participation including antipsychotics (*n* = 4), selective serotonin reuptake inhibitors (SSRIs; *n* = 3), benzodiazepine agonists (*n* = 2), amphetamines (*n* = 2), anti-epileptics (*n* = 2) and non-SSRI mood stabilizers (*n* = 2).

Instruments

Nowatch

The Nowatch is a wearable biosensor in the form of a wristband that was used to provide biofeedback to the participants. It continuously measures skin conductance and provides a 'stress level score' based on 'Biosensing EDA software and algorithms' (Philips, n.d.). Stress level scores range from 0 to 1000, where 0 indicates no current stress and 1000 indicates a high level of current stress (Van der Mee et al., 2021). Furthermore, it measures PR (photoplethysmography, PPG), heart rate variability (HRV) at rest, breathing rate, sleep, steps and movement (accelerometer). The Nowatch wristband operates with a phone application which was installed on an iPhone SE provided to participants during the intervention period. The application included two scale questions measuring mood and stress, based on the Affective Slider (Betella et al., 2016). By moving a slider on a scale with an emoticon at each end of the scale, participants reported the level of experienced stress and pleasure. This resulted in a score ranging from 0 to 100, where 0 indicates very high stress and very bad mood and 100 indicates no stress and very good mood.

EmbracePlus

The EmbracePlus is a wearable biosensor in the form of a wristband that was used to continuously measure PR and EDA. We used the two pre-processed metrics PR and EDA, calculated by Empatica's algorithms from the EmbracePlus sensor raw data. The PR algorithm uses data obtained from the EmbracePlus PPG sensor and the signals collected by the 3-axis accelerometer sensor embedded in the device. PR values are expressed in beats per minute (bpm). The EDA algorithm provides a measure of SCL every 1 min in μSiemens (μS).

X-system

X-system is a music technology that uses algorithms to predict the effect of a song on a person's level of autonomic arousal (Osborne et al., 2017). We used the combined arousal feature (for more information see Appendix 5A) of X-system to select and order songs with the aim to gradually lead the listener to a state of relaxation. The combined arousal feature has a range of 0.00 -1.00 where lower values predict relaxation effects on the user while higher values predict excitement. X-system playlists were compiled with songs that were preferred by the listener. On average, the playlists started with a combined arousal value of 0.76 and ended with a value of 0.24, had a mean difference of 0.08 in combined arousal between consecutive songs, had an average duration of 25 min in total and consisted of an average of 7 songs (with a range between 6 and 9).

Spotify

Spotify premium was used with a phone application as a streaming platform for both the X-system playlist and music selected by participants at that moment. The listening history of the Spotify data was retrieved and analysed. These data included the end times of listening sessions, the total time of each listening session in milliseconds, and the specific artists and songs listened to.

Questionnaire to evaluate the use of X-system playlists

A short questionnaire was created for the current study to evaluate the experiences of participants and their caregivers with using the X-system playlist. It consisted of four items for participants ('It was fun using the playlist'; 'I found it difficult to use'; 'Listening to the playlist helped me to relax'; 'I would like to use this for relaxation in the future') and one item for caregivers ('Listening to the playlist helped the client to relax'), with a 5-point Likert scale (1 = strongly disagree; 2 = disagree; 3 = neither agree nor disagree; 4 = agree; 5 = strongly agree).

Music intervention

Participants were instructed to wear the Nowatch on the dominant hand and the EmbracePlus on the non-dominant hand during the day. When participants had a 'high stress level score', as measured by the Nowatch, they received a vibration through the wristband and a notification of this high stress level on their research phone. The stress value at which the participant received a notification was personalized per individual to prevent too many or too few notifications (i.e. less than 1 per day or more than 3 per day). For the majority of participants, the value at which they received a notification was initially set at 500 but the value was lowered to 350 because participants received too few notifications and instructions to listen to music. The low number of vibrations could be due to technical factors, such as intermittent connection between the Nowatch and the phone (for example because a participant accidentally turned Bluetooth off), or the Nowatch not always detecting stress in every instance among individuals with MID-BIF during daily activities (false negatives). To prevent too many consecutive vibrations, an additional minimum of 30 min between notifications was applied. After a 'stress notification', participants answered two scale questions about their mood and subjective stress level at that moment. They were then instructed to listen to music for approximately 25 min. They alternately received an instruction in the Nowatch application to listen to their X-system playlist or to select music themselves at that moment. After 30 minutes, participants received a follow-up notification to complete the scale questions on mood and subjective stress again. To account for subjective feelings of stress, and to overcome the problem of false negatives, participants were instructed to also listen to music if they felt stressed but did not receive a stress notification from the wearable.

Procedure

This study was approved by the Ethics Committee of the Faculty of Social Sciences of the Radboud University (ECSW-2022-018). Participants were informed about the aims of the study by an information video and information letter. The participants, and if applicable their legal representative, signed informed consent, after which they were instructed to participate in the intervention for two weeks. In the event of any technical problems or infrequent wearing of the devices, the research period was extended to a maximum of 4 weeks. After the intervention period the participants and their caregivers completed a questionnaire to evaluate their experiences with using the X-system playlist. The data were collected from July 2023 to March 2024. The participants received a voucher worth 10 euros for participating in the study.

Data analysis

The music listening history was analysed and all occasions with 10 or more consecutive minutes of music listening were included as separate sessions. For every listening session, 60 min of physiological data, PR and SCL were extracted, that is 30 min before the start of music listening and 30 min from the start of music listening. Every minute of this timeframe was coded as no music, self-selected music, X-system music or X-system in random order (in case the participant shuffled songs). Sessions in which the participant listened to music in the baseline (30 minutes before music listening) were excluded. The time music was listened to was rounded to minutes because physiological data were provided in minutes.

To filter out artefacts mean PR < 30 or > 200 bpm were discarded in line with Koenig et al. (2023) and SCL values under 0.15 µS were discarded because low-intensity SCL measurements often indicate that the electrodes may have not fully coupled with the skin tissue, or the device is not worn correctly (Empatica, personal communication, May 15, 2024).

First, the following descriptives were calculated: mean PR and SCL during the 30 min before music listening, during X-system music and self-selected music listening for each participant, separately. Furthermore, mean combined arousal values of all X-system music and all self-selected music were calculated and an independent samples *t*-test was used to test whether the combined arousal values of X-system music differed significantly from those of self-selected music.

To examine whether listening to music in general (combining X-system and self-selected music) was associated with a reduction in PR and SCL, two 3-level mixed regression analyses were used (i.e. one for SCL and one for PR), as the data were hierarchical in nature. Repeated measures, SCL/PR (Level 1) were nested within sessions (Level 2), and nested within the participant (Level 3). A model was fitted with a random intercept for sessions and participants and fixed effects for time (standardized), music condition (music or no music listened), and the interaction time x music condition. Note that the combination of a limited number of repeated measurements within sessions, a limited number of sessions within participants and a limited number of participants was not sufficient to estimate random slope effects. Partial eta squared ($\eta p\,2$) was calculated with 0.01 indicating a small effect size, 0.06 indicating a medium effect size, and 0.14 indicating a large effect size (Yagin et al., 2024).

To examine whether PR and SCL over time differed between the X-system condition and self-selected music condition, two mixed regression model analyses (SCL and PR) with the same hierarchical structure as aforementioned were used, while adding a nominal dummy variable for self-selected music versus X-system (excluding sessions where X-system playlists were listened in random order).

Furthermore, the relationship between combined arousal values of all songs, from both the X-system and self-selected music condition, and PR and SCL was explored using mixed regression model analysis. All mixed regression analyses were performed using R (R Core Team, 2024) version 4.4.1 [packages: lme4, lmerTest] (Bates et al., 2015; Kuznetsova et al., 2017). Visualizations were created with the ggplot2 package (Wickham, 2016).

To examine changes in subjective stress and mood after listening to music, average scores of pre-test measures and post-test measures from the scale questions were calculated for 'no-music', 'self-selected music', 'X-system music' and 'combination of X-system and self-selected music'. The scores on stress were mirrored, with 0 indicating no stress and 100 indicating very high stress. These results were plotted in a line graph. No statistical tests were used because of insufficient (post-test) responses of participants on the scale questions.

RESULTS

The total number of music listening sessions was 103 (19 following biofeedback instruction and 84 without instruction but when participant felt stressed themselves). Participants listened to their X-system playlist (partially) in 15 of these sessions (in correct order), 67 of the sessions included music self-selected at that moment, 16 sessions consisted of a combination of X-system and self-selected music and in 5 sessions participants listened to the X-system playlist in random order. The combined arousal value from listened songs ranged from 0.09 to 0.94, and the mean combined arousal of self-selected songs (M = 0.59, SD = 0.18) was significantly higher than the mean combined arousal value of X-system music (M = 0.52, SD = 0.16), as calculated by the algorithms from X-system (t(1172,86) = 9.79, p < .001). The mean SCL and PR during no music (30 minutes preceding music listening), listening to X-system music and to self-selected music per participant are displayed in Table 1.

Table 1

Mean SCL and PR per Participant

Participant	No music		X-system		Self-selected	
	M(SD) SCL	M(SD) PR	M(SD) SCL	M(SD) PR	M(SD) SCL	M(SD) PR
1	0.64(0.78)	88.69(16.96)	- [a]	77.28(7.46)	0.86(0.94)	75.05(9.31)
2	2.59(2.70)	90.50(16.35)	2.40(2.68)	85.64(13.05)	4.87(1.99)	82.51(13.78)
3	1.06(0.66)	99.26(22.16)	1.35(0.94)	96.51(12.57)	1.80(2.07)	103.43(15.20)
4	1.27(0.08)	85.27(8.90)	- [b]	- [b]	1.48(0.01)	74.55(6.25)
5	1.86(2.43)	105.68(17.35)	3.08(2.82)	106.36(12.56)	2.31(2.06)	110.62(24.96)
6	0.15(0.01)	89.16(24.97)	- [b]	- [b]	0.18(0.01)	89.00(11.67)
7	- [a]	81.68(21.82)	- [a]	78.75(15.15)	- [a]	76.38(15.71)
8	0.18(0.04)	90.39(16.20)	- [a]	78.58(11.46)	- [a]	86.80(14.79)
9	- [a]	78.43(12.24)	- [a]	78.45(3.05)	- [a]	- [a]
10	- [a]	87.22(9.37)	- [b]	- [b]	0.54(0.08)	85.30(6.51)
11	2.24(2.38)	89.39(13.64)	1.02(0.60)	80.52(7.93)	1.82(2.30)	80.16(8.65)

Note. M = mean, SD = standard deviation

[a] *No mean SCL and correlation could be calculated because of missing values and/or because of exclusively low SCL values (<0.15), which were excluded from analyses due to possible artefacts.*

[b] *Did not listen to X-system or only listened to X-system in random order.*

Changes in physiology during listening to music

Mixed regression analysis showed that SCL significantly increased over time within a session (*b* = 0.02), but there was no main effect for music versus no music on SCL. The interaction effect (*b* = -0.04) shows that listening to music (X-system, self-selected or a combination) resulted in a significant decrease in SCL over time compared to not listening to music (see Table 2). PR significantly increased over time within a session. PR was lower during music listening compared to the half hour before music listening (*b* = -1.62), but there was no significant main effect for music on PR. The significant interaction effect (*b* = -0.37) shows that listening to music (X-system, self-selected or a combination) resulted in a decrease in PR over time in comparison to the preceding half hour without listening to music. The ICC shows that the correlation between sessions within participants is low (SCL = 0.19; PR = 0.20), indicating that the effect over time of SCL and PR and the effect of music versus no music within a participant can differ greatly from each other. See Figure 1 and 2 for the mean SCL and PR before and during music listening for each individual participant.

Table 2

Results from the Mixed Regression Analysis Predicting the SCL and PR

	Estimates	SE	df	t	p	ηp^2
		SCL				
Intercept	1.33	0.41	14.47	3.27	.005**	.386
Time standardized	0.02	0.01	1723.11	3.34	.001**	.007
Music	0.23	0.17	1715.52	1.35	.176	.066
Time standardized* Music	-0.04	0.01	1723.60	-3.61	<.001***	.011
		PR				
Intercept	90.47	3.15	17.09	28.75	<.001***	.319
Time standardized	0.10	0.04	4485.07	2.56	.011*	.000
Music	-1.62	0.92	4484.08	-1.75	.080	.006
Time standardized* Music	-0.37	0.06	4490.46	-6.18	<.001***	.001

Note. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

Figure 1

Mean SCL Before and During Music Listening for Each Individual Participant with Loess Line

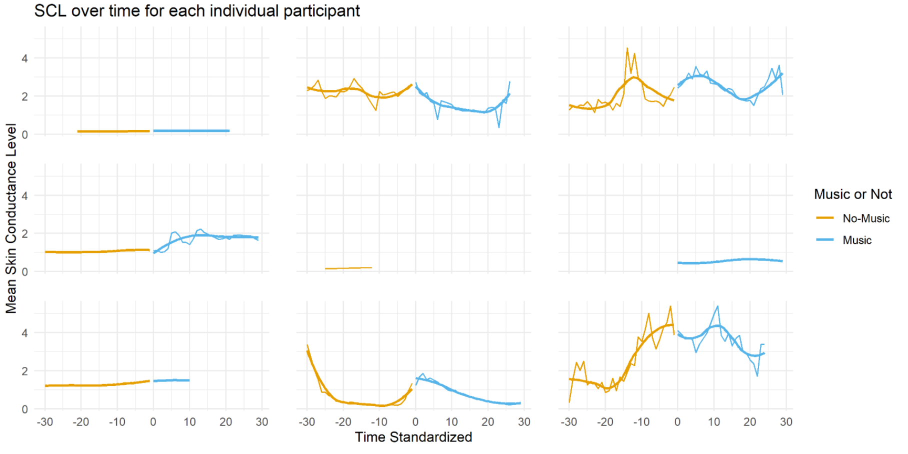
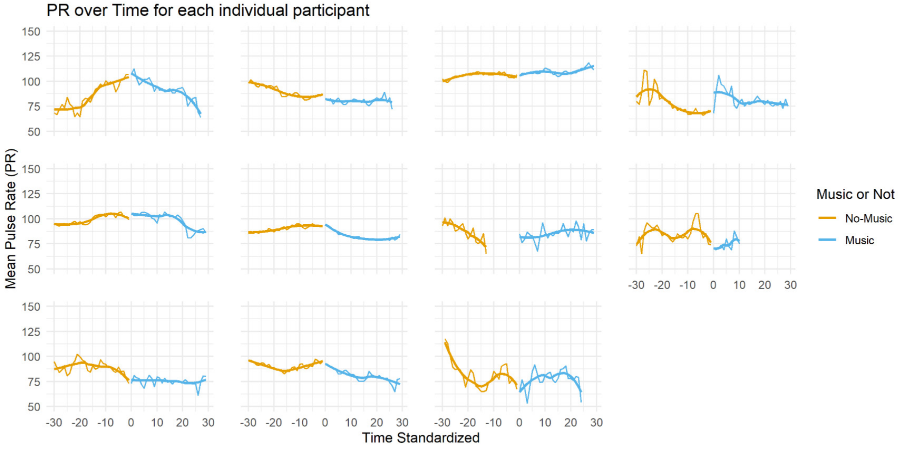


Figure 2

Mean PR Before and During Music Listening for Each Individual Participant with Loess Line



Changes in physiology during listening to X-system compared to self-selected music

When X-system sessions and self-selected music sessions were analysed separately, results of SCL showed that there was no significant main effect: neither for time, for no music compared to self-selected music nor for X-system compared to self-selected music (see Table 3). In line with the hypothesis, SCL increases significantly more over time (*b* = 0.04) during no music compared to self-selected music. Contrary to the hypothesis, no significant interaction effect was found with X-system, indicating no difference in SCL over time when X-system was compared to self-selected music.

Results showed that PR decreased significantly over time (see Table 3). PR was higher during no music compared to self-selected music and lower during X-system compared to self-selected music, but these main effects were not significant. In line with the hypothesis, a significant interaction effect with no music (*b* = 0.27) was found which showed that PR increased more over time during no music compared to self-selected music. Contrary to the hypothesis, no significant interaction effect was found with X-system, indicating no difference in PR over time when X-system was compared to self-selected music. The ICC shows that the correlation between sessions within participants is low (SCL = 0.19; PR = 0.22), indicating that the effect over time of X-system versus self-selected music on SCL and PR within a participant can differ greatly from each other.

Table 3

Results from the Mixed Regression Analysis Predicting the SCL and PR Comparing X-system and Self-selected Music

	Estimates	SE	Df	t	p	ηp^2
			SCL			
Intercept	1.61	0.41	15.23	3.90	.001**	.456
Time standardized	-0.02	0.01	1681.56	-1.74	.083	.005
No music	-0.26	0.19	1676.47	-1.42	.155	.075
X-system	0.00	0.31	1692.45	-0.01	.988	.001
Time standardized* no music	0.04	0.01	1681.47	3.32	.001**	.011
Time standardized* X-system	0.00	0.02	1684.48	-0.18	.860	.001
			PR			
Intercept	87.59	3.29	17.17	26.62	<.001***	.309
Time standardized	-0.18	0.06	4363.95	-3.07	.002**	.001
No music	1.91	1.02	4355.80	1.87	.061	.007
X-system	-2.96	1.67	4371.09	-1.77	.077	.010
Time standardized* no music	0.27	0.07	4361.25	3.99	<.001***	.001
Time standardized* X-system	-0.05	0.11	4371.67	-0.47	.640	.000

Note. Self-selected music is reference category.

** = p < .05; ** = p < .01; *** = p < .001*

Relation combined arousal values and physiology

When the relationship between all combined arousal values (from X-system music and self-selected music) and physiology over time was assessed, significant main effects of time and combined arousal and a significant interaction effect was found for SCL. For PR a trend ($b = -0.28$, $p = .052$) toward a significant main effect of time and a significant main effect of combined arousal was found. In line with the hypothesis, it was found that listening to songs with a lower combined arousal value was related to lower SCL and PR (see Table 4). This suggests that X-system can be used to select songs that have a relaxing or energizing effect in terms of physiology. The ICC shows a strong correlation for PR (0.96), indicating that the relationship between combined arousal and PR is very similar for the different sessions within the participants. The ICC for SCL was moderate (0.31).

Table 4

Results from the Mixed Regression Analysis Exploring the Relationship Between Combined Arousal Value and SCL and PR

SCL	Estimates	SE	df	t	p	ηp^2
SCL						
Time standardized	0.05	0.02	197.09	2.11	.036*	.012
Combined arousal	2.18	0.61	95.60	3.59	.001**	.555
Time standardized*Combined arousal	-0.11	0.04	265.37	-2.68	.008**	.028
PR						
Time standardized	-0.28	0.15	1972.86	-1.94	.052	.000
Combined arousal	9.78	3.85	1971.61	2.54	.011*	.001
Time standardized*Combined arousal	0.09	0.24	1972.63	0.38	.706	.000

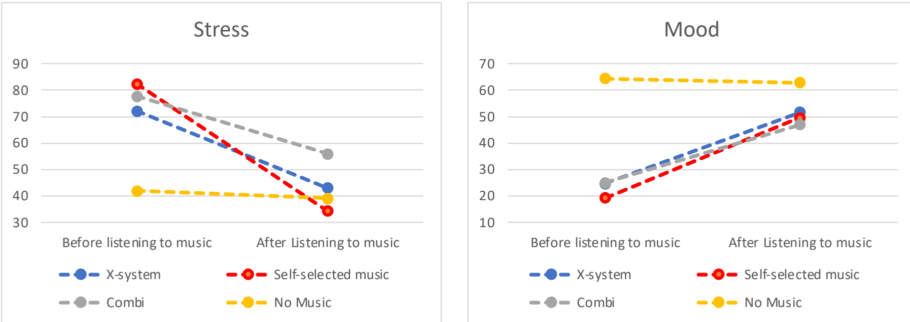
Note. * = $p < .05$; ** = $p < .01$.

Changes in subjective stress and mood after listening to music

Four of the 11 participants completed the pre-test and post-test measure of the affective slider on stress and mood one or more times (min = 1, max = 6 times) (see Figure 3). The number of measures for each condition are as follows: X-system = 5, self-selected music = 3, combination of X-system and self-selected music = 4 and no music = 14. Based on visual inspection, for all music listening conditions, self-reported stress was lower at post-test and mood improved at post-test compared to pre-test. Stress decreased slightly stronger after listening to self-selected music compared to X-system. When no music was listened to, stress and mood levels were already better at pre-test (stress: $M = 42.00$, SD = 35.28; mood: $M = 64.43$, $SD = 31.48$) and remained relatively stable at post-test (stress: $M = 39.21$, $SD = 34.15$; mood: $M = 62.92$, $SD = 34.25$).

Figure 3

Affective Slider Scores Concerning Stress and Mood, Pre- and Post-Listening to Music and No Music (n = 4)



Note. Combi = combination of X-system and self-selected music

Evaluation

Ten participants and caregivers filled in the questionnaire after the intervention period to evaluate the use of X-system playlists (see Table 5 in Appendix 5A). The results indicate that the majority of the participants found it relatively easy to use the X-system playlist, enjoyed using the X-system playlist, thought it helped them to reduce stress and would like to use the playlist in the future. Most caregivers reported that listening to the X-system playlist helped their client to reduce stress.

DISCUSSION

Main findings

This study explored changes in physiological stress during and subjective stress and mood following listening to music in adults with MID-BIF in a naturalistic setting. The main results showed reductions in PR and SCL during listening to music, and indications were found for reductions in subjective stress after listening to music. Listening to music compiled with X-system was not more effective than listening to self-selected music. However, the combined arousal values of songs, a feature of X-system, were positively associated with physiological changes. That is, listening to songs with lower combined arousal values predicted lower PR and SCL, indicating that X-system can be used to select songs that have a relaxing or energizing effect.

The results are in line with the results of the study by Nijman et al. (2023). We also found a significant reduction in physiological and subjective stress during listening to music but no accelerated reduction during the X-system playlist. Several explanations could explain the latter result. First, the equal effect of X-system and self-selected music despite a positive correlation between combined arousal values and physiology could indicate that participants themselves were able to adequately select music for stress reduction (although there was a great intra- and inter-personal variability). However, the mean combined arousal of self-selected songs was higher than the mean combined arousal of songs in X-system playlists. Second, the absence of larger stress reducing effects of X-system playlists compared to self-selected music might be explained by participants being insufficiently compliant to the X-system condition. Participants often did not listen to the complete X-system playlist and the largest decrease was expected in the second part of the playlist, because playlists consisted of songs with combined arousal values ranging from high to low. Third, although the X-system playlists were created together with the participants based on their preferred songs, it limited participants' choice of music at the moment. Possibly, the songs in the playlist did not always meet the participants mood and preference at that specific moment, which could have negatively affected stress reduction. Participants indeed seemed to prefer self-selected music over the X-system playlists: they listened more often to self-select music than to X-system music, although they were instructed to alternate self-selected music and the X-system playlist. This could also be related to the possibility that participants were less familiar with all songs in their X-system playlist because the X-system songs needed to comply a specific range of combined arousal value.

The effect of music within participants differed greatly between sessions. Possibly, the participants played different types of music in these sessions, which could explain the variability between sessions within participants. Another possible explanation is related to the level of stress at the start of the music listening session. Participants did not always listen to music directly after the notification of possible elevated stress level and their stress level might already have decreased at the time of music listening resulting in less stress reduction during music listening.

Although stress reduction was observed after listening to music, it should be noted that two participants dropped out on the first day due to lack of motivation and increase in stress related to participating. The latter was mainly explained by the use of wearables and not the music listening itself. However, five participants never listened to music longer than 10 min. This could indicate that music listening was not effective for all participants, it was not always possible to listen longer than 10 min to music or perhaps more support from caregivers through monitoring and direct instruction is needed to apply this strategy for stress reduction.

Strengths and limitations

This study is the first study that tested X-system in daily life in people with MID-BIF. However, due to the small sample and the experienced challenges with the implementation of the intervention (related to the naturalistic setting), the findings are preliminary and should be interpreted with caution. Moreover, the effect sizes, reflecting the proportion of variance explained for music over time, when controlling for the other predictors (main effects of time and music), were small. Changes in physiology could also be influenced by changes in movement intensity (participants may have sat more during music listening) or participants might have withdrawn from the context to listen to music, which could have reduced contextual stressors or demands. Moreover, several participants used psychotropic medication (regular or 'as required') that can influence the autonomic nervous system. It was unknown when medication was taken and if and how this may have influenced physiology during listening to music during the study period. Despite these limitations, this study included multi-modal assessment (physiology, ecological momentary assessment, and post-measurement through self-report and care-giver report) which all showed stress reduction when listening to music.

Conducting the study in the daily life of participants without continuous monitoring and direct instruction also led to challenges in compliance with the study protocol, resulting in reduced music listening sessions, missing physiological data and missing subjective ratings. When participants listened to music, they often did not alternate X-system and self-selected music across sessions correctly, regularly did not listened the complete X-system playlist and sometimes turned on shuffle which led to incorrect order of X-system songs. Moreover, participants regularly listened consecutive periods of both X-system and self-selected music *within* one session, which could have led to crossover effects. The results of the comparison of X-system music and self-selected music should therefore be interpreted with caution. Furthermore, because of the small sample size caution is needed when generalizing the results.

Finally, physiology measures, particularly SCL, were prone to artefacts and missing data. SCL often had values below 0.15 that were filtered out because they could indicate that the electrodes had not fully coupled with the skin tissue, or the device was not worn correctly. This problem has been observed in previous studies (Milstein & Gordon, 2020).

Future research

Based on the experiences and results of this research, the following suggestions for future research questions and methodological approaches are proposed. First, further research is required to evaluate music technology and artificial intelligence for selecting music, such as X-system. Software or a streaming application that automatically selects music for the participant with the correct order and number of songs could improve future assessments of the effectiveness of X-system. Moreover, future research could consider conducting a pilot study in another population, for example caregivers or people without intellectual disabilities and use this knowledge for developing a larger scale study in people with MID-BIF. Future research should control for movement and additional factors which could influence stress and physiology. To study these factors, observational information or information from structured interviews with participants or caregivers on factors which changed (for example the amount of stressful stimuli or demands) during music listening could be included. To investigate if changes in physiology during music listening can be attributed to listening to music future research could add a third condition where participants do not listen to music but receive care as usual which would function as a sort of control condition. Furthermore, the development of wearables for validly measuring SCL in naturalistic settings is needed.

Second, future research can explore how the effect of music listening on stress reduction can be optimized. For instance, studies can explore the possibilities of implementing technology for selecting music while retaining the experience of control by participants and variation in music, for example testing a dynamic system with the ability to change the X-system music every occasion. This could possibly improve the experience of participants, adherence to the study protocol and therewith stress reduction.

Third, future research can investigate for whom listening to music is effective and what causes drop-out. Qualitative methods (e.g., interviews, focus groups) are therefore needed to gain a better understanding in the experiences of participants with music listening with the aim of stress-reduction. Finally, future research can investigate potential explanations for the significant variation in effectiveness across different listening sessions. It can be explored whether type of music or stress-level at the start of a music listening session partly explain this variation.

Implications

The findings suggest that listening to music can be used by adults with MID-BIF as a strategy for the reduction of physiological and subjective stress. It is low-risk, low-cost (Kim & Stegemann, 2016) and clients could use it at various moments and places. It could be for example included in the individual treatment and intervention plan to reduce stress in specific situations, or at specific levels of stress. Self-selected music and X-system playlists were both associated with stress reduction and could both be used in clinical practice. However, combined arousal values were positively associated with physiological indices and therefore X-system could be used to make more informed choices about the selection of music together with the client. This can be especially relevant when client has difficulty selecting music for stress reduction, caregivers sometimes report concerns regarding the type of music participants select (e.g. hard rock). However, the findings suggest that it is important to consider sufficient control for the client, variety in music, and familiarity of the client with the music when using X-system.

What this paper adds

To our knowledge, this study is the first to examine changes in physiological stress during listening to music based on an innovative music selection system in people with MID-BIF in a naturalistic setting. Because of the naturalistic setting, not all factors which could have influenced physiology could be controlled for, therefore no casual attributions can be made about the relation between listening to music and physiology, and results should be considered preliminary. Suggestions for follow up research for improving the understanding of this relationship are given. Notwithstanding the reported limitations and exploratory nature of the study, its results suggests that music listening is associated with subjective and physiological stress-reduction in people with MID-BIF. X-system playlists were not more effective than self-selected music. However, PR and SCL decreased more after listening to songs with lower combined arousal values, indicating that X-system can be used to select songs that have a relaxing or energizing effect. Music listening could be an accessible, inexpensive and an empowering strategy which, through stress-reduction could positively influence emotion regulation, mental and physical health.

Appendix 5A

Explanation Combined arousal index

The X-System metric of 'Combined Arousal Index' uses a mathematically averaged combination of linear regression algorithms and a more recent neural network model. The linear regression algorithms are based on observations taken from large numbers of measured volunteer subjects and curators. The neural network model, which adjusts the algorithms in a non-linear fashion, is based on measured responses and edited by comparison with the linear method.

Table 5

Evaluation of the X-system Playlist by Participants and Caregivers (n = 10).

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Items participants					
It was fun using the playlist	0	1	0	5	4
I found it complicated to use	6	3	1	0	0
Listening to the playlist helped me relax	1	0	1	4	4
I would like to use this for relaxation in the future	2	0	1	5	2
Item caregivers					
Listening to the playlist helped the client to relax	0	1	0	7	2

Note. Number of scores for each answer category per item.

Chapter 6

General discussion

SUMMARY OF MAIN FINDINGS

Aggressive behaviour and self-harm are common in people with mild intellectual disability or borderline intellectual functioning (MID-BIF) in residential (treatment) facilities and have a negative impact on both the client, other clients as well as caregivers. Previous research showed that aggressive behaviour and self-harm often co-occur, a phenomenon known as dual-harm. However, at the start of the current thesis little was known on the co-occurrence and overlapping risk factors in adults with MID-BIF. In **Chapter 2**, we therefore first studied the association and shared risk factors between aggressive behaviour and non-suicidal self-injury (NSSI) in 125 adults with MID-BIF using incident reports and standard clinical measurements. Results showed a weak correlation between aggressive behaviour and NSSI, between impulsivity and total aggressive behaviour, as well as between poorer coping skills and aggressive behaviour and NSSI. NSSI, impulsivity and coping skills did, however, not predict aggressive behaviour. The latter is not in line with previous studies in this area, which could be explained by differences between people with and without intellectual disability, or the use of different measurement instruments compared to previous studies. These results also suggest it is important to assess other overlapping risk factors of aggressive behaviour and self-harm, besides coping and impulsivity.

One such overlapping risk factor is arousal. Altered (increased or decreased) arousal has repeatedly been linked to aggressive behaviour and self-harm (e.g. De Looff, Cornet, et al., 2022b; Nock, 2010; Weinberg & Klonsky, 2012). Over the past decades, an increasing number of studies were published on physiological arousal (e.g. hear rate; HR) in relation to aggressive behaviour and self-harm. Multiple systematic reviews have examined the literature on the relationship between aggressive behaviour and arousal. Remarkably, no comprehensive overview of studies on the relation between self-harm and arousal existed yet. In **Chapter 3**, we therefore conducted a systematic review of studies on the association between physiological parameters and self-harm, including 46 studies with diverse research designs, participants and measurements. Results showed that most studies which compared EDA and HR parameters in people with and without self-harm, found no clear indications for a relation between physiology and self-harm. Interestingly, one finding was that studies on HRV showed indications for lower HRV during recovery in people with self-harm. These studies indicated preserved arousal and suggested that this could reflect a poor ability to regulate physiologically following a painful or stressful experience in people who self-harm. Previous research found that lower HRV is associated to difficulties in emotion regulation (e.g., see the Integrated model of autonomic functioning of Beauchaine, 2001; Beauchaine, 2015). Imagery studies reported changes in HR and EDA across different stages of self-harm imagery scripts with an increased HR between scene setting and approach stages, and a decrease in HR and EDA during and after actual self-harm was imagined. The authors suggested that a decrease in physiological arousal can be considered as a (negative intrapersonal) reinforcer for self-harm behaviour. However, it should be noted that physiological arousal is influenced by many factors as described by Everly and Lating Jr (2019). No consistent findings were found when physiology was studied before, during or after actual occurrences of self-harm, but this was examined in very few studies. Although wearable technology has improved substantially in recent years, results showed that the majority of studies to date are lab-studies. Future research on the relation between self-harm and physiology need to be done with wearables in real-life (Johnson & Picard, 2020).

The overwhelming majority of physiological lab studies compared to real-life studies is also evident in aggressive behaviour. To our knowledge, no research existed on physiological trajectories measured in daily life following aggressive incidents in (forensic) inpatient treatment facilities. Understanding the mechanisms of increasing and decreasing arousal and gaining insight in the course of physiological arousal after aggressive behaviour could provide important knowledge for the prevention of recurrent aggressive behaviour, because a decreased ability to attenuate arousal might result in a higher probability for recurrent aggressive behaviour. In **Chapter 4**, we therefore studied changes in physiological arousal in the 30 minutes following aggressive behaviour, which was studied in a naturalistic settings with a large multicentre study. We analysed 165 incidents in which HR, SCL and SCR were measured with wearable technology. Primary analyses showed that, on average, all three parameters decreased over time, as hypothesized. However, the decreases were relatively small and showed high interpersonal and intrapersonal variability. Secondary latent growth curve analysis identified distinct subgroups of incidents with varying patterns of change in physiological parameters over time. In addition to decreasing trajectories, subgroups with stable and even increasing trajectories were identified, highlighting the heterogeneity of physiological reactions following aggressive behaviour. Interestingly, the type of aggressive behaviour did not predict the overall average decrease in the physiological parameters, but was significantly associated with the differential latent classes that an incident belonged to.

The identification of distinct trajectories of physiological arousal after aggression emphasizes the need for individualized approaches to managing aggressive incidents. For instance, if an individual's arousal has already decreased and they have returned to a calmer state, introducing additional interventions may be unnecessary or even counterproductive. On the other hand, individuals who continue to experience heightened arousal after aggression may benefit from targeted interventions, such as relaxation exercises or biofeedback. However, intervention decisions should not rely solely on physiological data but must complement the clinical judgment of support staff members.

In **Chapter 5**, we studied subjective and physiological arousal during a biofeedback and music listening intervention for stress reduction. X-system predicts the effect of music on autonomic arousal and can be used to select and order songs to construct a playlist based on preferred songs of the user, that is optimized for relaxation (or activation, depending on the needs of the listener). We studied whether using X-system for selecting music was associated with larger reductions in subjective and physiological arousal compared to self-selected music. To investigate this during stressful moments of participants, the study was conducted in real-life and biofeedback was used as participants often have difficulties in recognizing stress in themselves (i.e., interoceptive awareness). Participants were instructed to listen to music as soon as they received biofeedback on their increased stress level (as measured by a wearable biosensor called Nowatch) or when they themselves felt stressed. We collected 103 music listening sessions over a period of 2-4 weeks for 11 participants throughout their daily routines. Participants listened either to self-selected music or to a personalised playlist compiled with X-system. Results showed reductions in HR and SCL during listening to music, and indications were found for reductions in subjective stress and improvement of mood after music listening. However, the reductions were small, and physiology varied substantially both between and within participants during music listening. Listening to music compiled with X-system was not more effective than listening to self-selected music but lower combined arousal values based on a feature of X-system, predicted lower HR and SCL. This indicates that this feature can be used to select songs that have a relaxing effect. These preliminary results indicated that music listening is associated with both subjective and physiological stress reduction. Listening to music might be an accessible, inexpensive and empowering strategy for stress reduction and improving emotion regulation, which could also benefit mental and physical health.

REFLECTIONS AND DIRECTIONS FOR FUTURE RESEARCH

In the following paragraphs I will discuss two overarching key findings of this thesis.

Key finding 1: Physiological differences between self-harm and aggressive behaviour

Across the studies in this thesis and in the literature, we found indications for specificity in physiological characteristics of self-harm and aggressive behaviour. For self-harm the main, albeit preliminary, finding was that there were differences in HRV (lower at rest and during recovery) in persons with a history of self-harm compared to controls. In contrast, previous studies on aggressive behaviour identified lower resting HR as the most consistent physiological predictor of aggressive behaviour (De Looff, Cornet, et al., 2022b). Furthermore, our multicentre study conducted in clinical practice found decreases in EDA following aggressive behaviour, but stable or increases in EDA following self-harm. These findings could indicate that physiological trajectories towards and following self-harm and aggressive behaviour may be distinct.

In the following paragraphs, I provide several definitions of aggressive behaviour, physiological associations of self-harm and aggressive behaviour, and highlight overlap and distinctions in physiology between the two constructs. To further grasp the nuanced findings, I describe the heterogeneity of studies on the association between physiology, self-harm and aggressive behaviour. I argue that the distinct patterns may not hold when examined in the context of study-specific features. Finally, I conclude with directions for future research.

Definitions of aggressive behaviour

As mentioned in the general introduction of this thesis, several definitions of aggressive behaviour are available of which some include self-harm while others do not. The National Collaborating Centre for Mental Health (2015) defines aggressive behaviour as "a range of behaviours or actions that can result in harm, hurt or injury to another person, regardless of whether the violence or aggression is physically or verbally expressed, physical harm is sustained or the intention is clear" (p.6). Based on this definition, self-harm and aggressive behaviour are two distinct types of behaviour. On the other hand, for example, Bjorkqvist and Niemela (1992) defined aggressive behaviour as ''an act done with the intention to harm another person, oneself, or an object'' (p.4). Violence, a subtype of aggressive behaviour, is defined by the World Health Organization as follows: "the intentional use of physical force or power, threatened or actual, against oneself, another person, or against a group or community, that either results in or has a high likelihood of resulting in injury, dead, psychological harm, maldevelopment, or deprivation" (p. 1084; Krug et al., 2002). Similar to the latter definitions, the Modified Overt Aggression Scale+ (MOAS+) includes self-harm as a subtype of aggressive behaviour (Drieschner et al., 2013; Kay et al., 1988). The MOAS+ was used in Chapter 4 as an instrument to measure the frequency and severity of aggressive behaviour and includes auto-aggression (self-harm) as a subscale.

The provided definitions and distinction between self-harm and aggressive behaviour raises the question if there might be additional evidence that the constructs overlap or are distinct, and if yes, to what extent. The supposed overlap between aggressive behaviour towards others and aggressive behaviour towards oneself is for example further demonstrated by the frequently reported co-occurrence of these behaviours, which is known as dual-harm (Slade et al., 2020). The cognitive-emotional model of dual-harm states that the same pathways can result in aggressive behaviour and self-harm. The type of behaviour (i.e., aggressive behaviour versus self-harm) is influenced by the person's social context and his or her expectancies regarding the consequences of the behaviour. In Chapter 2, we investigated the co-occurrence between self-harm (NSSI) and aggressive behaviour in persons with MID-BIF. We found a significant but weak correlation, and self-harm (NSSI) did not predict physical aggressive behaviour. This contrasts with earlier studies on the association as they reported a positive association between self-harm and physical aggressive behaviour (mainly in persons without intellectual disabilities). The preceding paragraphs, and the outcomes of the study reported in Chapter 2 illustrate that heterogeneous definitions exist and findings from earlier studies are not equivocal. It is therefore useful to compare physiology related to aggressive behaviour and self-harm, which will be discussed in the following paragraphs. This knowledge could also contribute to assessment and treatment of these harmful behaviours.

Physiological associations with self-harm and aggressive behaviour

In Chapter 3, we systematically reviewed the literature on the association between self-harm and physiology. Results showed that the majority of studies in which participants with a history of self-harm were compared with participants without a history of self-harm, found no differences in physiology. However, when comparing the different physiological parameters (HR, HRV, and EDA), results showed more pronounced differences in HRV than in HR and EDA within the self-harm group compared to controls. This is consistent with findings from meta-analyses by Bellato et al. (2023) and Goreis et al. (2023). Several studies included in our systematic review and the meta-analysis by Goreis et al. (2023) reported lower HRV at recovery. This indicates a slower recovery to baseline levels after a painful or stressful experience in people who self-harm. Previous research found that lower HRV is associated with difficulties in emotions regulation (e.g., Beauchaine, 2015). These relatively longer recovery periods of neurophysiological responses may contribute to maladaptive strategies for stress reduction (Everly & Lating Jr, 2019), such as self-harm. It is interesting that we also found evidence for stable and increasing physiological (EDA) trajectories, and thus lower recovery following self-harm in the multicentre study. This seems to support these earlier lab findings, but should be further studied because only few self-harm incidents were observed in this study. Besides lower recovery HRV, Goreis et al. (2023) reported lower HRV in rest. These results suggest that compared to controls, HRV indices are different in people who self-harm in the sense that they showed lower HRV at rest and longer HRV recovery times following a stressor. However, there was considerable heterogeneity in the results. Not all studies included in our systematic review could replicate the finding by Goreis et al. (2023) of different HRV in people who self-harm compared to controls. Moreover, studies that did report differences in HRV in rest reported both lower and higher HRV in rest in the self-harm group compared to controls.

Regarding aggressive behaviour, a meta-analysis that included 11 studies reporting on HRV in rest found that aggressive behaviour was associated with decreased, but non-significant, HRV in rest *(De Looff, Cornet, et al., 2022b)*. Results regarding HR are, however, more consistent. HR in rest has repeatedly been found to be the best physiological predictor of aggressive behaviour (De Looff, Cornet, et al., 2022b; Lorber, 2004; Ortiz & Raine, 2004; Portnoy & Farrington, 2015), although it should be noted that effect sizes are typically small. The largest effect sizes for low resting HR were found among the more serious antisocial groups, such as serious offenders and subjects with psychopathic traits (De Looff, Cornet, et al., 2022b). This might be interpreted as support for the low-arousal theory, which states that antisocial behaviour is related to a low ANS activity, which might increase the risk for sensation-seeking behaviour and fearlessness (Jansen, 2022). However, not all studies found lower resting HR to be associated with aggressive behaviour (Oldenhof et al., 2019; Zijlmans et al., 2021). One possible explanation lies in the heterogeneity in types of aggressive behaviour. For instance, individuals with high callous-unemotional (CU) traits have low arousal levels, as evident from a low HR and EDA (e.g. in response to emotionally evocative films, when anticipating aversive stimuli, viewing others in pain or, peer provocation; Frick et al., 2014). CU traits are characterized by lack of empathy, guilt or remorse, and a shallow or deficient affect, while individuals with higher levels of anxiety and reactive aggressive behaviour have been reported to show higher arousal levels (Jansen, 2022).

To summarise, previous meta-analyses on self-harm found lower HRV (rest and recovery). Our systematic review also found indications for lower HRV at recovery in participants with self-harm compared to controls, whereas studies on aggressive behaviour have predominantly shown lower HR levels at rest. HR is associated with dually innervated SNS and PNS activity and HRV (in terms of the reported RSA and RMSSD) is primarily associated with PNS activation. These differences in physiological parameters could indicate that self-harm and aggressive behaviour have different associations with ANS functioning, highlighting their distinctive characters. This knowledge might also be useful for assessment of these behaviours.

Heterogeneity of studies and directions for future research

The differences in physiological parameters typically linked to self-harm and aggressive behaviour (i.e., HRV and HR, respectively) may also stem from study characteristics, such as study design (e.g., comparison and laboratory versus naturalistic), operationalization of the target behaviour, and the number of studies conducted.

First, if we dive deeper into the two recent meta-analyses on self-harm (Bellato et al., 2023; Goreis et al., 2023), our systematic review on self-harm (see Chapter 3) and the meta-analyses on aggressive behaviour, it became evident that the findings differ across the specific comparisons that were made. Studies on self-harm using a between-subject comparison (mostly comparing a group with self-harm to a group without self-harm) found indications of decreased HRV in the self-harm group. In contrast, multiple imagery studies using a within-subject comparison design show other parameters, i.e. EDA and HR, changing across stages of personalised imagery scripts in self-harm groups. Within-subject comparison studies into physiology before, during and after actual occurrences of self-harm found contradictory results (decreases, increases in HR, EDA as well as no changes were reported before, during and after self-harm), but the number of this type of studies is small. Thus, although between-subject comparison studies show a general lower HRV related to self-harm and lower HR related to aggressive behaviour, within-subject comparison studies paint a heterogeneous picture. Actual (and imaged) self-harm and actual aggressive behaviour may not be (solely) related to lower HRV and HR, respectively, but also to differences in other physiological parameters. There seems to be a clear need for real-life studies that investigate physiology before, during and after actual occurrences of aggressive behaviour and self-harm.

Second, differences in physiological parameters related to self-harm versus aggressive behaviour might be related to operationalization of the target behaviour or subtypes of the behaviour. In Chapter 4 we explored the variability in physiological arousal following aggressive incidents by assessing whether (latent) classes could be identified which follow similar patterns of physiological arousal. Results showed distinct classes that varied in physiological parameters (HR, SCL and SCR) at the time of the incident (intercept), as well as the trajectories (slope) of physiological parameters over time following the incident. Type of aggressive behaviour (e.g. aggressive behaviour towards others and towards oneself) predicted the class that an incident belonged to. Verbal, physical and aggressive behaviour towards objects had a higher probability to belong to classes with a decreasing EDA slope, whereas self-harm more often belonged to classes with a stable EDA slope. With regard to HR, classes showed a different distribution across types of aggressive behaviour (see Chapter 4). However, these results need to be interpreted with caution, as only few self-harm incidents were observed, which could be due to the fact that self-harm is often underreported (Hulsmans et al., 2024; Preyde et al., 2012). Regardless of the limitation we were, however, able to conduct a real-life multicentre study into physiological changes in both aggressive behaviour and self-harm.

Third, the number of studies conducted on the different physiological parameters in relation to self-harm versus aggressive behaviour should be considered. HRV is less often studied than HR in relation to aggressive behaviour. From the four meta-analyses on physiology related to aggressive behaviour (Lorber, 2004; Ortiz & Raine, 2004; Portnoy & Farrington, 2015), only the meta-analysis by De Looff, Cornet, et al. (2022b) included HRV. Moreover, these HRV indices represented only a small proportion of studies compared to the number of studies on HR. The observed difference in physiology in aggressive behaviour compared to self-harm (HR, HRV) may thus partly be explained by the number of studies, study design, and other study characteristics. More research into HRV and more studies with idiographic methods are necessary to unravel these complex interactions. Last but not least, a meta-analysis in which physiology of aggressive behaviour and self-harm are compared directly could improve our insight in (differential) physiological characteristics (see Nijman et al., in preparation).

Key finding 2: Small effects and high variability: call for in-depth idiographic research with multi-source integration

A second overarching finding of studies in this thesis is that groupwise research designs on (only) physiological data yield contradictory results, and relatively limited implications for assessment and treatment of aggressive behaviour and self-harm in clinical practice. First, I will explain why solely groupwise comparisons are insufficient. Second, I will explain why additional idiographic approaches are needed. Third, I will explain the importance of multi-source integration and describe what type of information needs to be included in future research to improve clinical usefulness of physiological arousal measures.

Groupwise comparisons resulted in several contradictory results with limited clinical implications. First, in our systematic review on physiological parameters associated with self-harm (see Chapter 3), we found that most studies failed to find a significant association between physiological arousal and self-harm. Second, studies that did find alterations in physiology in the self-harm group found both increased and decreased physiological arousal. Third, in Chapter 5 we found a large variation between and within participants' physiological arousal during music listening, though results indicated that there was a significant but small decrease in physiological arousal. Fourth, in Chapter 4, we investigated changes in physiological arousal following aggressive incidents. The main analysis showed again significant but small decreases in physiological arousal with a large variation. The need for idiographic approaches was evident from the secondary analysis. We found that physiological arousal followed distinct trajectories after aggressive incidents and different latent classes could be distinguished (four to six, depending on the physiological parameter). These latent classes varied in the mean value of physiological arousal at the time of an aggressive incident as well as in the trajectories of physiological arousal over time following an incident. This subgroup analysis provided valuable insight in the different trajectories following aggressive behaviour. Besides a decrease in physiological arousal during the 30 minutes following an aggressive incident, also clusters of incidents (classes) with stable trajectories and even trajectories with increasing physiological arousal were found. These different trajectories could explain the small average (i.e., groupwise) decrease in physiological arousal. Furthermore, results showed that the type of aggressive behaviour predicted the trajectory of physiological arousal following the aggressive incident. Thus, the main groupwise analyses of these studies showed contradictory outcomes or small effect sizes with large variability.

Notably, the subgroup analysis (Chapter 4) improved our understanding of the variety in physiological arousal following aggressive behaviour.

However, randomized controlled trials (RCT's), a groupwise study design, are considered the golden standard method for studying effectiveness because of the high internal validity (Tomlinson et al., 2015). RCT's alone might not be sufficient because of several reasons. First, studies such as RCT's that seek causal factors groupwise may not generate findings that apply to the individual client (Fedor et al., 2023). Second, data analytics that are based on identifying patterns across people will have to account for many parameters and thus require very large samples with long time series and high-quality data. Third, if these requirements are met, groupwise comparison designs are still potentially problematic as they may average out or ignore meaningful individual differences, as was seen in the main analysis of the study presented in Chapter 4. Fourth, knowledge from large (RCT) studies takes a long time to be disseminated and implemented in clinics. The arguments presented above give rise to a call for within subject and preferably idiographic approaches.

Idiographic approaches

Aggressive behaviour, self-harm and physiological arousal are complex phenomena, since they are dynamic and multi-causal, and the mechanisms manifest idiosyncratically (Hulsmans et al., 2024). Several authors argue that individualizing prevention and treatment require in-depth idiographic approaches to disentangle complexity and generating new insights relevant for clinical practice (Fedor et al., 2023; Hekler et al., 2019; Hulsmans et al., 2024). Single case experimental designs (SCED's), latent class analyses (such as in Chapter 4), complex systems analyses (e.g. Hasselman, 2023), multilevel modelling (Chapter 4), and ecological momentary assessment (such as in Chapter 5) seem to be useful. This might also improve the implementation of biological information in the assessment and treatment of aggressive behaviour and self-harm in persons with (and without) MID-BIF. Several participants (professionals working in forensic mental healthcare and experts on technology in forensic mental healthcare) in the study by Kip et al. (2020) indicated that for example SCED's might be more suitable than RCT's to study eHealth in forensic mental health care from a practical and methodological point of view. The authors state that more such studies are needed.

Another advantage of idiographic approaches is that knowledge from idiographic approaches can rapidly be used to improve the care for an individual in the moment, or shortly thereafter, whereas RCT's don't have this advantage (Hekler et al., 2019; Staddon, 2001). This was illustrated by Fedor et al. (2023) who described a single case in which a patient with a depressive disorder uses several sources of information from a wearable to track the progression of the therapy and discuss valuable insights with the therapist. Because idiographic approaches improve the possibilities of giving the client (and the caregiver or therapist) valuable personalised feedback, they may also improve motivation and compliance for participating in scientific studies.

New data, such as collected with wearables, advance individualisation of prevention and treatment (Hekler et al., 2019). Wearables can be valuable measurement instruments for studies with idiographic approaches as they enable high-frequency, longitudinal monitoring of individual clients. These data can be for example used to evaluate treatment effectiveness through *N*=1 designs (Fedor et al., 2023).

Knowledge retrieved from *N*=1 studies can not only help that specific individual but, through the identification of clusters across people and change mechanisms, generalizable insights can be generated as well (Hekler et al., 2019). Moreover, groupwise and idiographic approaches can be used complementary. The studies in this thesis followed predominantly a groupwise approach, although the studies in Chapter 4 and 5 followed to some extent more idiographic approaches. As described by Hekler et al. (2019) these studies are excellent for providing insights for 'a warm start' in this study field/topic and providing important directions for future research. In contrast, idiographic approaches are useful for moving forward towards an increasingly more individualised understanding, of in this case, physiological patterns related to aggressive behaviour and self-harm. Also, Heirbaut et al. (2024) underscore the importance of idiographic approaches (SCED's) at this stage and state that a large RCT is more appropriate once the use of wearables is widespread across multiple facilities.

Multi-source integration: gaining insight in variability in physiology by including non-physiological data

Future research using idiographic approaches can further improve our understanding of physiological arousal in relation to self-harm and aggressive behaviour by also including non-physiological data. Arousal and physiological parameters are influenced by many factors and thus interpretation is not always straightforward. Including information on the type, function, and severity of the target behaviour, as well as contextual factors and subjective experiences, could enhance the clinical usefulness of physiological arousal measures. Below, I elaborate on this topic and explain why including this information is important to understand (variation in) physiology in relation to self-harm and aggressive behaviour.

Type, function and severity of behaviour

Future research into the association between self-harm and physiological arousal should specify information on the type, function and severity of self-harm. The null results (especially for HR and EDA) and contrasting results of the study presented in Chapter 3 may not be caused by an absence of an association between physiological arousal and self-harm. Alternatively, it could possibly be attributed to the levels of severity and the various types of self-harm, which likely served different behavioural functions, that were included in most studies. Different functions of self-harm may be characterised by different (contrasting) physiological arousal profiles. For instance, downregulation of high arousal and upregulation of low arousal presumably result in different physiological patterns. Similarly, self-harm as a means to regulate arousal or self-harm to regulate interpersonal situations might result in different physiological profiles (Nock, 2010). For instance, Nock and Mendes (2008) found that persons who reported engaging in self-harm to escape hyperarousal experienced the strongest physiological arousal during a distressing task. Indications that it is advisable to distinguish between types of self-harm also comes from Bellato et al. (2023) who found no association between EDA and self-harm when all types of self-harm were considered together. However, specifically non-suicidal self-harm showed increased EDA, while studies that did not distinguish between different types of self-harm (i.e. suicidal ideation, suicidal self-harm, and non-suicidal self-harm) reported reduced EDA.

Different types of aggressive behaviour are also associated to differences in physiological parameters, as reported by the meta-analysis of *De Looff, Cornet, et al. (2022b)*. In Chapter 4, we included information on the type of aggressive incident, categorised by topography (e.g. verbal or physical aggressive behaviour). However, no information was available on type of aggressive behaviour, categorised by function, for example proactive (i.e., goal-directed, instrumental, unemotional) versus reactive (i.e., emotional, impulsive) aggressive behaviour. Information on the behavioural function, measured with validated questionnaires, of aggressive behaviour might help understand the different physiological trajectories (e.g. decreasing, stable and increasing) we found following aggressive incidents (Blankenstein et al., 2022; Hubbard et al., 2002; Jansen, 2022; Murray-Close et al., 2017; Schoorl et al., 2016).

Furthermore, including severity scores of self-harm or aggression incidents might help understanding the variability of trajectories of physiological arousal within a type of behaviour (e.g. self-harm) which was reported in Chapter 4. For example, Wells et al. (1999) found different trajectories in HR between stages of self-cutting and skin-picking imagery scripts. Moreover, De Looff et al.'s meta-analysis on the association between physiological parameters and aggressive behaviour found the largest effect sizes for low resting HR among the most violent offenders. It is therefore essential to include information on these factors in the analyses, preferably in addition to person specific contextual information.

Contextual information

To enhance our understanding of different trajectories in physiological arousal following (preceding and during) aggressive behaviour future research could also include contextual information (De Looff, Duursma, et al., 2022; Dunn et al., 2025; Fedor et al., 2023). In Chapter 4 we studied changes in physiological arousal after aggressive incidents however, it was not clear what caused these changes. Future research could collect data on the response of caregivers following the incidents, and study the effect of different responses on the physiological trajectory following incidents. For instance, coercive or restrictive measures are often perceived as stressful by the client (Chieze et al., 2019), and may impact physiological responses after incidents. Research including this type of contextual information can also explore the effectiveness of different management strategies for the reduction of physiological arousal. This knowledge might guide professionals in selecting personalised interventions given the context in which the behaviour occurs.

Adding contextual information could also improve our understanding of variation in physiological stress reduction within and between subjects during music listening. Changes in physiology during music listening could be influenced by contextual changes. For example, if a participant withdrew from the context to listen to music (e.g. by going to a comfort room or retreating in his room), this could have reduced contextual stressors or demands and therewith physiological arousal.

Psychosocial, psychological and subjective experiences

Including psychosocial information together with physiological data could also improve the prediction of self-harm. This is illustrated by the study of Nelson et al. (2023), who found that resting HR was not associated with self-harm in main effects analyses, but the interaction between peer conflict and resting HR at baseline significantly predicted the frequency of self-harm at follow-up. Besides psychosocial information, psychological and subjective experiences are important to consider when studying physiology. The following study illustrates the importance of psychological context in the interpretation of physiological arousal. Contrary to their expectations, Tutunji et al. (2023) found decreased physiological arousal during an examination week in medical students. However, it was found that this relationship was mediated by the students' affect. Using machine learning, Tutunji et al. (2023) tested whether wearables and/or ecological momentary assessment (EMA) could best classify whether data stemmed from examination or control weeks. The combination of both types of measures resulted in the best classification, illustrating the importance of combining physiological data with psychological data from for example self-report.

In our biofeedback study on (physiological) stress reduction during music listening (see Chapter 5) we also combined physiological data with self-report (EMA) data. We measured subjective experiences with a slider before and after music listening, a form of EMA, and a short questionnaire after the intervention period. Results of the subjective measures helped to interpret the physiological outcomes and supported the conclusion of stress reduction. However, there were many missing data from the EMA sliders. By contrast, Hulsmans et al. (2023) studied the feasibility of daily monitoring in adolescents and young adults with MID-BIF and concluded that EMA is feasible in outpatient and residential care.

A possible reason why these authors on average reached good compliance is that the diary entries were used for feedback in treatment. Possibly, in our study participants did not fill in the sliders because of insufficient (direct) feedback on their answers and insufficient imbedding of these measures in their treatment program. Lack of compliance might also be related to involuntary care. Hulsmans et al. (2023) reported low compliance (19%) in juvenile detention setting, data of the study described in Chapter 4 was partially collected in forensic residential care facilities. In addition to or as an alternative to using a slider to measure subjective experiences, future research could include more in-depth interviews to explore participants' experiences. This might improve our understanding of the large variation we found in physiological stress reduction within and between subjects during music listening. More information on the subjective experiences of the participants could also provide knowledge for whom and when biofeedback and music listening could be effectively used for stress reduction and how it can be improved. This is important because, besides variation in stress-reduction, we observed large differences between participants in adherence to the research protocol and the need for support with participating in the intervention (using the biosensors and listening to music). The need for more research to better understand why and for whom wearable technology works was also reported by Heirbaut et al. (2024) and Kip et al. (2020).

Notwithstanding the importance of combining physiological and self-report data, this can also introduce new complexities. For instance, interpreting contrasting physiological and self-report data is challenging. For example, in Chapter 5 participants frequently reported discrepancies between self-experienced stress and the notifications on increased stress levels (both false positives and false negatives) which they received. It is important to note that available stress algorithms are mainly tested in laboratory settings, with non-clinical populations (e.g. Westerink et al., 2020). Also, no golden standard, as with HR (i.e. electrocardiography), exists to validate stress-algorithms. To evaluate physiological stress algorithms, we therefore depend on for example self-report data on stress. However, using self-report data to assess validity of stress algorithms is challenging in the population which could benefit most from biofeedback on stress, that is persons with limited interoceptive awareness *(see also, Bellemans et al., 2019; Heirbaut et al., 2024).*

Combining physiological parameters

Lastly, besides mixed methods combining quantitative physiological arousal and (qualitative) psychological (self-report and/or interview) data a combination of physiological parameters (e.g. EDA and HR) could improve insight into aggressive behaviour, self-harm and stress levels. Results of the study by Goodwin et al. (2019) with persons with an autism spectrum disorder and of whom most had a comorbid intellectual disability suggest that aggressive behaviour to others can be predicted 1 minute before it occurs using 3 minutes of data collected with wearables including a combination of cardiovascular activity (blood volume pulse and interbeat interval), EDA, and motion. Imbiriba et al. (2023) conducted a comparable study in persons with an autism spectrum disorder (of whom almost half were diagnosed with intellectual disability) on self-harm and reported similar results. Also, research into detecting stress using machine learning with a combination of physiological parameters has achieved high accuracy in classifying stress levels (in patients with metabolic syndrom and children with autism spectrum disorders; Akbulut et al., 2020; Masino et al., 2019).

To summarise, results of the studies presented in this thesis showed contradictory outcomes or small effect sizes with large variability. More in-dept studies using idiographic approaches and multi-source integration are needed to move the field of physiological arousal forward and to improve the implementation of biological information in the assessment and treatment (of aggressive behaviour and self-harm) in mental health and disability care.

Potential, pitfalls and future directions of wearables in clinical practice

Potential

The implementation of wearables in the (forensic) care and treatment of individuals with intellectual disabilities with behavioural and mental health problems has many potential benefits. Besides the aforementioned benefits such as the possibilities for individualized approaches to managing aggressive incidents, the following benefits are interesting to mention. First, wearables can give the client and his caregiver or therapist insight in his arousal levels (besides movement and sleep), in addition to self-reported data. Wearables can bring important information from daily life into 'the treatment room' (Fedor et al., 2023; Heirbaut et al., 2024). For clients it can be difficult to report on their arousal levels and self-report is prone to various biases. It is known that it can be difficult to recall one's (physiological) state of previous days (e.g. Fedor et al., 2023). Physiological data collected with wearables can function as a prompt for a conversation between clients and their caregiver or therapist (Fedor et al., 2023; Heirbaut et al., 2024). Fedor et al. (2023) reported an example of a client with major depressive disorder who reported to her therapist that she had been sleeping great the past two weeks. However, the wearable measured a very irregular sleep pattern. When the therapist discussed this, the client recognised this and together they could further evaluate and explore the presence and severity of sleep problems. It should however be mentioned that the reliability and validity of wearables is still topic of research (van der Mee et al., 2021), and it is important for the clinician to understand what is measured and what is aggregated (see the section 'Pitfalls' for more details; Altini, 2024; Doherty et al., 2025). Besides bias in recall, reduced interoceptive abilities or social desirable answers, which are frequently observed in people with MID-BIF, can cause bias and reduced accuracy of self-report data (Havercamp et al., 2022). The addition of wearables requires less cognitive reflection and language proficiency (Heirbaut et al., 2024; Kip et al., 2020), which can be especially beneficial for persons with cognitive impairments. Second, besides bringing information from daily life into the treatment room, wearables can support in the transfer of the treatment from the treatment room to daily life, and therewith enhancing the empowerment of the client (Heirbaut et al., 2024).

Third, wearables can capture the attention of caregivers through (real-time) biofeedback. Caregivers may easily overlook non-verbal signs of rising arousal levels in real-time situations, due to high workload. Wearables can support them in monitoring tasks, making sure their caring expertise is focused on the client at relevant times (Noordzij & Laroy-Noordzij, 2015). Fourth, using wearables is a 'passive' way of measuring, which enables frequent and longitudinal measurements (Fedor et al., 2023), and which can be a valuable addition to more 'active' measuring methods. Active measuring methods can be prone to missing or biased data, as illustrated by the missing responses to the scale questions in Chapter 5. Finally, the apps used with wearables are potentially destigmatizing because apps are commonplace and socially accepted (Bucci et al., 2019).

Pitfalls

Besides the various potential benefits of using wearables in clinical practice, multiple pitfalls should be considered. First, there is the issue of validity of the data and their interpretation. Using wearable devices produces large amount of raw data which first need to be processed, 'cleaned' and screened for artefacts before they can be interpreted (De Looff, Duursma, et al., 2022; Fedor et al., 2023). Measuring physiological arousal with wearables is prone to artefacts. For example, regarding measuring EDA, invalid low values caused by insufficient coupling of the electrodes has been reported several times in the literature (e.g. Milstein & Gordon, 2020). In Chapter 4, one of the latent classes had a very low mean SCL which might suggest that the recorded data of this class were mostly noise. Some physiological parameters however can be measured more valid than others. Altini (2024) states that it is important to know which physiological parameters are measured and which are estimated. For example, HR(V) at rest can be measured validly with wearables (Schuurmans et al., 2020). However, stress is estimated with an estimation algorithm, and difficult to capture as no golden standard to measure physiological stress exists. In our studies, we experienced difficulties in measuring stress in a valid way using wearables: participants frequently reported discrepancies between self-experienced stress and the notifications they received on increased stress levels (both false positives and false negatives, Chapter 5). This has also been reported with different wearables (Heirbaut et al., 2024; Ter Harmsel, Noordzij, et al., 2023). Furthermore, some wearables do not (accurately) measure stress levels during physical activity or movement (e.g. walking or cycling), or activities which naturally elevate heart rate (like doing household chores). Thus, physiological measures during these activities can falsely appear as stress (Chapter 5; Heirbaut et al., 2024). Wearables should therefore be considered a tool to increase awareness of bodily changes, not as a normative measurement tool indicating stress. However, interpreting these biofeedback signals and adding nuance to this interpretation can be complicated for clients (Heirbaut et al., 2024). A factor that complicates using valid wearable technology is that technological developments go fast but scientific research is relatively slow. By the time that the validity of devices is properly tested, the devices are already outdated (Kip et al., 2020).

A second pitfall is the usability of wearable devices. De Looff et al. (2021) investigated the usability and acceptance of four wearable devices in forensic clients with MID-BIF. Results from qualitative questionnaires, in general, indicated that the devices were easy to use and gave clear information. However, a more recent study by De Looff et al. (2024) on the usability of four different wearable devices in forensic clients with quantitative measures reported that the devices did not have usability scores that would typically lead people to recommend that technology. Another study on the perspectives of forensic psychiatric clients and therapists on one of these wearable devices, using the same questionnaire as De Looff et al. (2024), found acceptable usability scores (Ter Harmsel, Smulders, et al., 2023), though also below the threshold typically needed for recommendation. Currently, there is no wearable device that fully meets all the requirements of forensic clinical practice (Heirbaut et al., 2024). However, sufficient support by caregivers can lead to successful implementation of wearables. For example, caregivers can address usability issues by configuring the wearable together with the client, answering questions about its use during sessions, and assisting the client in understanding the notifications (Heirbaut et al., 2024).

A third pitfall applies to the acceptance of technology by caregivers. Some caregivers report reluctance to move away from more traditional, face-to-face methods of care or report a lack of confidence in using technology (Bucci et al., 2019; Heirbaut et al., 2024). However, the majority of professionals working in forensic mental healthcare, current or former forensic psychiatric patients, and experts on technology in forensic mental healthcare, consider the use of wearables promising (Heirbaut et al., 2024; Kip et al., 2020).

A fourth pitfall is the lack of guidelines on how and when to use wearables in clinical and mental health care for persons with psychiatric disorders and/or intellectual disabilities (Kip et al., 2020). Besides guidelines, personalisation is very important to fit the different needs and characteristics of this diverse population (Heirbaut et al., 2024; Kip et al., 2020).

A final pitfall are the costs of wearables. The wearables used in the studies of this thesis cost for example between 600 and 2000 euro's. Currently, the use of wearables in mental health and intellectual disability care is not standard reimbursed and clients are often not able to afford to buy the technology (Bucci et al., 2019). Wearables might be reimbursed in the future, effectivity research showing positive effects of wearables in mental health and intellectual disability care can contribute to this.

Future directions of wearables in clinical practice

To improve clinical interpretability of large data outputs of wearables, tools making the data easily accessible are needed (Bucci et al., 2019). Visualization tools showing the longitudinal pattern of clinically relevant information should be developed (Dunn et al., 2025; Fedor et al., 2023). De Looff, Duursma, et al. (2022) developed a "Wearables" R package and a Shiny "E4 dashboard" application for the wearable Empatica E4. The package and Shiny application can be used to visualise the relationship between physiological signals and real-life stressors or stimuli, but can also be used to pre-process physiological data, detect artefacts, and extract relevant features for further analysis. The dashboard is currently being further developed and expanded for two other wearables (the EmbracePlus and the Nowatch). These visualisation tools can be used by the clinician to discuss the data with the client, which can improve the client's engagement in treatment (Fedor et al., 2023).

When using wearable devices in (forensic) mental health or disability care it is important to integrate it into treatment protocols or existing situations (Heirbaut et al., 2024; Kip et al., 2020; Ter Harmsel, Smulders, et al., 2023). For example, wearables can be added to an existing cognitive behavioural treatment program for reduction of aggressive behaviour, which has been studied by Heirbaut et al. (2024) and is currently studied by Nijman et al. (in preparation). For further recommendations of implementing wearable devices in the care for various populations (e.g. forensic adult clients or children with severe developmental problems) see Adam et al. (2024); Hagoort et al. (2025); Heirbaut et al. (2024); Kip et al. (2020); Ter Harmsel, Smulders, et al. (2023).

To improve the usability of wearable devices De Looff et al. (2024) suggest improvements in gamification and motivational aspects of wearable technology. Furthermore technical developments can improve the usability, such as improvement of connectivity and notifications, increase in possibilities of personalization of settings and improved battery life (Heirbaut et al., 2024; Ter Harmsel, Noordzij, et al., 2023). The development of wearables requires continuous development, customization and co-design with professionals and clients (Adam et al., 2024; Kip et al., 2020; Schouten et al., 2022; Woodward et al., 2023). Co-designing wearables with clients ensures a better alignment of the wearables with their needs and preferences and therewith improves the compliance when using the wearables.

With regard to validly measuring stress with wearables, in the past years and currently multiple studies are conducted (e.g., Immanuel et al., 2023), leading to the expectation that the quality of the algorithms utilised by these devices will significantly improve in the coming years. 'Stress in action' is a current large longitudinal multicentre project that aims to develop low-burden, ecologically valid ambulatory assessment toolkit to quantify stress in daily life (Schoenmakers et al., 2025).

Finally, the ethical, privacy and data security aspects of the use of wearables in disability and mental health care need to be considered (Adam et al., 2024; Almubairik & Khan, 2024; Fedor et al., 2023; Kip et al., 2020). This is important because wearables measure many health-related signals that are considered to be sensitive (Almubairik & Khan, 2024). In Chapter 5, we selected for example the Nowatch because of its good data policy, i.e. leaving the client the owner of the data and not saving their data unless requested otherwise. Giving clients control over their data may increase their rust (Ginsburg et al., 2024). However, several commercial wearables do save their users data or do not give their users access to the raw data. Clarity and guidelines on topics such as privacy, data security and ownership of data are needed (Ginsburg et al., 2024; Kip et al., 2020).

CONCLUSION

In the current thesis, the co-occurrence of non-suicidal self-injury (NSSI) and aggressive behaviour and overlapping risk factors in persons with MID-BIF was studied. It also provided a broad overview of the current literature on the association between self-harm and physiological parameters. Furthermore, in daily life situations, physiology following aggressive behaviour was investigated with a large multicentre study. The final chapter examined changes in subjective and physiological stress during music listening in daily life situations among individuals with MID-BIF, comparing the effects of self-selected music and music selected and ordered by an innovative music system.

Advancements in technology enable the continuous monitoring of physiological arousal. The use of wearables has many potential benefits for scientific research. More research measuring physiology with wearables before, during and after actual aggressive behaviour and self-harm in daily life situations is needed. Furthermore, more in-dept studies using idiographic approaches and multi-source integration are needed to move the field of physiological arousal forward and to improve the implementation of biological information in assessment and treatment (of aggressive behaviour and self-harm) in mental health, forensic and disability care.

Monitoring with wearables also has promising potential in clinical practice for unravelling complex issues, particularly in diagnosis and treatment. For example, physiological data collected with wearables can function as a prompt for a conversation between clients and clinicians. Or wearables can capture the attention of caregivers in residential settings, supporting them in monitoring tasks. Furthermore, insights into physiological arousal patterns, both prior to and following aggressive behaviour, may help to clarify the function of aggressive behaviour in clients. This understanding could be used to inform personalised treatment strategies aimed at addressing the underlying causes of aggressive behaviour. However, assessments and intervention decisions should not rely solely on physiological data but must complement the clinical judgment of caregivers and therapists. Currently we stand at the start of many new developments in this area, that are expected to progress rapidly. Guidelines on how and when to use wearables in (forensic) care for persons with psychiatric disorders and/or intellectual disabilities need to be developed. To improve clinical interpretability of large data outputs of wearables, tools making valid data easily accessible are needed. Moreover, with the use of wearables it is important that ethical, privacy and data security aspects are considered carefully.

Follow-up Jason

Jason participated in the biofeedback study into physiological and subjective stress reduction during music listening. He successfully completed the two weeks research period. Afterwards, his experiences with using wearable devices and music listening (with X-system) was evaluated with him and his therapist. Jason explained that before he found it challenging to recognise increased levels of arousal on his own. The wearable however offered him significant support in recognizing levels of arousal. He described moments when the device alerted him to increases in his arousal, even when he was unaware of experiencing stress. At other times, he did recognise feelings of stress and noted that the device supported his ability to better identify these moments. He mentioned an example where he walked away from a stressful social experience to his room and listened to music and returned after his stress had dissipated. The biofeedback thus led to a deeper awareness of his stress levels, something he had not focused on much in the past. Even after discontinuing using the wearable, he reported being better able to recognise and reflect on his stress patterns, something which was recognised by his therapist. No aggressive incidents occurred during the period he used the wearables. His therapist used the study results to further discuss and understand stressful situations and to plan strategies for future stress detection and reduction together with Jason.

References

Adam, E., Meiland, F., Frielink, N., Meinders, E., Smits, R., Embregts, P., & Smaling, H. (2024). User Requirements and Perceptions of a Sensor System for Early Stress Detection in People With Dementia and People With Intellectual Disability: Qualitative Study. *JMIR Formative Research*, *8*, e52248. https://doi.org/10.2196/52248

Agorastos, A., & Chrousos, G. P. (2022). The neuroendocrinology of stress: the stress-related continuum of chronic disease development. *Molecular Psychiatry*, *27*(1), 502-513. https://doi.org/10.1038/s41380-021-01224-9

Akbulut, F. P., Ikitimur, B., & Akan, A. (2020). Wearable sensor-based evaluation of psychosocial stress in patients with metabolic syndrome. *Artificial Intelligence in Medicine*, *104*, 101824. https://doi.org/10.1016/j.artmed.2020.101824

Aldrich, J. T., Wielgus, M. D., & Mezulis, A. H. (2018). Low physiological arousal and high impulsivity as predictors of self-injurious thoughts and behaviors among adolescents. *Journal of Adolescence*, *62*, 55-60. https://dx.doi.org/10.1016/j.adolescence.2017.11.006

Almubairik, N. A., & Khan, F. A. (2024). Systematic Literature Review on Wearable Digital Forensics: Acquisition Methods, Analysis Techniques, Tools, and Future Directions. *IEEE Internet of Things Journal*. https://doi.org/10.1109/JIOT.2024.3485027

Altini, M. (2024). Making sense of wearable data. *Sports Science Exchange*, *37*(250), 1-6.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5* (Fifth edition ed.). American Psychiatric Association. http://dsm.psychiatryonline.org/book.aspx?bookid=556

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev. ed.). https://doi.org/10.1176/appi.books.9780890425787

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Clinical Psychology*, *53*, 27-51. https://doi.org/10.1146/annurev.psych.53.100901.135231

Anderson, C. A., & Bushman, B. J. (2018). Media Violence and the General Aggression Model. *The Journal of Social Issues*, *74*(2), 386-413. https://doi.org/10.1111/josi.12275

Armey, M. F., Crowther, J. H., & Miller, I. W. (2011). Changes in ecological momentary assessment reported affect associated with episodes of nonsuicidal self-injury. *Behavior Therapy*, *42*(4), 579-588. https://doi.org/10.1016/j.beth.2011.01.002

Arnetz, J. E. P. M. P. H. P. T., Hamblin, L. M. A., Ager, J. P., Luborsky, M. P., Upfal, M. J. M. D. M. P. H., Russell, J. B. S. N., & Essenmacher, L. M. P. H. (2015). Underreporting of workplace violence: Comparison of self-report and actual documentation of hospital incidents. *Workplace Health & Safety*, *63*(5), 200-210. https://doi.org/10.1177/2165079915574684

Ataie, S., Gharraee, B., & Shahboulaghi, F. M. (2022). The Effectiveness of Emotion Regulation Therapy for Adolescent Females With Nonsuicidal Self-Injury Disorder: A Mixed-Method Study. *The Journal of Nervous and Mental Disease*, *210*(6), 411-417. https://dx.doi.org/10.1097/NMD.0000000000001476

Bali, A., & Jaggi, A. S. (2015). Clinical experimental stress studies: methods and assessment. *Reviews in the Neurosciences*, *26*(5), 555-579. https://doi.org/doi:10.1515/revneuro-2015-0004

Baron, R. A. (1983). The control of human aggression: An optimistic perspective. *Journal of Social and Clinical Psychology*, *1*(2), 97-119. https://doi.org/10.1521/jscp.1983.1.2.97

Barrera, F. J., Violo, R. A., & Graver, E. E. (2007). On the form and function of severe self-injurious behavior. *Behavioral Interventions*, *22*(1), 5-33. https://dx.doi.org/10.1002/bin.228

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1 - 48. https://doi.org/10.18637/jss.v067.i01

Beauchaine, T. (2001). Vagal tone, development, and Gray's motivational theory: Toward an integrated model of autonomic nervous system functioning in psychopathology. *Development and Psychopathology*, *13*(2), 183-214. https://doi.org/10.1017/S0954579401002012

Beauchaine, T. P. (2015). Respiratory Sinus Arrhythmia: A Transdiagnostic Biomarker of Emotion Dysregulation and Psychopathology. *Current Opinion in Psychology*, *3*, 43-47. https://doi.org/10.1016/j.copsyc.2015.01.017

Beauchaine, T. P., & Thayer, J. F. (2015). Heart rate variability as a transdiagnostic biomarker of psychopathology. *International Journal of Psychophysiology*, *98*(2), 338-350. https://doi.org/10.1016/j.ijpsycho.2015.08.004

Behnke, M., Kreibig, S. D., Kaczmarek, L. D., Assink, M., & Gross, J. J. (2022). Autonomic nervous system activity during positive emotions: a meta-analytic review. *Emotion Review*, *14*(2), 132-160. https://doi.org/10.1177/17540739211073084

Bellato, A., Admani, M. A., Deak, C., Farhat, L. C., Fontana Antunes de Oliveira, M. C., Vasconcelos, R., Malanchini, M., Shephard, E., & Michelini, G. (2023). Autonomic dysregulation and self-injurious thoughts and behaviours in children and young people: A systematic review and meta-analysis. *JCPP Advances*, *3*(3), e12148. https://doi.org/10.1016/j.jaac.2023.01.017

Bellemans, T., Didden, R., van Busschbach, J. T., Hoek, P. T., Scheffers, M. W., Lang, R. B., & Lindsay, W. R. (2019). Psychomotor therapy targeting anger and aggressive behaviour in individuals with mild or borderline intellectual disabilities: A systematic review. *Journal of Intellectual & Developmental Disability*, *44*(1), 121-130. https://doi.org/10.3109/13668250.2017.1326590

Bellemans, T., Peters-Scheffer, N., Didden, R., Traas, R., & van Busschbach, J. T. (2022). Psychomotor therapy for individuals with mild intellectual disabilities or borderline intellectual functioning presenting anger regulation problems and/or aggressive behaviour: A qualitative study on clients' experiences. *Journal of Intellectual and Developmental Disability*, *47*(1), 74-86. https://doi.org/10.3109/13668250.2021.1899561

Berkowitz, L. (1989). Frustration-aggression hypothesis: examination and reformulation. *Psychological Bulletin*, *106*(1), 59. https://doi.org/10.1037/0033-2909.106.1.59

Berkowitz, L. (1993). Aggression: Its causes, consequences, and control. *McGrawll Hill*.

Berntson, G. G., Norman, G. J., Hawkley, L. C., & Cacioppo, J. T. (2008). Cardiac autonomic balance versus cardiac regulatory capacity. *Psychophysiology*, *45*(4), 643-652. https://doi.org/10.1111/j.1469-8986.2008.00652.x

Betella, A., Verschure, P. F. M. J., & Tran, U. S. E. (2016). The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions. *PloS one*, *11*(2). https://doi.org/10.1371/journal.pone.0148037

Bjorkqvist, K., & Niemela, P. (1992). *Of mice and women: Aspects of female aggression*. Academic Press. *https://doi.org/10.5860/CHOICE.30-5284*

Blankenstein, N. E., Vandenbroucke, A. R., de Vries, R., Swaab, H., Popma, A., & Jansen, L. (2022). Understanding aggression in adolescence by studying the neurobiological stress system: A systematic review. *Motivation Science*, *8*(2), 133. https://doi.org/10.1037/mot0000259

Bøe, A. S., Mehlum, L., Melle, I., & Qin, P. (2022). Psychiatric disorders among adults treated for deliberate self-harm in general hospital: A national register study. *Journal of Affective Disorders*, *319*, 490-496. https://doi.org/10.1016/j.jad.2022.09.066

Bohus, M., Limberger, M., Ebner, U., Glocker, F. X., Schwarz, B., Wernz, M., & Lieb, K. (2000). Pain perception during self-reported distress and calmness in patients with borderline personality disorder and self-mutilating behavior. *Psychiatry Research*, *95*(3), 251-260. https://doi.org/10.1016/S0165-1781(00)00179-7

Boucsein, W. (2012). *Electrodermal activity*. Springer Science & Business Media.

Bowers, L., Stewart, D., Papadopoulos, C., Dack, C., Ross, J., Khanom, H., & Jeffery, D. (2011). Inpatient violence and aggression: a literature review. *Report from the Conflict and Containment Reduction Research Programme, Kings College, London, available at: www. kcl. ac. uk/iop/depts/hspr/research/ciemh/mhn/projects/litreview/LitRevAgg. pdf*.

Bowring, D. L., Totsika, V., Hastings, R. P., Toogood, S., & Griffith, G. M. (2017). Challenging behaviours in adults with an intellectual disability: A total population study and exploration of risk indices. *British Journal of Clinical Psychology*, *56*(1), 16-32. https://doi.org/10.1111/bjc.12118

Brain, K., Haines, J., & Williams, C. (1998). The psychophysiology of self-mutilation: Evidence of tension reduction. *Archives of Suicide Research* *4*(3), 227-242. https://doi.org/10.1080/13811119808258298

Brain, K. L., Haines, J., & Williams, C. L. (2002). The psychophysiology of repetitive self-mutilation. *Archives of Suicide Research*, *6*(3), 199-210. https://dx.doi.org/10.1080/13811110214140

Brown, S., & Kalaitzidis, E. (2013). Barriers preventing high-quality nursing care of people with disabilities within acute care settings: a thematic literature review. *Disability & Society*, *28*(7), 937-954. https://doi.org/10.1080/09687599.2012.748646

Bucci, S., Berry, N., Morris, R., Berry, K., Haddock, G., Lewis, S., & Edge, D. (2019). "They are not hard-to-reach clients. We have just got hard-to-reach services." Staff views of digital health tools in specialist mental health services. *Frontiers in Psychiatry*, *10*, 344. https://doi.org/10.3389/fpsyt.2019.00344

Bushman, B. J. (2002). Does venting anger feed or extinguish the flame? Catharsis, rumination, distraction, anger, and aggressive responding. *Personality and Social Psychology Bulletin*, *28*(6), 724-731. *https://doi.org/10.1177/0146167202289002*

Byrne, G. (2018). Prevalence and psychological sequelae of sexual abuse among individuals with an intellectual disability: A review of the recent literature. *Journal of Intellectual Disabilities*, *22*(3), 294-310. https://doi.org/10.1177/1744629517698844

Calamari, J. E., McNally, R. J., Benson, D. S., & Babington, C. M. (1990). Case study: Use of propranolol to reduce aggressive behavior in a woman who is mentally retarded. *Behavioral Residential Treatment*, *5*(4), 287-296. https://doi.org/10.1002/bin.2360050407

Calver, J., Balogh, R., & Rudoler, D. (2021). Incidence of injury in children and adolescents with intellectual and developmental disability. *Journal of Safety Research*, *77*, 56-60. https://doi.org/10.1016/j.jsr.2021.02.003

Carroll, R., Metcalfe, C., Gunnell, D., & Gluud, L. L. (2014). Hospital Presenting Self-Harm and Risk of Fatal and Non-Fatal Repetition: Systematic Review and Meta-Analysis. *PloS one*, *9*(2). https://doi.org/10.1371/journal.pone.0089944

Cassels, M., Neufeld, S., Goodyer, I., & Wilkinson, P. (2022). Prospective Pathways From Impulsivity to Non-Suicidal Self-Injury Among Youth. *Archives of Suicide Research*, *26*(2), 534-547. https://doi.org/10.1080/13811118.2020.1811180

Challinor, A., Naylor, K., & Verstreken, P. (2021). The effects of COVID-19 on self-harm in a high-secure psychiatry hospital. *The Journal of Forensic Practice*, *23*(3), 230-239. https://doi.org/10.1108/JFP-04-2021-0019

Chieze, M., Hurst, S., Kaiser, S., & Sentissi, O. (2019). Effects of seclusion and restraint in adult psychiatry: a systematic review. *Frontiers in Psychiatry*, *10*, 491. https://doi.org/10.3389/fpsyt.2019.00491

Christopoulos, G. I., Uy, M. A., & Yap, W. J. (2019). The body and the brain: Measuring skin conductance responses to understand the emotional experience. *Organizational Research Methods*, *22*(1), 394-420. https://doi.org/10.1177/1094428116681073

Cohen, I. L., & Tsiouris, J. A. (2020). Triggers of aggressive behaviors in intellectually disabled adults and their association with autism, medical conditions, psychiatric disorders, age and sex: a large-scale study. *Journal of Autism and Developmental Disorders*, *50*(10), 3748-3762. https://doi.org/10.1007/s10803-020-04424-4

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Cowman, S., Björkdahl, A., Clarke, E., Gethin, G., & Maguire, J. (2017). A descriptive survey study of violence management and priorities among psychiatric staff in mental health services, across seventeen european countries. *BMC Health Services Research*, *17*, 1-10. https://doi.org/10.1186/s12913-017-1988-7

Crocker, A. G., Mercier, C., Lachapelle, Y., Brunet, A., Morin, D., & Roy, M. E. (2006). Prevalence and types of aggressive behaviour among adults with intellectual disabilities. *Journal of Intellectual Disability Research*, *50*(9), 652-661. https://doi.org/10.1111/j.1365-2788.2006.00815.x

Crowell, S. E., Beauchaine, T. P., Hsiao, R. C., Vasilev, C. A., Yaptangco, M., Linehan, M. M., & McCauley, E. (2012). Differentiating adolescent self-injury from adolescent depression: possible implications for borderline personality development. *Journal of Abnormal Child Psychology*, *40*(1), 45-57. https://dx.doi.org/10.1007/s10802-011-9578-3

Crowell, S. E., Beauchaine, T. P., McCauley, E., Smith, C. J., Stevens, A. L., & Sylvers, P. (2005). Psychological, autonomic, and serotonergic correlates of parasuicide among adolescent girls. *Development and Psychopathology*, *17*(4), 1105-1127. https://dx.doi.org/10.1017/S0954579405050522

Crowell, S. E., Butner, J. E., Wiltshire, T. J., Munion, A. K., Yaptangco, M., & Beauchaine, T. P. (2017). Evaluating emotional and biological sensitivity to maternal behavior among self-injuring and depressed adolescent girls using nonlinear dynamics. *Clinical Psychological Science*, *5*(2), 272-285. https://dx.doi.org/10.1177/2167702617692861

Daffern, M., & Howells, K. (2009). The function of aggression in personality disordered patients. *Journal of Interpersonal Violence*, *24*(4), 586-600. https://doi.org/10.1177/0886260508317178

Daukantaitė, D., Lundh, L.-G., Wångby-Lundh, M., Claréus, B., Bjärehed, J., Zhou, Y., & Liljedahl, S. I. (2021). What happens to young adults who have engaged in self-injurious behavior as adolescents? A 10-year follow-up. *European Child & Adolescent Psychiatry*, *30*, 475-492. https://doi.org/10.1007/s00787-020-01533-4

De Geus, E. J., Willemsen, G. H., Klaver, C. H., & van Doornen, L. J. (1995). Ambulatory measurement of respiratory sinus arrhythmia and respiration rate. *Biological Psychology*, *41*(3), 205-227. https://doi.org/10.1016/0301-0511(95)05137-6

De Looff, P., Cornet, L., De Kogel, C., Fernández-Castilla, B., Embregts, P., Didden, R., & Nijman, H. (2022b). Heart rate and skin conductance associations with physical aggression, psychopathy, antisocial personality disorder and conduct disorder: An updated meta-analysis. *Neuroscience & Biobehavioral Reviews*, *132*, 553-582. https://doi.org/10.1016/j.neubiorev.2021.11.003

De Looff, P., Duursma, R., Noordzij, M., Taylor, S., Jaques, N., Scheepers, F., de Schepper, K., & Koldijk, S. (2022). Wearables: an R package with accompanying shiny application for signal analysis of a wearable device targeted at clinicians and researchers. *Frontiers in Behavioral Neuroscience*, *16*, 856544. https://doi.org/10.3389/fnbeh.2022.856544

De Looff, P., Nijman, H., Didden, R., & Embregts, P. (2018). Burnout symptoms in forensic psychiatric nurses and their associations with personality, emotional intelligence and client aggression: A cross-sectional study. *Journal of Psychiatric and Mental Health Nursing*, *25*(8), 506-516. https://doi.org/10.1111/jpm.12496

De Looff, P., Noordzij, M. L., Moerbeek, M., Nijman, H., Didden, R., & Embregts, P. (2019). Changes in heart rate and skin conductance in the 30 min preceding aggressive behavior. *Psychophysiology*, *56*(10). https://doi.org/10.1111/psyp.13420

De Looff, P. C., Nijman, H., Didden, R., & Noordzij, M. L. (2021). Usability and acceptance of wearable biosensors in forensic psychiatry: cross-sectional questionnaire study. *JMIR Formative Research*, *5*(5), e18096. https://doi.org/10.2196/18096

De Looff, P. C., Noordzij, M. L., Nijman, H. L., Goedhard, L., Bogaerts, S., & Didden, R. (2024). Putting the usability of wearable technology in forensic psychiatry to the test: a randomized crossover trial. *Frontiers in Psychiatry*, *15*, 1330993. https://doi.org/10.3389/fpsyt.2024.1330993

De Vries, S., van Oost, F., Smaling, H., de Knegt, N., Cluitmans, P., Smits, R., & Meinders, E. (2023). Real-time stress detection based on artificial intelligence for people with an intellectual disability. *Assistive Technology*, *36*(3), 232-240. https://doi.org/10.1080/10400435.2023.2261045

De Witte, M., Spruit, A., van Hooren, S., Moonen, X., & Stams, G.-J. (2020). Effects of music interventions on stress-related outcomes: a systematic review and two meta-analyses. *Health Psychology Review*, *14*(2), 294-324. https://doi.org/10.1080/17437199.2019.1627897

De Witte, N. A., Buyck, I., & Van Daele, T. (2019). Combining biofeedback with stress management interventions: A systematic review of physiological and psychological effects. *Applied Psychophysiology and Biofeedback*, *44*(2), 71-82. https://doi.org/10.1007/s10484-018-09427-7

Deeks, J. J., Dinnes, J., D'Amico, R., Sowden, A. J., Sakarovitch, C., Song, F., Petticrew, M., Altman, D., Group, I. S. T. C., & Group, E. C. S. T. C. (2003). Evaluating non-randomised intervention studies. *Health Technology Assessment*, *7*(27), iii-173. https://doi.org/10.3310/hta7270

Delforterie, M., Hesper, B., & Didden, R. (2020). Psychometric properties of the Dynamic Risk Outcome Scales (DROS) for individuals with mild intellectual disability or borderline intellectual functioning and externalizing behaviour problems. *Journal of Applied Research in Intellectual Disabilities*, *33*(4), 662-672. https://doi.org/10.1111/jar.12546

Didden, R., & Mevissen, L. (2022). Trauma in individuals with intellectual and developmental disabilities: Introduction to the Special Issue. *Research in Developmental Disabilities*, *120*, 104122. https://doi.org/10.1016/j.ridd.2021.104122

Didden, R., Nijman, H., Delforterie, M., & Keulen-De Vos, M. (2019). Treatment of Anger and Violence in Individuals with Intellectual Disability. In *The Wiley Handbook on What Works for Offenders with Intellectual and Developmental Disabilities* (pp. 297-309). https://doi.org/10.1002/9781119316268.ch15

Dixon-Gordon, K., Harrison, N., & Roesch, R. (2012). Non-Suicidal Self-Injury Within Offender Populations: A Systematic Review. *International Journal of Forensic Mental Health*, *11*(1), 33-50. *https://doi.org/10.1080/14999013.2012.667513*

Dodge, K. A. (1980). Social cognition and children's aggressive behavior. *Child Development*, 162-170. https://doi.org/10.2307/1129603

Doherty, C., Baldwin, M., Lambe, R., Burke, D., & Altini, M. (2025). Readiness, recovery, and strain: an evaluation of composite health scores in consumer wearables. *Translational Exercise Biomedicine*(0). https://doi.org/10.1515/teb-2025-0001

Dollard, J., Coob, L., Miller, N., Mowrer, O., & Sears, R. (1939). Frustration and Aggression. Yale University. *Press, New Haven*.

Drieschner, K. H., & Hesper, B. L. (2008). *Dynamic risk outcome scales.* Trajectum.

Drieschner, K. H., Marrozos, I., & Regenboog, M. (2013). Prevalence and risk factors of inpatient aggression by adults with intellectual disabilities and severe challenging behaviour: A long-term prospective study in two Dutch treatment facilities. *Research in Developmental Disabilities*, *34*(8), 2407-2418. https://doi.org/10.1016/j.ridd.2013.04.008

Duarte, T. A., Paulino, S., Almeida, C., Gomes, H. S., Santos, N., & Gouveia-Pereira, M. (2020). Self-harm as a predisposition for suicide attempts: A study of adolescents' deliberate self-harm, suicidal ideation, and suicide attempts. *Psychiatry Research*, *287*, 112553. https://doi.org/10.1016/j.psychres.2019.112553

Dunn, J., Mishra, V., Shandhi, M. M. H., Jeong, H., Yamane, N., Watanabe, Y., Chen, B., & Goodwin, M. S. (2025). Building an open-source community to enhance autonomic nervous system signal analysis: DBDP-autonomic. *Frontiers in Digital Health*, *6*, 1467424. https://doi.org/10.3389/fdgth.2024.1467424

Eron, L. D., Walder, L. O., & Lefkowitz, M. M. (1971). Learning of aggression in children.

Everly, S. G., & Lating Jr, M. (2019). *A clinical guide to the treatment of the human stress response*. Springer.

Fanneran, T., Brimblecombe, N., Bradley, E., & Gregory, S. (2015). Using workload measurement tools in diverse care contexts: the experience of staff in mental health and learning disability inpatient settings. *Journal of Psychiatric and Mental Health Nursing*, *22*(10), 764-772. https://doi.org/10.1111/jpm.12263

Fedor, S., Lewis, R., Pedrelli, P., Mischoulon, D., Curtiss, J., & Picard, R. W. (2023). Wearable technology in clinical practice for depressive disorder. *New England Journal of Medicine*, *389*(26), 2457-2466. https://doi.org/10.1056/NEJMra2215898

Ferguson, B. J., Hamlin, T., Lantz, J. F., Villavicencio, T., Coles, J., & Beversdorf, D. Q. (2019). Examining the association between electrodermal activity and problem behavior in severe autism spectrum disorder: A feasibility study. *Frontiers in Psychiatry*, *10*. https://dx.doi.org/10.3389/fpsyt.2019.00654

Ferguson, C. J., & Dyck, D. (2012). Paradigm change in aggression research: The time has come to retire the General Aggression Model. *Aggression and Violent Behavior*, *17*(3), 220-228. https://doi.org/10.1016/j.avb.2012.02.007

Ferguson, C. J., Maguire, R., & Lemar, S. (2018). Pick your poison: Choice of activity determines mood management following a stressful task. *Journal of Aggression, Maltreatment & Trauma*, *27*(3), 332-346. https://doi.org/10.1080/10926771.2017.1322656

Ferguson, C. J., Rueda, S. M., Cruz, A. M., Ferguson, D. E., Fritz, S., & Smith, S. M. (2008). Violent video games and aggression: Causal relationship or byproduct of family violence and intrinsic violence motivation? *Criminal Justice and Behavior*, *35*(3), 311-332. https://doi.org/10.1177/0093854807311719

Ferreira, A. F., da Silva, H. P., Alves, H., Marques, N., & Fred, A. (2023). Feasibility of Electrodermal Activity and Photoplethysmography Data Acquisition at the Foot Using a Sock Form Factor. *Sensors*, *23*(2). https://doi.org/10.3390/s23020620

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics: and sex and drugs and rock 'n' roll* (4th ed.). Sage.

Fish, R. M. (2000). Working with People Who Harm Themselves in a Forensic Learning Disability Service : Experiences of Direct Care Staff. *Journal of Learning Disabilities*, *4*(3), 193-207. https://doi.org/10.1177/146900470000400302

Fitzpatrick, Zeifman, R., Krantz, L., McMain, S., & Kuo, J. R. (2020). Getting specific about emotion and self-inflicted injury: An examination across emotion processes in borderline personality disorder. *Archives of Suicide Research*, *24*(sup1), 102-123. https://doi.org/10.1080/13811118.2019.1586605

Fliege, H., Lee, J. R., Grimm, A., & Klapp, B. F. (2009). Risk factors and correlates of deliberate self-harm behavior: a systematic review. *Journal of Psychosomatic Research* *66*(6), 477-493. https://doi.org/10.1016/j.jpsychores.2008.10.013

Flygare Wallén, E., Ljunggren, G., Wahlström, L., Pettersson, D., Carlsson, A. C., & Wändell, P. (2023). The prevalence of self-harm and mental disorders among individuals with intellectual disabilities. *Nordic Journal of Psychiatry*, *77*(7), 712-720. https://doi.org/10.1080/08039488.2023.2228292

Forte, M., Jahoda, A., & Dagnan, D. (2011). An anxious time? Exploring the nature of worries experienced by young people with a mild to moderate intellectual disability as they make the transition to adulthood. *British Journal of Clinical*

Psychology, *50*(4), 398-411. *https://doi.org/10.1111/j.2044-8260.2010.02002.x*
Fox, A. R., Hammond, L. E., & Mezulis, A. H. (2018). Respiratory sinus arrhythmia and adaptive emotion regulation as predictors of nonsuicidal self-injury in young adults. *International Journal of Psychophysiology*, *133*, 1-11. https://dx.doi.org/10.1016/j.ijpsycho.2018.09.006
Frankfurt, S., Frazier, P., Syed, M., & Jung, K. R. (2016). Using group-based trajectory and growth mixture modeling to identify classes of change trajectories. *The Counseling Psychologist*, *44*(5), 622-660. https://doi.org/10.1177/0011000016658097
Franz, P. J., Kleiman, E. M., & Nock, M. K. (2019). Why do people hurt themselves. *The Oxford Handbook of Positive Emotion and Psychopathology*, *453*.
Freeman, R., Horner, R., & Reichle, J. (1999). Relation between heart rate and problem behaviors. *American Journal of Mental Retardation*, *104*(4), 330-345. https://doi.org/10.1352/0895-8017(1999)104%3C0330:RBHRAP%3E2.0.CO;2
Freud, S. (1915). Instinct and their Vicisitudes. In *S. E. V. XIV (1914-1916)*. The Hogarth Press and the Institute of Psychoanalysis.
Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin*, *140*(1), 1. https://doi.org/10.1037/a0033076
Frowijn, I., Masthoff, E., Vermunt, J. K., & Bogaerts, S. (2024). Transgressive incidents targeted on staff in forensic psychiatric healthcare: a latent class analysis. *Frontiers in Psychiatry*, *15*, 1394535. https://doi.org/10.3389/fpsyt.2024.1394535
Furniss, F., & Biswas, A. (2012). Recent research on aetiology, development and phenomenology of self-injurious behaviour in people with intellectual disabilities: a systematic review and implications for treatment. *Journal of Intellectual Disability Research*, *56*(5), 453-475. https://doi.org/10.1111/j.1365-2788.2012.01534.x
Ghiasi, S., Greco, A., Barbieri, R., Scilingo, E. P., & Valenza, G. (2020). Assessing Autonomic Function from Electrodermal Activity and Heart Rate Variability During Cold-Pressor Test and Emotional Challenge. *Scientific Reports*, *10*(1), 5406. https://doi.org/10.1038/s41598-020-62225-2
Ghinea, D., Edinger, A., Parzer, P., Koenig, J., Resch, F., & Kaess, M. (2020). Non-suicidal self-injury disorder as a stand-alone diagnosis in a consecutive help-seeking sample of adolescents. *Journal of Affective Disorders*, *274*, 1122-1125. https://doi.org/10.1016/j.jad.2020.06.009
Gigi, K., Werbeloff, N., Goldberg, S., Portuguese, S., Reichenberg, A., Fruchter, E., & Weiser, M. (2014). Borderline intellectual functioning is associated with poor social functioning, increased rates of psychiatric diagnosis and drug use–A cross sectional population based study. *European Neuropsychopharmacology*, *24*(11), 1793-1797. https://doi.org/10.1016/j.euroneuro.2014.07.016
Giner-Bartolome, C., Mallorqui-Bague, N., Tolosa-Sola, I., Steward, T., Jimenez-Murcia, S., Granero, R., & Fernandez-Aranda, F. (2017). Non-suicidal Self-In-

jury in Eating Disordered Patients: Associations with Heart Rate Variability and State-Trait Anxiety. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01163

Ginsburg, G. S., Picard, R. W., & Friend, S. H. (2024). Key issues as wearable digital health technologies enter clinical care. *New England Journal of Medicine*, *390*(12), 1118-1127. https://doi.org/10.1056/NEJMra2307160

Giordano, A. L., Prosek, E. A., Schmit, E. L., & Schmit, M. K. (2023). Examining coping and nonsuicidal self-injury among adolescents: A profile analysis. *Journal of Counseling & Development*, *101*(2), 214-223. *https://doi.org/10.1002/jcad.12459*

Girasek, H., Nagy, V. A., Fekete, S., Ungvari, G. S., & Gazdag, G. (2022). Prevalence and correlates of aggressive behavior in psychiatric inpatient populations. *World Journal of Psychiatry*, *12*(1), 1. https://dx.doi.org/10.5498/wjp.v12.i1.1

Godfrey, D. A., Kaufman, E. A., & Crowell, S. E. (2022). Non-suicidal Self-injury History Moderates the Association Between Maternal Emotional Support and Adolescent Affect During Conflict. *Child Psychiatry and Human Development*, 415-425. https://dx.doi.org/10.1007/s10578-022-01417-5

Goodwin, M. S., Mazefsky, C. A., Ioannidis, S., Erdogmus, D., & Siegel, M. (2019). Predicting aggression to others in youth with autism using a wearable biosensor. *Autism Research*, *12*(8), 1286-1296. https://doi.org/10.1002/aur.2151

Gordon, K. H., Selby, E. A., Anestis, M. D., Bender, T. W., Witte, T. K., Braithwaite, S., Van Orden, K. A., Bresin, K., & Joiner Jr, T. E. (2010). The reinforcing properties of repeated deliberate self-harm. *Archives of Suicide Research*, *14*(4), 329-341. https://doi.org/10.1080/13811118.2010.524059

Goreis, A., Prillinger, K., Bedus, C., Lipp, R., Mayer, A., Nater, U. M., Koenig, J., Plener, P. L., & Kothgassner, O. D. (2023). Physiological stress reactivity and self-harm: A meta-analysis. *Psychoneuroendocrinology*, 106406. https://doi.org/10.1016/j.psyneuen.2023.106406

Gratz, K. L., Hepworth, C., Tull, M. T., Paulson, A., Clarke, S., Remington, B., & Lejuez, C. (2011). An experimental investigation of emotional willingness and physical pain tolerance in deliberate self-harm: The moderating role of interpersonal distress. *Comprehensive Psychiatry*, *52*(1), 63-74. https://doi.org/10.1016/j.comppsych.2010.04.009

Gratz, K. L., Richmond, J. R., Dixon-Gordon, K. L., Chapman, A. L., & Tull, M. T. (2019). Multimodal assessment of emotional reactivity and regulation in response to social rejection among self-harming adults with and without borderline personality disorder. *Personality Disorders*, *10*(5), 395-405. https://doi.org/https://dx.doi.org/10.1037/per0000334

Griffith, G. M., Hutchinson, L., & Hastings, R. P. (2013). "I'm not a patient, I'm a person": The Experiences of Individuals With Intellectual Disabilities and Challenging Behavior—A Thematic Synthesis of Qualitative Studies. *Clinical Psychology: Science and Practice*, *20*(4), 469-488. https://doi.org/10.1111/cpsp.12053

Groves, C. L., & Anderson, C. A. (2016). *Aversive events and aggression* (1 ed.). Rou-

tledge.
Guerdjikova, A. I., Gwizdowski, I. S., McElroy, S. L., McCullumsmith, C., & Suppes, P. (2014). Treating nonsuicidal self-injury. *Current Treatment Options in Psychiatry*, *1*, 325-334. https://doi.org/ 10.1007/s40501-014-0028-z
Hagoort, K., Smeets, K., Koldijk, S., Scheepers, F., & Velders, F. (2025). Collecting Real-Life Psychophysiological Data via Wearables to Better Understand Child Behavior in a Children's Psychiatric Center: Mixed Methods Study on Feasibility and Implementation. *JMIR Formative Research*, *9*(1), e65559. https://doi.org/10.2196/65559
Hagopian, L. P., & Leoni, M. (2017). Self-injurious behavior among individuals with intellectual and developmental disabilities. *Acta Psychopathologica*, *3*(5), 1-9.
Hall, S. S., Hammond, J. L., & Hustyi, K. M. (2013). Examining the relationship between heart rate and problem behavior: a case study of severe skin picking in Prader-Willi syndrome. *American Journal on Intellectual and Developmental Disabilities*, *118*(6), 460-474. https://dx.doi.org/10.1352/1944.7558-118.6.460
Hallett, N., Whittington, R., Richter, D., & Eneje, E. (2024). *Coercion and Violence in Mental Health Settings*. Springer. *https://doi.org/10.1007/978-3-031-61224-4*
Hamza, C. A., Willoughby, T., & Heffer, T. (2015). Impulsivity and nonsuicidal self-injury: A review and meta-analysis. *Clinical Psychology Review*, *38*, 13-24. https://dx.doi.org/10.1016/j.cpr.2015.02.010
Hartley, S. L., & MacLean Jr, W. E. (2008). Coping strategies of adults with mild intellectual disability for stressful social interactions. *Journal of Mental Health Research in Intellectual Disabilities*, *1*(2), 109-127. https://doi.org/10.1080/19315860801988426
Hasselman, F. (2023). Understanding the complexity of individual developmental pathways: A primer on metaphors, models, and methods to study resilience in development. *Development and Psychopathology*, *35*(5), 2186-2198. https://doi.org/10.1017/S0954579423001281
Hassiotis, A., Emerson, E., Wieland, J., & Bertelli, M. O. (2022). Borderline Intellectual Functioning. In *Textbook of Psychiatry for Intellectual Disability and Autism Spectrum Disorder* (pp. 95-106). Springer.
Hassiotis, A., Strydom, A., Hall, I., Ali, A., Lawrence-Smith, G., Meltzer, H., Head, J., & Bebbington, P. (2008). Psychiatric morbidity and social functioning among adults with borderline intelligence living in private households. *Journal of Intellectual Disability Research*, *52*(2), 95-106. https://doi.org/10.1111/j.1365-2788.2007.01001.x
Haugvaldstad, M. J., & Husum, T. L. (2016). Influence of staff's emotional reactions on the escalation of patient aggression in mental health care. *International Journal of Law and Psychiatry*, *49*, 130-137. https://doi.org/10.1016/j.ijlp.2016.09.001
Havercamp, S. M., Barnhill, L. J., Bonardi, A., Chapman, R. A., Cobranchi, C., Fletcher, R. J., Rabidoux, P., Seeley, J. R., Tassé, M. J., Health, N. C. R. o., & Function. (2022). Straight from the horse's mouth: Increasing self-report in mental health assess-

6

ment in individuals with intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, *35*(2), 471-479. https://doi.org/10.1111/jar.12952

Heck, R., & Thomas, S. L. (2020). *An introduction to multilevel modeling techniques: MLM and SEM approaches*. Routledge.

Heirbaut, T., Kip, H., Beerlage-de Jong, N., Vroegop, M., Meerman, S., van IJzendoorn, S., & de Looff, P. (2024). *Meten is weten: Een pilotstudie naar de integratie van wearables in de behandeling van agressie in de ambulante de forensische zorg*. KFZ.

Hekler, E. B., Klasnja, P., Chevance, G., Golaszewski, N. M., Lewis, D., & Sim, I. (2019). Why we need a small data paradigm. *BMC Medicine*, *17*(1), 133. https://doi.org/10.1186/s12916-019-1366-x

Hensel, J. M., Lunsky, Y., & Dewa, C. S. (2014). The mediating effect of severity of client aggression on burnout between hospital inpatient and community residential staff who support adults with intellectual disabilities. *Journal of Clinical Nursing*, *23*(9-10), 1332-1341.

Herzog, R., Álvarez-Pasquin, M. J., Díaz, C., Del Barrio, J. L., Estrada, J. M., & Gil, Á. (2013). Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review. *BMC Public Health*, *13*, 1-17. https://doi.org/10.1186/1471-2458-13-154

Hoch, J., Symons, F., & Sng, S. (2013). Sequential analysis of autonomic arousal and self-injurious behavior. *American Journal on Intellectual and Developmental Disabilities*, *118*(6), 435-446. https://dx.doi.org/10.1352/1944.7558-118.6.435

Hokanson, J. E., & Burgess, M. (1962). The effects of three types of aggression on vascular processes. *The Journal of Abnormal and Social Psychology*, *64*(6), 446. https://doi.org/10.1037/h0040236

Hooley, J. M., & Franklin, J. C. (2018). Why Do People Hurt Themselves? A New Conceptual Model of Nonsuicidal Self-Injury. *Clinical Psychological Science*, *6*(3), 428-451. https://doi.org/10.1177/2167702617745641

Hossain, M. B., Kong, Y., Posada-Quintero, H. F., & Chon, K. H. (2022). Comparison of Electrodermal Activity from Multiple Body Locations Based on Standard EDA Indices' Quality and Robustness against Motion Artifact. *Sensors*, *22*(9). https://doi.org/10.3390/s22093177

Hu, B., Peng, H., Zhao, Q., Hu, B., Majoe, D., Zheng, F., & Moore, P. (2015). Signal quality assessment model for wearable EEG sensor on prediction of mental stress. *IEEE Transactions on Nanobioscience*, *14*(5), 553-561. https://doi.org/10.1109/TNB.2015.2420576

Hubbard, J. A., Smithmyer, C. M., Ramsden, S. R., Parker, E. H., Flanagan, K. D., Dearing, K. F., Relyea, N., & Simons, R. F. (2002). Observational, physiological, and self–report measures of children's anger: Relations to reactive versus proactive aggression. *Child Development*, *73*(4), 1101-1118. https://doi.org/10.1111/1467-8624.00460

Huesmann, L. R. (1982). Information processing models of behavior. *Multivariate*

Applications in the Social Sciences, 261-288.

Hulsmans, D. H., Otten, R., Poelen, E. A., van Vonderen, A., Daalmans, S., Hasselman, F., Olthof, M., & Lichtwarck-Aschoff, A. (2024). A complex systems perspective on chronic aggression and self-injury: case study of a woman with mild intellectual disability and borderline personality disorder. *BMC Psychiatry*, *24*(1), 378. https://doi.org/10.1186/s12888-024-05836-7

Hulsmans, D. H., Poelen, E. A., Lichtwarck-Aschoff, A., & Otten, R. (2023). The feasibility of daily monitoring in adolescents and young adults with mild intellectual disability or borderline intellectual functioning. *Journal of Applied Research in Intellectual Disabilities*, *36*(4), 847-858. https://doi.org/10.1111/jar.13102

Imbiriba, T., Demirkaya, A., Singh, A., Erdogmus, D., & Goodwin, M. S. (2023). Wearable biosensing to predict imminent aggressive behavior in psychiatric inpatient youths with autism. *JAMA Network Open*, *6*(12), e2348898-e2348898. https://doi.org/10.1001/jamanetworkopen.2023.48898

Immanuel, S., Teferra, M. N., Baumert, M., & Bidargaddi, N. (2023). Heart rate variability for evaluating psychological stress changes in healthy adults: a scoping review. *Neuropsychobiology*, *82*(4), 187-202. *https://doi.org/10.1159/000530376*

James, K., Stewart, D., & Bowers, L. (2012). Self-harm and attempted suicide within inpatient psychiatric services: A review of the Literature. *International Journal of Mental Health Nursing*, *21*(4), 301-309. https://doi.org/10.1111/j.1447-0349.2011.00794.x

James, K. M., Balderrama-Durbin, C., Israel, E., Feurer, C., & Gibb, B. E. (2024). Self- and co-regulation of physiological activity during mother-daughter interactions: The role of adolescent non-suicidal self-injury. *Journal of Child Psychology and Psychiatry*, *65*(1), 91-99. https://doi.org/10.1111/jcpp.13859

Jansen, L. M. (2022). The neurobiology of antisocial behavior in adolescence; current knowledge and relevance for youth forensic clinical practice. *Current Opinion in Psychology*, *47*, 101356. https://doi.org/10.1016/j.copsyc.2022.101356

Jarczok, M. N., Guendel, H., McGrath, J. J., & Balint, E. M. (2019). Circadian rhythms of the autonomic nervous system: Scientific implication and practical implementation. *Chronobiology—The science of biological time structure*, 51-70. https://doi.org/10.5772/intechopen.86822

Jarczok, M. N., Jarczok, M., Mauss, D., Koenig, J., Li, J., Herr, R. M., & Thayer, J. F. (2013). Autonomic nervous system activity and workplace stressors—A systematic review. *Neuroscience & Biobehavioral Reviews*, *37*(8), 1810-1823. https://doi.org/10.1016/j.neubiorev.2013.07.004

Jiang, J., Zhou, L., Rickson, D., & Jiang, C. (2013). The effects of sedative and stimulative music on stress reduction depend on music preference. *The Arts in Psychotherapy*, *40*(2), 201-205. https://doi.org/10.1016/j.aip.2013.02.002

Johnson, K. T., & Picard, R. W. (2020). Advancing neuroscience through wearable devices. *Neuron*, *108*(1), 8-12. *https://doi.org/10.1016/j.neuron.2020.09.030*

Kaess, M., Hille, M., Parzer, P., Maser-Gluth, C., Resch, F., & Brunner, R. (2012). Alterations in the neuroendocrinological stress response to acute psychosocial stress in adolescents engaging in nonsuicidal self-injury. *Psychoneuroendocrinology*, *37*(1), 157-161. https://doi.org/10.1016/j.psyneuen.2011.05.009

Kahle, S., Miller, J. G., Lopez, M., & Hastings, P. D. (2016). Sympathetic recovery from anger is associated with emotion regulation. *Journal of Experimental Child Psychology*, *142*, 359-371. https://doi.org/10.1016/j.jecp.2015.10.004

Kamath, M. V., Watanabe, M., & Upton, A. (2016). *Heart rate variability (HRV) signal analysis: clinical applications*. CRC Press.

Kaufman, E., Puzia, M., Godfrey, D., & Crowell, S. (2020). Physiological and behavioral effects of interpersonal validation: A multilevel approach to examining a core intervention strategy among self-injuring adolescents and their mothers. *Journal of Clinical Psychology*, *76*(3), 559-580. https://doi.org/10.1002/jclp.22902

Kaufman, E. A., Crowell, S. E., Coleman, J., Puzia, M. E., Gray, D. D., & Strayer, D. L. (2018). Electroencephalographic and cardiovascular markers of vulnerability within families of suicidal adolescents: A pilot study. *Biological Psychology*, *136*, 46-56. https://dx.doi.org/10.1016/j.biopsycho.2018.05.007

Kay, S. R., Wolkenfeld, F., & Murrill, L. M. (1988). Profiles of aggression among psychiatric patients: I. Nature and prevalence. *The Journal of Nervous and Mental Disease*, *176*(9), 539-546.

Kim, J., & Stegemann, T. (2016). Music listening for children and adolescents in health care contexts: A systematic review. *The Arts in Psychotherapy*, *51*, 72-85. https://doi.org/10.1016/j.aip.2016.08.007

Kind, N., Eckert, A., Steinlin, C. l., Fegert, J. r. M., & Schmid, M. (2018). Verbal and physical client aggression - A longitudinal analysis of professional caregivers' psychophysiological stress response and burnout. *Psychoneuroendocrinology*, *94*, 11-16. https://doi.org/10.1016/j.psyneuen.2018.05.001

Kip, H., Oberschmidt, K., & Bierbooms, J. J. (2020). eHealth technology in forensic mental healthcare: recommendations for achieving benefits and overcoming barriers. *International Journal of Forensic Mental Health*, *20*(1), 31-47. https://doi.org/10.1080/14999013.2020.1808914

Kivimäki, M., Bartolomucci, A., & Kawachi, I. (2023). The multiple roles of life stress in metabolic disorders. *Nature Reviews Endocrinology*, *19*(1), 10-27. https://doi.org/10.1038/s41574-022-00746-8

Kleckner, I. R., Jones, R. M., Wilder-Smith, O., Wormwood, J. B., Akcakaya, M., Quigley, K. S., Lord, C., & Goodwin, M. S. (2017). Simple, transparent, and flexible automated quality assessment procedures for ambulatory electrodermal activity data. *IEEE Transactions on Biomedical Engineering*, *65*(7), 1460-1467. https://doi.org/10.1109/TBME.2017.2758643

Klimes-Dougan, B., Mirza, S. A., Babkin, E., & Lanning, C. (2022). Biased reporting of past self-injurious thoughts and behaviors: A literature review. *Journal of Affective Disorders*, *308*, 596-606. https://doi.org/10.1016/j.jad.2022.04.073

Klonsky, E. D., May, A. M., & Glenn, C. R. (2013). The relationship between nonsuicidal self-injury and attempted suicide: converging evidence from four samples. *Journal of Abnormal Psychology*, *122*(1), 231. https://doi.org/10.1037/a0030278

Knotter, M. (2019). The Whole Is More, a Contextual Perspective on Attitudes and Reactions of Staff Towards Aggressive Behaviour of Clients with Id in Residential Institutions. *Journal of Intellectual Disability Research*, *63*(7), 687-687. <Go to ISI>://WOS:000474486301079

Knowles, S. E., & Townsend, E. (2012). Implicit and explicit attitudes toward self harm: Support for a functional model. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(2), 730-736. https://doi.org/https://dx.doi.org/10.1016/j.jbtep.2011.10.007

Koenig, J., Lischke, A., Bardtke, K., Heinze, A.-L., Kröller, F., Pahnke, R., & Kaess, M. (2023). Altered psychobiological reactivity but no impairment of emotion recognition following stress in adolescents with non-suicidal self-injury. *European Archives of Psychiatry and Clinical Neuroscience*, *273*(2), 379-395. https://doi.org/10.1007/s00406-022-01496-4

Koenig, J., Rinnewitz, L., Parzer, P., Resch, F., Thayer, J. F., & Kaess, M. (2017). Resting cardiac function in adolescent non-suicidal self-injury: The impact of borderline personality disorder symptoms and psychosocial functioning. *Psychiatry Research*, *248*, 117-120. https://dx.doi.org/10.1016/j.psychres.2016.12.024

Koenig, J., Rinnewitz, L., Warth, M., Hillecke, T. K., Brunner, R., Resch, F., & Kaess, M. (2017). Psychobiological response to pain in female adolescents with nonsuicidal self-injury. *Journal of Psychiatry and Neuroscience*, *42*(3), 189-199. https://dx.doi.org/10.1503/jpn.160074

Koenig, J., Weise, S., Rinnewitz, L., Parzer, P., Resch, F., & Kaess, M. (2018). Longitudinal covariance of resting-state cardiac function and borderline personality disorder symptoms in adolescent non-suicidal self-injury. *World Journal of Biological Psychiatry*, *19*(2), 152-157. https://dx.doi.org/10.1080/15622975.2017.1342046

Konstantopoulou, G., Iliou, T., Karaivazoglou, K., Iconomou, G., Assimakopoulos, K., & Alexopoulos, P. (2020). Associations between (sub) clinical stress- and anxiety symptoms in mentally healthy individuals and in major depression: a cross-sectional clinical study. *BMC Psychiatry*, *20*(1), 428. https://doi.org/10.1186/s12888-020-02836-1

Krug, E. G., Mercy, J. A., Dahlberg, L. L., & Zwi, A. B. (2002). The world report on violence and health. *The Lancet*, *360*(9339), 1083-1088.

Kume, S., Nishimura, Y., Mizuno, K., Sakimoto, N., Hori, H., Tamura, Y., Yamato, M., Mitsuhashi, R., Akiba, K., & Koizumi, J. (2017). Music improves subjective feelings leading to cardiac autonomic nervous modulation: a pilot study. *Frontiers in Neuroscience*, *11*, 108. https://doi.org/10.3389/fnins.2017.00108

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1 - 26. https://doi.org/10.18637/jss.v082.i13

Landgren, V., Hedman, E., Lindblad, I., Gillberg, C., & Fernell, E. (2024). Adult psychiatric and psychosocial outcomes of children with mild intellectual disability: a register follow-up of a population-based cohort. *Journal of Intellectual Disability Research*, *68*(1), 34-44. https://doi.org/10.1111/jir.13087

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer Pub. Co.

Li, M., Li, Q., & Wu, C. (2023). Mechanisms and intervention strategies of emotional dysregulation in individuals with nonsuicidal self-injury. *Interdisciplinary Nursing Research*, *2*(4), 257-263. https://doi.org/10.1097/NR9.0000000000000041

Lin, B., Kaliush, P. R., Conradt, E., Terrell, S., Neff, D., Allen, A. K., Smid, M. C., Monk, C., & Crowell, S. E. (2019). Intergenerational transmission of emotion dysregulation: Part I. Psychopathology, self-injury, and parasympathetic responsivity among pregnant women. *Development and Psychopathology*, *31*(3), 817-831. https://dx.doi.org/10.1017/S0954579419000336

Lokman, S., Van Oorsouw, W. M., Didden, R., & Embregts, P. J. (2022). Setting up a new team of support staff for people with mild intellectual disability or borderline intellectual functioning and severe challenging behaviour: A concept mapping study. *Journal of Applied Research in Intellectual Disabilities*, *35*(6), 1348-1359. https://doi.org/10.1111/jar.13023

Long, K., Felton, J. W., Lilienfeld, S. O., & Lejuez, C. W. (2014). The role of emotion regulation in the relations between psychopathy factors and impulsive and premeditated aggression. *Personality and Individual Differences*, *5*(4), 390-396. https://doi.org/10.1037/per0000085

Lorber, M. F. (2004). Psychophysiology of aggression, psychopathy, and conduct problems: a meta-analysis. *Psychological Bulletin*, *130*(4), 531. https://doi.org/10.1037/0033-2909.130.4.531

Losilla, J.-M., Oliveras, I., Marin-Garcia, J. A., & Vives, J. (2018). Three risk of bias tools lead to opposite conclusions in observational research synthesis. *Journal of Clinical Epidemiology*, *101*, 61-72. *https://doi.org/10.1016/j.jclinepi.2018.05.021*

Lucena, N. L., Rossi, T. A., Azevedo, L. M. G., & Pereira, M. (2022). Self-injury prevalence in adolescents: A global systematic review and meta-analysis. *Children and Youth Services Review*, *142*, 106634. https://doi.org/10.1016/j.childyouth.2022.106634

Lunsky, Y., & Bramston, P. (2006). A preliminary study of perceived stress in adults with intellectual disabilities according to self-report and informant ratings. *Journal of Intellectual and Developmental Disability*, *31*(1), 20-27. https://doi.org/10.1080/13668250500488660

Luther, L., Gearhart, T., Fukui, S., Morse, G., Rollins, A. L., & Salyers, M. P. (2017). Working overtime in community mental health: Associations with clinician burnout and perceived quality of care. *Psychiatric Rehabilitation Journal*, *40*(2), 252. https://psycnet.apa.org/doi/10.1037/prj0000234

Lutz, J., & Krahé, B. (2018). Inducing sadness reduces anger-driven aggressive behavior: A situational approach to aggression control. *Psychology of Violence*, *8*(3), 358. https://doi.org/10.1037/vio0000167

Lydon, S., Healy, O., & Dwyer, M. (2013). An examination of heart rate during challenging behavior in Autism Spectrum Disorder. *Journal of Developmental and Physical Disabilities*, *25*(1), 149-170. https://dx.doi.org/10.1007/s10882-012-9324-y

Lynar, E., Cvejic, E., Schubert, E., & Vollmer-Conna, U. (2017). The joy of heartfelt music: An examination of emotional and physiological responses. *International Journal of Psychophysiology*, *120*, 118-125. https://doi.org/10.1016/j.ijpsycho.2017.07.012

Martínez, C. A. G., Quintana, A. O., Vila, X. A., Touriño, M. J. L., Rodríguez-Liñares, L., Presedo, J. M. R., & Penín, A. J. M. (2017). Heart rate variability analysis with the R package RHRV.

Masino, A. J., Forsyth, D., Nuske, H., Herrington, J., Pennington, J., Kushleyeva, Y., & Bonafide, C. P. (2019, 5-7 June 2019). m-Health and Autism: Recognizing Stress and Anxiety with Machine Learning and Wearables Data. 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS),

Matson, J. L., & Turygin, N. C. (2012). How do researchers define self-injurious behavior? *Research in Developmental Disabilities*, *33*(4), 1021-1026. https://doi.org/10.1016/j.ridd.2012.01.009

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*(3), 276-282. https://hrcak.srce.hr/89395

McLafferty, M., Armour, C., Bunting, B., Ennis, E., Lapsley, C., Murray, E., & O'Neill, S. (2019). Coping, stress, and negative childhood experiences: The link to psychopathology, self-harm, and suicidal behavior. *PsyCh Journal*, *8*(3), 293-306. https://doi.org/10.1002/pchj.301

Mevissen, L., Didden, R., & de Jongh, A. (2016). Assessment and treatment of PTSD in people with intellectual disabilities. *Comprehensive Guide to Post-traumatic Stress Disorder*, 281-299. https://doi.org/10.1007/978-3-319-08613-2_95-2

Milstein, N., & Gordon, I. (2020). Validating measures of electrodermal activity and heart rate variability derived from the empatica E4 utilized in research settings that involve interactive dyadic states. *Frontiers in Behavioral Neuroscience*, *14*, 148. https://doi.org/10.3389/fnbeh.2020.00148

Morgan, C., Webb, R. T., Carr, M. J., Kontopantelis, E., Green, J., Chew-Graham, C. A., Kapur, N., & Ashcroft, D. M. (2017). Incidence, clinical management, and mortality risk following self harm among children and adolescents: cohort study in primary care. *BMJ*, *359*. https://doi.org/10.1136/bmj.j4351

Morley, D. S. (2000). Psychophysiological reactivity to stress in nail biters. *International Journal of Neuroscience*, *103*(1-4), 139-154. https://doi.org/10.3109/00207450009003260

Müerner-Lavanchy, I., Koenig, J., Reichl, C., Josi, J., Cavelti, M., & Kaess, M. (2024). The quest for a biological phenotype of adolescent non-suicidal self-injury: a machine-learning approach. *Translational Psychiatry*, *14*(1). https://doi.org/10.1038/s41398-024-02776-4

Muller, D. L., Buser, T. J., Farag, M. S., & Buser, J. K. (2020). Experiences of Unintentionally Severe Harm during Nonsuicidal Self-Injury among College Students. *Journal of College Counseling*, *23*(3), 234-246. https://doi.org/10.1002/jocc.12168

Munn, Z., Barker, T. H., Moola, S., Tufanaru, C., Stern, C., McArthur, A., Stephenson, M., & Aromataris, E. (2020). Methodological quality of case series studies: an introduction to the JBI critical appraisal tool. *JBI Evidence Synthesis*, *18*(10), 2127-2133. https://doi.org/10.11124/JBISRIR-D-19-00099

Murray-Close, D., Holterman, L. A., Breslend, N. L., & Sullivan, A. (2017). Psychophysiology of proactive and reactive relational aggression. *Biological Psychology*, *130*, 77-85. https://doi.org/10.1016/j.biopsycho.2017.10.005

Naoum, J., Kleindienst, N., Baumgartner, U., Willis, F., Mancke, F., Treede, R. D., Bohus, M., & Schmahl, C. (2019). Effects of a Painful Stimulus on Stress Regulation in Male Patients With Borderline Personality Disorder: A Pilot Study. *Journal of Personality Disorders*, *33*(3), 394-412. https://dx.doi.org/10.1521/pedi_2018_32_351

Naoum, J., Reitz, S., Krause-Utz, A., Kleindienst, N., Willis, F., Kuniss, S., Baumgärtner, U., Mancke, F., Treede, R.-D., & Schmahl, C. (2016). The role of seeing blood in non-suicidal self-injury in female patients with borderline personality disorder. *Psychiatry Research*, *246*, 676-682. https://doi.org/10.1016/j.psychres.2016.10.066

National Collaborating Centre for Mental Health. (2015). National Institute for Health and Care Excellence: Guidelines. In *Violence and Aggression: Short-Term Management in Mental Health, Health and Community Settings: Updated edition*. British Psychological Society (UK).

Neimeijer, E. G., Delforterie, M. J., Roest, J. J., van der Helm, P., & Didden, R. (2021). Group climate, aggressive incidents and coercion in a secure forensic setting for individuals with mild intellectual disability or borderline intellectual functioning: A multilevel study. *Journal of Applied Research in Intellectual Disabilities*, *34*(4), 1026-1036. https://doi.org/10.1111/jar.12841

Nelson, B. W., Pollak, O. H., Clayton, M. G., Telzer, E. H., & Prinstein, M. J. (2023). An RDoC-based approach to adolescent self-injurious thoughts and behaviors: The interactive role of social affiliation and cardiac arousal. *Development and Psychopathology*, 1-11. https://dx.doi.org/10.1017/S0954579423000251

Nezamabadi, K., Sardaripour, N., Haghi, B., & Forouzanfar, M. (2022). Unsupervised ECG analysis: A review. *IEEE Reviews in Biomedical Engineering*, *16*, 208-224. https://doi.org/10.1109/RBME.2022.3154893

Nieuwenhuis, J. G., Lepping, P., Mulder, C. L., Nijman, H. L. I., & Noorthoorn, E. O. (2022). Aggressive behaviour of psychiatric patients with mild and borderline intellectual disabilities in general mental health care. *PloS one*, *17*(10), e0272502. *https://doi.org/10.1371/journal.pone.0272502*

Nieuwenhuis, J. G., Lepping, P., Mulder, N. L., Nijman, H. L., Veereschild, M., & Noorthoorn, E. O. (2021). Increased prevalence of intellectual disabilities in higher-intensity mental healthcare settings. *BJPsych Open*, *7*(3), e83. https://doi.org/10.1192/bjo.2021.28

Nijhof, K., de Looff, P., Nijman, H., van Loon, M., Klop, J., Strijbosch, E., Konijn, C., Popma, A., & Noordzij, M. (2023). Ambulatory physiological stress measurements prior to aggression incidents in residential youth care. https://doi.org/10.31234/osf.io/auwyq

Nijman, H., Jakobs, A., Waters, M., Osborne, N., Moerbeek, M., Herstel, A., & de Looff, P. (2023). A randomized crossover study on the physiological arousal reducing effects of music in forensic psychiatry. *Psychology of Music*, *51*(3), 764-781. https://doi.org/10.1177/03057356221111823

Nock, M. K. (2010). *Self-Injury* (Vol. 6). https://doi.org/10.1146/annurev.clinpsy.121208.131258

Nock, M. K., & Mendes, W. B. (2008). Physiological Arousal, Distress Tolerance, and Social Problem-Solving Deficits Among Adolescent Self-Injurers. *Journal of Consulting and Clinical Psychology*, *76*(1), 28-38. https://dx.doi.org/10.1037/0022-006X.76.1.28

Noordzij, M. L., & Laroy-Noordzij, M. (2015). Moodradar-david: Providing caretakers with real-time, personalized insight in electrodermal activity of clients who show severely challenging behavior. *Journal of Intellectual Disability Research*, *59*(SUPPL. 1), 100. https://dx.doi.org/10.1111/jir.12214 (10th International Congress of the Eamhid. Florence Italy.)

Noordzij, M. L., & Laroy-Noordzij, M. E. (2013). Longitudinal electrodermal recordings of mentally disabled individuals and their caretakers.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural equation modeling: A multidisciplinary Journal*, *14*(4), 535-569. https://doi.org/10.1080/10705510701575396

O'Shea, L. E., Picchioni, M. M., Mason, F. L., Sugarman, P. A., & Dickens, G. L. (2014). Predictive validity of the HCR-20 for inpatient self-harm. *Comprehensive Psychiatry*, *55*(8), 1937-1949. https://doi.org/10.1016/j.comppsych.2014.07.010

O'Donnell, O., House, A., & Waterman, M. (2015). The co-occurrence of aggression and self-harm: Systematic literature review. *Journal of Affective Disorders*, *175*, 325-350. https://doi.org/10.1016/j.jad.2014.12.051

Oldenhof, H., Prätzlich, M., Ackermann, K., Baker, R., Batchelor, M., Baumann, S., Bernhard, A., Clanton, R., Dikeos, D., & Dochnal, R. (2019). Baseline autonomic nervous system activity in female children and adolescents with conduct disorder: Psychophysiological findings from the FemNAT-CD study. *Journal of Criminal Justice*, *65*, 101564. https://doi.org/10.1016/j.jcrimjus.2018.05.011

Oliver, P., Crawford, M., Rao, B., Reece, B., & Tyrer, P. (2007). Modified Overt Aggression Scale (MOAS) for people with intellectual disability and aggressive challenging behaviour: a reliability study. *Journal of Applied Research in Intellectual Disabilities*, *20*(4), 368-372. https://doi.org/10.1111/j.1468-3148.2006.00346.x

Ortiz, J., & Raine, A. (2004). Heart rate level and antisocial behavior in children and adolescents: A meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*(2), 154-162. https://doi.org/10.1097/00004583-200402000-00010

Osborne, N., Ashcroft, R., Robertson, P., & Kingsley, P. (2017). *Method and system for analysing sound* (U.S. Patent No. 9736603B2). https://patents.google.com/patent/US9736603B2/en

Otto, A., Jarvers, I., Kandsperger, S., Reichl, C., Ando, A., Koenig, J., Kaess, M., & Brunner, R. (2023). Stress-induced alterations in resting-state functional connectivity among adolescents with non-suicidal self-injury. *Journal of Affective Disorders*, *339*, 162-171. https://doi.org/10.1016/j.jad.2023.07.032

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*. https://doi.org/10.1136/bmj.n71

Påhlsson-Notini, A., Liu, S., Tideman, M., Latvala, A., Serlachius, E., Larsson, H., Hirvikoski, T., Taylor, M. J., Kuja-Halkola, R., & Lichtenstein, P. (2024). Substance use-related problems in mild intellectual disability: A Swedish nationwide population-based cohort study with sibling comparison. *JCPP Advances*, e12225. https://doi.org/10.1002/jcv2.12225

Papadopoulos, C., Ross, J., Stewart, D., Dack, C., James, K., & Bowers, L. (2012). The antecedents of violence and aggression within psychiatric in-patient settings. *Acta Psychiatrica Scandinavica*, *125*(6), 425-439. https://doi.org/10.1111/j.1600-0447.2012.01827.x

Patlar Akbulut, F., Ikitimur, B., & Akan, A. (2020). Wearable sensor-based evaluation of psychosocial stress in patients with metabolic syndrome. *Artificial Intelligence in Medicine*, *104*, 101824. https://doi.org/10.1016/j.artmed.2020.101824

Pattyn, E., Lutin, E., Van Kraaij, A., Thammasan, N., Tourolle, D., Kosunen, I., Tump, D., De Raedt, W., & Van Hoof, C. (2023). Annotation-Based Evaluation of Wrist EDA Quality and Response Assessment Techniques. 16th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2023) -Biosignals,

Pelto-Piri, V., Warg, L.-E., & Kjellin, L. (2020). Violence and aggression in psychiatric inpatient care in Sweden: a critical incident technique analysis of staff descriptions. *BMC Health Services Research*, *20*, 1-11. https://doi.org/10.1186/s12913-020-05239-w

Peltopuro, M., Ahonen, T., Kaartinen, J., Seppälä, H., & Närhi, V. (2014). Borderline intellectual functioning: a systematic literature review. *Intellectual and Developmental Disabilities*, *52*(6), 419-443. https://doi.org/10.1352/1934-9556-52.6.419

Pham, T., Lau, Z. J., Chen, S. H. A., & Makowski, D. (2021). Heart Rate Variability in Psychology: A Review of HRV Indices and an Analysis Tutorial. *Sensors*, *21*(12), 3998. https://doi.org/10.3390/s21123998

Philips. (n.d.). *A unique solution for measuring stress.* Retrieved January 20, 2024 from https://www.philips.com/a-w/about/innovation/ips/ip-licensing/programs/biosensing-by-eda.html

Plotner, A. J., Ezell, J. A., VanHorn-Stinnitt, C., & Rose, C. A. (2020). An investigation of coping strategies used by college students with intellectual disability. *Inclusion*, *8*(3), 194-209. https://doi.org/10.1352/2326-6988-8.3.194

Popma, A., & Raine, A. (2006). Will future forensic assessment be neurobiologic? *Child and Adolescent Psychiatric Clinics*, *15*(2), 429-444. http://dx.doi.org/10.1016/j.chc.2005.11.004

Porges, S. W. (2007). The polyvagal perspective. *Biological Psychology*, *74*(2), 116-143. https://doi.org/10.1016/j.biopsycho.2006.06.009

Porta, M. (2014). *A dictionary of epidemiology*. Oxford university press.

Portnoy, J., & Farrington, D. P. (2015). Resting heart rate and antisocial behavior: An updated systematic review and meta-analysis. *Aggression and Violent Behavior*, *22*, 33-45. https://doi.org/10.1016/j.avb.2015.02.004

Preyde, M., Watkins, H., Csuzdi, N., Carter, J., Lazure, K., White, S., Penney, R., Ashbourne, G., Cameron, G., & Frensch, K. (2012). Non-suicidal self-injury and suicidal behaviour in children and adolescents accessing residential or intensive home-based mental health services. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *21*(4), 270.

Pulopulos, M. M., Baeken, C., & De Raedt, R. (2020). Cortisol response to stress: The role of expectancy and anticipatory stress regulation. *Hormones and Behavior*, *117*, 104587. https://doi.org/10.1016/j.yhbeh.2019.104587

Quigley, J. M., Thompson, J. C., Halfpenny, N. J., & Scott, D. A. (2019). Critical appraisal of nonrandomized studies—a review of recommended and commonly used tools. *Journal of Evaluation in Clinical Practice*, *25*(1), 44-52. https://doi.org/10.1111/jep.12889

Reitz, S., Kluetsch, R., Niedtfeld, I., Knorz, T., Lis, S., Paret, C., Kirsch, P., Meyer-Lindenberg, A., Treede, R. D., Baumgartner, U., Bohus, M., & Schmahl, C. (2015). Incision and stress regulation in borderline personality disorder: Neurobiological mechanisms of self-injurious behaviour. *British Journal of Psychiatry*, *207*(2), 165-172. https://dx.doi.org/10.1192/bjp.bp.114.153379

Reitz, S., Krause-Utz, A., Pogatzki-Zahn, E. M., Ebner-Priemer, U., Bohus, M., & Schmahl, C. (2012). Stress regulation and incision in borderline personality disorder--a pilot study modeling cutting behavior. *Journal of Personality Disorders*, *26*(4), 605-615. https://doi.org/10.1521/pedi.2012.26.4.605

Reyes del Paso, G. A., Langewitz, W., Mulder, L. J., Van Roon, A., & Duschek, S. (2013). The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: a review with emphasis on a reanalysis of previous studies. *Psychophysiology*, *50*(5), 477-487. https://doi.org/10.1111/psyp.12027

Richter, D., & Whittington, R. (2006). *Violence in mental health settings*. Springer.

Rippon, T. J. (2000). Aggression and violence in health care professions. *Jour-*

nal of Advanced Nursing, *31*(2), 452-460. https://doi.org/10.1046/j.1365-2648.2000.01284.x

Roberton, T., Daffern, M., & Bucks, R. S. (2012). Emotion regulation and aggression. *Aggression and Violent Behavior*, *17*(1), 72-82. https://doi.org/10.1016/j.avb.2011.09.006

Robson, H., & Attard, C. (2019). Staff burnout and staff turnover on inpatient wards. *Oxford Textbook of Inpatient Psychiatry*, 277.

Rouleaux, M., Peters-Scheffer, N., Lindauer, R., Mevissen, L., & Didden, R. (2024). Stressful life events, post-traumatic stress disorder symptoms and mental health in individuals with intellectual disabilities: a scoping review. *Journal of Intellectual Disability Research*, *68*(10), 1087-1113. https://doi.org/10.1111/jir.13178

Sabry, F., Eltaras, T., Labda, W., Alzoubi, K., & Malluhi, Q. (2022). Machine learning for healthcare wearable devices: the big picture. *Journal of Healthcare Engineering*, *2022*(1), 4653923. https://doi.org/10.1155/2022/4653923

Samways, B., Heslop, P., & Dowling, S. (2024). Self-reported explanations for self-injury by people with intellectual disabilities: a systematic review of qualitative studies. *International Journal of Developmental Disabilities*, *70*(3), 382-396. https://doi.org/10.1080/20473869.2022.2098665

Schalock, R. L. (2011). The evolving understanding of the construct of intellectual disability. *Journal of Intellectual and Developmental Disability*, *36*(4), 227-237. https://doi.org/10.3109/13668250.2011.624087

Schoenmakers, M., Saygin, M., Sikora, M., Vaessen, T., Noordzij, M., & de Geus, E. (2025). Stress in action wearables database: A database of noninvasive wearable monitors with systematic technical, reliability, validity, and usability information. *Behavior Research Methods*, *57*(6), 1-23. https://doi.org/10.3758/s13428-025-02685-4

Schoorl, J., Van Rijn, S., De Wied, M., Van Goozen, S. H., & Swaab, H. (2016). Variability in emotional/behavioral problems in boys with oppositional defiant disorder or conduct disorder: the role of arousal. *European Child & Adolescent Psychiatry*, *25*, 821-830. https://doi.org/10.1007/s00787-015-0790-5

Schouten, S. E., Kip, H., Dekkers, T., Deenik, J., Beerlage-de Jong, N., Ludden, G. D., & Kelders, S. M. (2022). Best-practices for co-design processes involving people with severe mental illness for eMental health interventions: a qualitative multi-method approach. *Design for Health*, *6*(3), 316-344. https://doi.org/10.1080/24735132.2022.2145814

Schulz, A., & Vögele, C. (2015). Interoception and stress. *Frontiers in Psychology*, *6*, 133987. https://doi.org/10.3389/fpsyg.2015.00993

Schuurmans, A. A., De Looff, P., Nijhof, K. S., Rosada, C., Scholte, R. H., Popma, A., & Otten, R. (2020). Validity of the Empatica E4 wristband to measure heart rate variability (HRV) parameters: A comparison to electrocardiography (ECG). *Journal of Medical Systems*, *44*, 1-11. https://doi.org/10.1007/s10916-020-01648-w

Selenius, H., & Strand, S. (2017). Experiences of self-injury and aggression among women admitted to forensic psychiatric care. *Nordic Journal of Psychiatry*, *71*(4), 304-311. https://doi.org/10.1080/08039488.2017.1283443

Shaffer, F., McCraty, R., & Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. *Frontiers in Psychology*, *5*, 1040. https://doi.org/10.3389/fpsyg.2014.01040

Shafti, M., Taylor, P. J., Forrester, A., & Pratt, D. (2021). The Co-occurrence of Self-Harm and Aggression: A Cognitive-Emotional Model of Dual-Harm. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.586135

Shiner, A. (2008). Self-harm in adolescence. *InnovAiT*. https://doi.org/10.1093/innovait/inn119

Sigrist, C., Jakob, H., Beeretz, C. J., Schmidt, S. J., Kaess, M., & Koenig, J. (2024). Diurnal variation of cardiac autonomic activity in adolescent non-suicidal self-injury. *European Archives of Psychiatry and Clinical Neuroscience*, *274*(3), 609-628. https://doi.org/10.1007/s00406-023-01574-1

Simons, R., Koordeman, R., de Looff, P., & Otten, R. (2021). Physiological measurements of stress preceding incidents of challenging behavior in people with severe to profound intellectual disabilities: Longitudinal study protocol of single-case studies. *JMIR Research Protocols*, *10*(7), e24911. https://doi.org/10.2196/24911

Slade, K., Forrester, A., & Baguley, T. (2020). Coexisting violence and self-harm: Dual harm in an early-stage male prison population. *Legal and Criminological Psychology*, *25*(2), 182-198. https://dx.doi.org/10.1111/lcrp.12169

Staddon, J. E. (2001). *Adaptive dynamics: The theoretical analysis of behavior*. MIT press.

Stang, A. (2010). Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *European Journal of Epidemiology*, *25*, 603-605. https://doi.org/10.1007/s10654-010-9491-z

Steinert, T., & Whittington, R. (2013). A bio-psycho-social model of violence related to mental health problems. *International Journal of Law and Psychiatry*, *36*(2), 168-175. https://doi.org/10.1016/j.ijlp.2013.01.009

Stevenson, C., & Taylor, J. (2020). Nurses' perspectives of factors that influence therapeutic relationships in secure inpatient forensic hospitals. *Journal of Forensic nursing*, *16*(3), 169-178. https://doi.org/10.1097/JFN.0000000000000274

Tassé, M. J., Schalock, R. L., Balboni, G., Bersani Jr, H., Borthwick-Duffy, S. A., Spreat, S., Thissen, D., Widaman, K. F., & Zhang, D. (2012). The construct of adaptive behavior: Its conceptualization, measurement, and use in the field of intellectual disability. *American Journal on Intellectual and Developmental Disabilities*, *117*(4), 291-303. https://doi.org/10.1352/1944-7558-117.4.291

Tatnell, R., Hasking, P., Lipp, O. V., Boyes, M., & Dawkins, J. (2018). Emotional responding in NSSI: examinations of appraisals of positive and negative emotional stimuli, with and without acute stress. *Cognition and Emotion*, *32*(6), 1304-1316. https://dx.doi.org/10.1080/02699931.2017.1411785

Taylor, J. L., & Novaco, R. W. (2005). *Anger treatment for people with developmental disabilities: A theory, evidence and manual based approach*. John Wiley & Sons.

Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., & Picard, R. (2015). Automatic identification of artifacts in electrodermal activity data. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),

Teodosio, G., & Ugo, S. (2019). Psychoanalysis and Affective Neuroscience. The Motivational/Emotional System of Aggression in Human Relations. *Frontiers in Psychology*, *9*. http://dx.doi.org/10.3389/fpsyg.2018.02475

Ter Harmsel, J. F., Noordzij, M. L., Goudriaan, A. E., Dekker, J. J., Swinkels, L. T., van der Pol, T. M., & Popma, A. (2021). Biocueing and ambulatory biofeedback to enhance emotion regulation: a review of studies investigating non-psychiatric and psychiatric populations. *International Journal of Psychophysiology*, *159*, 94-106. https://doi.org/10.1016/j.ijpsycho.2020.11.009

Ter Harmsel, J. F., Noordzij, M. L., van der Pol, T. M., Swinkels, L. T., Goudriaan, A. E., & Popma, A. (2023). Exploring the effects of a wearable biocueing app (Sense-IT) as an addition to aggression regulation therapy in forensic psychiatric outpatients. *Frontiers in Psychology*, *14*, 983286. https://doi.org/10.3389/fpsyg.2023.983286

Ter Harmsel, J. F., Smulders, L. M., Noordzij, M. L., Swinkels, L. T., Goudriaan, A. E., Popma, A., & van der Pol, T. M. (2023). Forensic psychiatric outpatients' and therapists' perspectives on a wearable biocueing app (Sense-IT) as an addition to aggression regulation therapy: qualitative focus group and interview study. *JMIR Formative Research*, *7*, e40237. https://doi.org/10.2196/40237

Theis, N., Campbell, N., De Leeuw, J., Owen, M., & Schenke, K. C. (2021). The effects of COVID-19 restrictions on physical activity and mental health of children and young adults with physical and/or intellectual disabilities. *Disability and Health Journal*, *14*(3), 101064. https://doi.org/10.1016/j.dhjo.2021.101064

Tomlinson, M., Ward, C. L., & Marlow, M. (2015). Improving the efficiency of evidence-based interventions: The strengths and limitations of randomised controlled trials. *South African Crime Quarterly*, *51*, 43-52. https://doi.org/10.4314/sacq.v51i0.5

Tuna, E., & Gencoz, T. (2021). Pain perception, distress tolerance and self-compassion in Turkish young adults with and without a history of non-suicidal self-injury. *Current Psychology*, *40*(8), 4143-4155. https://doi.org/10.1007/s12144-020-00634-2

Turner, A. I., Smyth, N., Hall, S. J., Torres, S. J., Hussein, M., Jayasinghe, S. U., Ball, K., & Clow, A. J. (2020). Psychological stress reactivity and future health and disease outcomes: A systematic review of prospective evidence. *Psychoneuroendocrinology*, *114*, 104599. https://doi.org/https://doi.org/10.1016/j.psyneuen.2020.104599

Turner, B. J., Arya, S., & Chapman, A. L. (2015). Nonsuicidal self-injury in Asian versus

Caucasian university students: Who, how, and why? *Suicide and Life-Threatening Behavior*, *45*(2), 199-216. https://doi.org/10.1111/sltb.12113

Tutunji, R., Kogias, N., Kapteijns, B., Krentz, M., Krause, F., Vassena, E., & Hermans, E. J. (2023). Detecting Prolonged Stress in Real Life Using Wearable Biosensors and Ecological Momentary Assessments: Naturalistic Experimental Study. *Journal of Medical Internet Research*, *25*, e39995. https://doi.org/10.2196/39995

Ulrich-Lai, Y. M., & Herman, J. P. (2009). Neural regulation of endocrine and autonomic stress responses. *Nature Reviews Neuroscience*, *10*(6), 397-409. https://doi.org/10.1038/nrn2647

Ursin, H., & Eriksen, H. R. (2010). Cognitive activation theory of stress (CATS). *Neuroscience & Biobehavioral Reviews*, *34*(6), 877-881. https://doi.org/10.1016/j.neubiorev.2009.03.001

van den Akker, N., Kroezen, M., Wieland, J., Pasma, A., & Wolkorte, R. (2021). Behavioural, psychiatric and psychosocial factors associated with aggressive behaviour in adults with intellectual disabilities: A systematic review and narrative analysis. *Journal of Applied Research in Intellectual Disabilities*, *34*(2), 327-389. https://doi.org/https://doi.org/10.1111/jar.12809

Van der Mee, D. J., Gevonden, M. J., Westerink, J. H. D. M., & de Geus, E. J. C. (2021). Validity of electrodermal activity-based measures of sympathetic nervous system activity from a wrist-worn device. *International Journal of Psychophysiology*, *168*, 52-64. https://doi.org/10.1016/j.ijpsycho.2021.08.003

van der Venne, P., Murner-Lavanchy, I., Hoper, S., Koenig, J., & Kaess, M. (2023). Physiological response to pain in female adolescents with nonsuicidal self-injury as a function of severity. *Journal of Affective Disorders*, *339*, 64-73. https://doi.org/https://dx.doi.org/10.1016/j.jad.2023.06.041

Van der Zwaag, M. D., Janssen, J. H., & Westerink, J. H. (2012). Directing physiology and mood through music: Validation of an affective music player. *IEEE Transactions on Affective Computing*, *4*(1), 57-68. https://doi.org/10.1109/T-AFFC.2012.28

van Duijvenbode, N., VanDerNagel, J. E., Didden, R., Engels, R. C., Buitelaar, J. K., Kiewik, M., & de Jong, C. A. (2015). Substance use disorders in individuals with mild to borderline intellectual disability: current status and future directions. *Research in Developmental Disabilities*, *38*, 319-328. https://doi.org/10.1016/j.ridd.2014.12.029

van Lier, H. G., Pieterse, M. E., Garde, A., Postel, M. G., de Haan, H. A., Vollenbroek-Hutten, M. M. R., Schraagen, J. M., & Noordzij, M. L. (2020). A standardized validity assessment protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor. *Behavior Research Methods*, *52*(2), 607-629. https://doi.org/10.3758/s13428-019-01263-9

van Swieten, M., de Looff, P., VanDerNagel, J., & Didden, R. (2024). The association between aggressive behaviour and non-suicidal self-injury and shared risk factors in adults with mild intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, *37*(6), e13288. https://doi.org/10.1111/jar.13288

Vansteelandt, K., Houben, M., Claes, L., Berens, A., Sleuwaegen, E., Sienaert, P., & Kuppens, P. (2017). The affect stabilization function of nonsuicidal self injury in borderline personality disorder: an ecological momentary assessment study. *Behaviour Research and Therapy*, *92*, 41-50. https://doi.org/10.1016/j.brat.2017.02.003

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, *18*(4), 450-469. https://doi:10.1093/pan/mpq025

Verona, E., & Sullivan, E. A. (2008). Emotional catharsis and aggression revisited: heart rate reduction following aggressive responding. *Emotion*, *8*(3), 331. https://psycnet.apa.org/doi/10.1037/1528-3542.8.3.331

Verstegen, N., de Vogel, V., Huitema, A., Didden, R., & Nijman, H. (2020). Physical Violence During Mandatory Psychiatric Treatment: Prevalence and Patient Characteristics. *Criminal Justice and Behavior*, *47*(7), 771-789. https://doi.org/10.1177/0093854820924691

Verstegen, N., Peters-Scheffer, N., Didden, R., & de Vogel, V. (2024). 'You Still Need to Care for the patient'. Experiences of Forensic Mental Health Professionals with Inpatient Aggression. *International Journal of Forensic Mental Health*, 1-12. https://doi.org/10.1080/14999013.2024.2354286

Vruwink, F. J., VanDerNagel, J. E., Noorthoorn, E. O., Nijman, H. L., & Mulder, C. L. (2022). "Disruptive Behavior" or "Expected Benefit" Are Rationales of Seclusion Without Prior Aggression. *Frontiers in Psychiatry*, *13*, 871525. https://doi.org/10.3389/fpsyt.2022.871525

Wagemaker, E., Huizenga, H. M., Dekkers, T. J., Collot d'Escury-Koenigs, A. L., Salemink, E., & Bexkens, A. (2020). When do those "risk-taking adolescents" take risks? The combined effects of risk encouragement by peers, mild-to-borderline intellectual disability and sex. *Journal of Abnormal Child Psychology*, *48*, 573-587. https://doi.org/10.1007/s10802-020-00617-8

Warburton, W. A., & Anderson, C. A. (2015). Aggression, social psychology of. *International Encyclopedia of the Social & Behavioral Sciences*, *1*(2), 373-380.

Weinberg, A., & Klonsky, E. D. (2012). The effects of self-injury on acute negative arousal: A laboratory simulation. *Motivation and Emotion*, *36*(2), 242-254. https://doi.org/doi:https://dx.doi.org/10.1007/s11031-011-9233-x

Welch, S. S., Linehan, M. M., Sylvers, P., Chittams, J., & Rizvi, S. L. (2008). Emotional Responses to Self-Injury Imagery Among Adults With Borderline Personality Disorder. *Journal of Consulting and Clinical Psychology*, *76*(1), 45-51. https://doi.org/doi:https://dx.doi.org/10.1037/0022-006X.76.1.45

Wells, G., Shea, B., O'connell, D., Petersen, J., Welch, V., Losos, M., & Tugwell, P. (2011). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses*. https://www.ohri.ca/programs/clinical_epi-

demiology/oxford.asp
Wells, J., Haines, J., Williams, C., & Brain, K. (1999). The self-mutilative nature of severe onychophagia: A comparison with self-cutting. *Canadian Journal of Psychiatry*, *44*(1), 40-47. https://doi.org/doi:10.1177/070674379904400105
Weltens, I., Bak, M., Verhagen, S., Vandenberk, E., Domen, P., van Amelsvoort, T., & Drukker, M. (2021). Aggression on the psychiatric ward: Prevalence and risk factors. A systematic review of the literature. *PloS one*, *16*(10), e0258346. https://doi.org/10.1371/journal.pone.0258346
Westerink, J. H. D. M., Rajae-Joordens, R. J. E., Ouwerkerk, M., van Dooren, M., Jelfs, S., Denissen, A. J. M., de Vries, E., & van Ee, R. (2020). Deriving a Cortisol-Related Stress Indicator From Wearable Skin Conductance Measurements: Quantitative Model & Experimental Validation. *Frontiers in Computer Science*, *2*, 1-11. https://doi.org/10.3389/fcomp.2020.00039
Whiteside, S. P., & Lynam, D. R. (2001). The Five Factor Model and impulsivity: using a structural model of personality to understand impulsivity. *Personality and Individual Differences*, *30*(4), 669-689. https://doi.org/10.1016/S0191-8869(00)00064-7
Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://doi.org/https://ggplot2.tidyverse.org
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., Grolemund, G., Hayes, A., Henry, L., & Hester, J. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686
Wielgus, M., Aldrich, J., Mezulis, A., & Crowell, S. (2016). Respiratory sinus arrhythmia as a predictor of self-injurious thoughts and behaviors among adolescents. *International Journal of Psychophysiology*, *106*, 127-134. https://doi.org/doi:10.1016/j.ijpsycho.2016.05.005
Wigham, S., & Emerson, E. (2015). Trauma and life events in adults with intellectual disability. *Current Developmental Disorders Reports*, *2*, 93-99. https://doi.org/10.1007/s40474-015-0041-y
Wilder, J. (1950). The law of initial values. Psychosomatic Medicine,
Wilkowski, B. M., & Robinson, M. D. (2010). The anatomy of anger: An integrative cognitive model of trait anger and reactive aggression. *Journal of Personality*, *78*(1), 9-38. https://doi.org/10.1111/j.1467-6494.2009.00607.x
Willis, F., Kuniss, S., Kleindienst, N., Lis, S., Naoum, J., Jungkunz, M., Neukel, C., Bohus, M., Treede, R., Baumgartner, U., & Schmahl, C. (2018). Stress reactivity and pain-mediated stress regulation in remitted patients with borderline personality disorder. *Brain and Behavior*, *8*(2), e00909. https://doi.org/10.1002/brb3.909
Willis, F., Kuniss, S., Kleindienst, N., Naoum, J., Reitz, S., Boll, S., Bohus, M., Treede, R., Baumgartner, U., & Schmahl, C. (2017). The role of nociceptive input and tissue injury on stress regulation in borderline personality disorder. *PAIN*, *158*(3), 479-487. https://doi:10.1097/j.pain.0000000000000787

Woittiez, I., Eggink, E., & Ras, M. (2019). Het aantal mensen met een licht verstandelijke beperking: een schatting.

Wolff, J. C., Thompson, E., Thomas, S. A., Nesi, J., Bettis, A. H., Ransford, B., Scopelliti, K., Frazier, E. A., & Liu, R. T. (2019). Emotion dysregulation and non-suicidal self-injury: A systematic review and meta-analysis. *European Psychiatry*, *59*, 25-36. https://doi:10.1016/j.eurpsy.2019.03.004

Woodward, K., Kanjo, E., Brown, D. J., McGinnity, T., & Harold, G. (2023). In the hands of users with intellectual disabilities: co-designing tangible user interfaces for mental wellbeing. *Personal and Ubiquitous Computing*, *27*(6), 2171-2191. https://doi.org/10.1007/s00779-023-01752-x

Yagin, F. H., Pinar, A., & de Sousa Fernandes, M. S. (2024). Statistical Effect Sizes in Sports Science. *Journal of Exercise Science & Physical Activity Reviews*, *2*(1), 164-171. https://doi.org/10.5281/zenodo.12601138%20

Yu, B., Funk, M., Hu, J., Wang, Q., & Feijs, L. (2018). Biofeedback for everyday stress management: A systematic review. *Frontiers in ICT*, *5*, 23. https://doi.org/10.3389/fict.2018.00023

Zhang, S., Hu, S., Chao, H. H., Ide, J. S., Luo, X., Farr, O. M., & Li, C.-s. R. (2014). Ventromedial prefrontal cortex and the regulation of physiological arousal. *Social Cognitive and Affective Neuroscience*, *9*(7), 900-908. https://doi.org/10.1093/scan/nst064

Zijlmans, J., Marhe, R., van Duin, L., Luijks, M.-J. A., Bevaart, F., & Popma, A. (2021). No association between autonomic functioning and psychopathy and aggression in multi-problem young adults. *Frontiers in Psychology*, *12*, 645089. https://doi.org/10.3389/fpsyg.2021.645089

Zillmann, D. (1971). Excitation transfer in communication-mediated aggressive behavior. *Journal of Experimental Social Psychology*, *7*(4), 419-434. https://doi.org/10.1016/0022-1031(71)90075-8

Zillmann, D., Katcher, A. H., & Milavsky, B. (1972). Excitation transfer from physical exercise to subsequent aggressive behavior. *Journal of Experimental Social Psychology*, *8*(3), 247-259. https://doi.org/10.1016/S0022-1031(72)80005-2

Appendix

Dutch summary (Nederlandse samenvatting)

Onderzoeksresultaten

Agressief gedrag en zelfbeschadiging komen vaak voor bij mensen met een licht verstandelijke beperking of zwakbegaafdheid (LVB) in (forensische behandel-) instellingen. Beide hebben een negatieve impact op zowel de cliënt, andere cliënten als zorgverleners. Agressief gedrag en zelfbeschadiging komen vaak samen voor. Er was echter minder bekend over het samen voorkomen en overlappende risicofactoren bij volwassenen met LVB. In **hoofdstuk 2** onderzochten we daarom eerst de associatie tussen agressief gedrag en niet-suïcidale zelfbeschadiging (NSSI) en overlappende risicofactoren, bij 125 volwassenen met LVB. Dit werd onderzocht aan de hand van incidentrapportages en metingen die onderdeel zijn van het reguliere behandelproces. De resultaten toonden een zwakke correlatie tussen agressief gedrag en NSSI, tussen impulsiviteit en agressief gedrag, evenals tussen coping vaardigheden en agressief gedrag en NSSI. NSSI, impulsiviteit en coping vaardigheden voorspelden agressief gedrag echter niet. Dit laatste is niet in lijn met eerdere studies. Dit zou verklaard kunnen worden door verschillen tussen mensen met en zonder LVB, of het gebruik van andere meetinstrumenten dan die in eerdere studies. Deze resultaten kunnen wijzen op het belang om, naast coping vaardigheden en impulsiviteit, ook andere overlappende risicofactoren van agressief gedrag en zelfbeschadiging in kaart te brengen, zoals arousal.

Verhoogde en verlaagde arousal zijn herhaaldelijk in verband gebracht met agressief gedrag en zelfbeschadiging. Er wordt steeds meer onderzoek gedaan naar *fysiologische* arousal (bijvoorbeeld hartslag) in relatie tot agressief gedrag en zelfbeschadiging. Meerdere systematische reviews onderzochten de relatie tussen agressief gedrag en arousal. Er bestond echter nog geen uitgebreid overzicht van studies naar de relatie tussen zelfbeschadiging en arousal. In **hoofdstuk 3** hebben we daarom een systematische review uitgevoerd van studies naar de associatie tussen fysiologische parameters en zelfbeschadiging. Er werden 46 studies met uiteenlopende onderzoeksopzetten, participanten en metingen opgenomen. De resultaten lieten zien dat de meeste studies, die huidgeleiding en hartslag vergeleken bij mensen met en zonder zelfbeschadiging, geen duidelijke aanwijzingen vonden voor een relatie tussen fysiologische parameters (arousal) en zelfbeschadiging. Studies naar hartslagvariatie toonden aanwijzingen voor een lagere hartslagvariatie bij mensen met zelfbeschadiging tijdens herstel na bijvoorbeeld stressverhogende taken in vergelijking tot mensen zonder zelfbeschadiging. Dit kan duiden op problemen in de emotieregulatie, wat werd ondersteund door resultaten van imaginatiestudies waarbij hartslag en huidgeleiding werden gemeten. Studies die fysiologische parameters onderzochten op momenten van daadwerkelijke zelfbeschadiging vonden tegenstrijdige resultaten, maar dit is in zeer weinig studies onderzocht. Hoewel de draagbare technologie (wearables) om fysiologie te meten is verbeterd, bleek dat de meerderheid van de studies tot nu toe laboratoriumstudies zijn.

Ook voor agressief gedrag zijn de meeste onderzoeken naar de relatie met fysiologische arousal tot nu toe voornamelijk laboratoriumstudies. Voor zover bekend was er nog geen onderzoek in de klinische praktijk gedaan naar fysiologische parameters bij volwassen met LVB-BIF na agressie incidenten. Inzicht in het verloop van fysiologische arousal na agressief gedrag zou belangrijke kennis kunnen opleveren voor de preventie van terugkerend agressief gedrag. In **hoofdstuk 4** onderzochten we daarom met een grote multicenter studie veranderingen in hartslag en huidgeleiding in de 30 minuten na agressief gedrag, gemeten op de afdeling. Arousal, gemeten met draagbare technologie, werd na 165 agressie incidenten geanalyseerd. De resultaten toonden zoals voorspeld een significante afname in hartslag en huidgeleiding na agressie incidenten. De afname in arousal was echter klein en er was een grote variatie tussen en binnen personen. Verdere exploratieve analyse (latente groeicurve-analyse) van het verloop van arousal na agressie toonde aan dat er verschillende patronen konden worden onderscheiden. Naast een afname in arousal waren er ook groepen incidenten met stabiele niveaus of zelfs een toename in arousal. Het type agressief gedrag hing samen met de verschillende latente klassen waartoe incidenten behoorden. De verschillende fysiologische patronen benadrukken het belang van gepersonaliseerde benaderingen bij het omgaan met agressie incidenten. Als bijvoorbeeld bij een cliënt de arousal al is afgenomen en de cliënt is teruggekeerd naar een kalmere staat, kan het inzetten van aanvullende interventies onnodig of zelfs contraproductief zijn. Aan de andere kant kunnen cliënten, waarbij na agressie de arousal langdurig verhoogd is, baat hebben bij gerichte interventies, zoals stressreducerende interventies of biofeedback.

Tot slot, omdat arousal gerelateerd is aan agressief gedrag en zelfbeschadiging, is het van belang om verhoogde arousal (tijdig) te verlagen. Met de studie in **hoofdstuk 5** van dit proefschrift onderzochten we subjectieve en fysiologische arousal tijdens een interventie voor stressreductie, namelijk muziek luisteren. Verder werd onderzocht of het gebruik van software voor het selecteren en rangschikken van muziek (gebaseerd op het voorspellen van het effect van een nummer op het autonome zenuwstelsel, X-systeem genaamd) geassocieerd was met grotere afnames in subjectieve en fysiologische arousal in vergelijking met zelfgeselecteerde muziek. Biofeedback werd gebruikt om dit te onderzoeken op stressvolle momenten in het dagelijks leven van cliënten, en omdat cliënten vaak moeite hebben om stress bij zichzelf te herkennen. Participanten werden geïnstrueerd om naar muziek te luisteren wanneer ze biofeedback kregen over hun verhoogde stressniveau (zoals gemeten door de draagbare biosensor Nowatch) of wanneer ze zich gestrest voelden. We verzamelden in totaal 103 luistersessies tijdens de dagelijkse routines van 11 participanten, over een periode van 2-4 weken. Participanten luisterden naar zelfgekozen muziek of naar een gepersonaliseerde afspeellijst die was samengesteld met X-system. De resultaten lieten een afname zien in hartslag en huidgeleiding tijdens het luisteren naar muziek. Daarnaast werden er aanwijzingen gevonden voor afname in zelf-gerapporteerde stress en verbetering van de stemming na muziek luisteren. De afname van hartslag en huidgeleiding tijdens het luisteren naar muziek was klein en er was veel variatie tussen en binnen participanten. Het luisteren naar muziek samengesteld met X-system was niet effectiever dan luisteren naar zelfgeselecteerde muziek. Echter, lagere 'gecombineerde arousal' waarden van nummers (een index van X-system) voorspelden lagere hartslag en huidgeleiding. Dit suggereert dat deze index gebruikt kan worden om nummers te selecteren die een ontspannend of activerend effect hebben. Luisteren naar muziek kan een laagdrempelige, goedkope en empowering strategie zijn voor stressvermindering, wat tevens een positieve invloed kan hebben op emotieregulatie en de mentale en lichamelijke gezondheid.

Vervolgonderzoek

Technologische ontwikkelingen in de vorm van wearables maken continue monitoring van fysiologische arousal in het dagelijks leven mogelijk. Het gebruik van wearables biedt veel mogelijkheden voor wetenschappelijk onderzoek. Er is onder andere meer onderzoek nodig waarbij fysiologie wordt gemeten met wearables in het dagelijks leven vóór, tijdens en na agressief gedrag en zelfbeschadiging. Daarnaast zijn meer studies met idiografische benaderingen nodig die meerdere databronnen integreren (o.a. verschillende fysiologie parameters, contextuele factoren, psychologische factoren en subjectieve ervaringen). Dit zal bijdragen aan de verdere ontwikkeling van het vakgebied van fysiologische arousal en de integratie van biologische gegevens in de diagnostiek en behandeling (van agressief gedrag en zelfbeschadiging) in de geestelijke gezondheidszorg, forensische psychiatrie en verstandelijk gehandicaptenzorg.

Klinische implicaties

Het gebruik van wearables biedt ook veelbelovende mogelijkheden in de klinische praktijk bij het ontrafelen van complexe vraagstukken in diagnostiek en behandeling. Zo kunnen fysiologische gegevens verzameld worden die een aanzet vormen tot een gesprek tussen cliënt en behandelaar. Daarnaast kunnen wearables in klinische settingen de aandacht van begeleiders trekken, bijvoorbeeld op momenten van verhoogde spanning bij cliënten, en hen daarmee ondersteunen bij het monitoren van hun cliënten. Verder kunnen inzichten in patronen van fysiologische arousal, zowel voorafgaand aan als na agressie, helpen om inzicht te krijgen in het ontstaan en de functie van agressie. Deze kennis kan gebruikt worden om gepersonaliseerde behandelstrategieën te ontwikkelen die gericht zijn op het aanpakken van de onderliggende oorzaken van agressie. Diagnostiek en indicatiestelling van interventies kunnen echter niet uitsluitend op fysiologische data worden gebaseerd, maar kunnen een aanvulling vormen op het klinisch oordeel van begeleiders en behandelaren.

Momenteel staan we aan het begin van veel nieuwe ontwikkelingen op dit gebied, die naar verwachting snel zullen doorgroeien. Er moeten richtlijnen worden ontwikkeld over hoe, wanneer en voor wie wearables ingezet kunnen worden in de (forensische) zorg voor mensen met psychiatrische problematiek en/of een verstandelijke beperking. Om de klinische interpretatie van de grote hoeveelheid data van wearables te verbeteren zijn tools nodig die valide data toegankelijk en visueel maken. Niet in de laatste plaats is het bij het gebruik van wearables belangrijk om ethische aspecten, privacy en gegevensbescherming zorgvuldig in acht te nemen.

Research data management and privacy statement

Ethical statement and privacy statement

The studies described in this thesis were conducted in compliance with the General Data Protection Regulation (GDPR) and all applicable laws and ethical guidelines. The studies in Chapter 2 and 3 did not require ethical assessment because we used publicly available data (systematic review, Chapter 3) or re-used existing (unpublished) data (Chapter 2). In Chapter 2 we used data from measurements which are part of the standard care during clinical treatment within Trajectum, from clients who signed informed consent for using the data for scientific research. For the study in Chapter 4, data from three studies reporting on physiology and aggressive behaviour were combined for secondary analyses (De Looff, et al., 2019; Hagoort et al., 2025; Nijhof et al., 2023). Ethical approval for the study by De Looff et al. was granted by the Ethics Committee of the Faculty of Social Sciences of the Radboud University at Nijmegen (ECSW2015-1901-282). The Medical Research Ethics Committee at UMCU, at Utrecht reviewed the study by Hagoort et al. (18-886). The study was evaluated as 'no research related medical ethical approval needed' because it imposed a minimal additional burden for the patient with no further interruption or change of their regular clinical program. The Medical Ethical Committee of Arnhem-Nijmegen reviewed the study by Nijhof et al. (2023) and evaluated the study was as complying with the Dutch Law on Medical Research in Humans, and no further ethical approval was needed (ref.no. NL63138.091.17). The study described in Chapter 5 received a positive advice from the Ethics Committee of the Faculty of Social Sciences (ECSS) after which the Dean of the Faculty formally approved this study (ECSW-2022-018).

The privacy of participants had been warranted using anonymous individual participant IDs. Data used for Chapter 2 and 4 were received from the treatment facilities without any identifiable personal information (e.g. name, BSN, telephone number or date of birth). For the study described in Chapter 5 no such identifiable personal information was collected. The consent forms collected for this study are stored in safes at the facilities and/or stored digitally in encrypted files at a secure network drive. All anonymous data are stored on a secure network drive.

Funding

This thesis was financially supported by the Behavioural Science Institute, Radboud University, Nijmegen. Peter de Looff is supported by a ZonMw Netherlands Fellowship grant (number 06360322210023). The studies described in Chapter 2 and 3 did not receive any additional funding. The studies of which data was reanalysed in Chapter 4 were funded by a grant from De Borg, DForZo (Directie Forensische Zorg) and ZonMw (number 729410003). The work described in Chapter 5 was also financially supported by De Borg. Furthermore, Noordelijk Platform Gehandicaptenzorg (NPGZ) provided financial support for this study.

Data management and availability

Radboud University and Behavioural Science Institute (BSI) have set strict conditions for the management of research data. Research Data Management was conducted according to the FAIR principles. All research data resulting from this dissertation were handled in accordance with the university's research data management policy (https://www.ru.nl/rdm/) and the BSI's research data management protocol (https://www.radboudnet.nl/bsi/rdm). The data that support the findings of Chapter 2 and 5 are available on reasonable request from the corresponding author. The data are not publicly available due to privacy of research participants. The data that support the findings of Chapter 4 are stored under restricted access and cannot be shared due to privacy of research participants. The data we extracted for the systematic review, Chapter 3, are available in Appendix 3 C and D.

Curriculum Vitae

Marlieke van Swieten was born in Leidschendam on September 12, 1994. She completed her secondary education (VWO) at Vlietland College in Leiden before pursuing Pedagogical Sciences at Radboud University, where she obtained her Bachelor's degree (Bene Meritum). At Radboud, Marlieke also participated in the Disciplinary Honours Programme, as part of which she conducted her Bachelor's thesis at Victoria University of Wellington, New Zealand. She went on to obtain her Master's degree (Cum Laude) in Education and Child Studies at Leiden University, specializing in Clinical Child and Adolescent Studies and Applied Neuroscience. During her Bachelor's and Master's degrees Marlieke also gained extensive practical experience working as a professional caregiver in a specialized daycare centre for children with severe developmental disabilities, a day activity centre supporting adults with mild to moderate intellectual disabilities, and a residential facility for adults with moderate to severe intellectual disabilities and epilepsy.

After obtaining her Master's degree, Marlieke gained intercultural experience as a behavioural specialist at Matoekoe in Suriname, a facility supporting children and adults with developmental disabilities. She subsequently worked as a cognitive therapist at Hersenz, treating individuals with acquired brain injuries. During this time, she also worked as a teacher and practicum coordinator at Erasmus University. In 2020, Marlieke began working at Trajectum, where she provided ambulatory psychological treatments for individuals with mild intellectual disabilities, and started her training as a mental health care psychologist (GZ-psycholoog). After one year, she also began her PhD at the Behavioural Science Institute of Radboud University in collaboration with Trajectum. In her PhD project she used her clinical experience in practice based research, within the innovative research field of physiology, measured with wearables, in relation to aggressive behaviour and self-harm in people with mild intellectual disabilities. This work culminated in this dissertation. Her research, described in chapter 5, was awarded the Stimuleringsprijs by the Noordelijk Platform Gehandicaptenzorg, an annual grant awarded for an innovative project that aims to improve the quality of life of people with an intellectual disability. During her PhD, alongside her research and clinical work she has been teaching for the faculty of Social Sciences of the Radboud University (Academic Skills and supervision of Master's theses).

In 2022 Marlieke completed her training as a mental health care psychologist (GZ-psycholoog) and worked at 's Heerenloo, at a residential living facility for youth with mild intellectual disabilities and children with severe developmental disabilities. Currently, she continues clinical work as a mental health care psychologist at BijBram, an innovative specialised mental health care facility in a rural setting.

Publication list

Journal contributions

Published

van Swieten, M., de Looff, P., VanDerNagel, J., & Didden, R. (2024). The association between aggressive behaviour and non-suicidal self-injury and shared risk factors in adults with mild intellectual disability. *Journal of Applied Research in Intellectual Disabilities, 37*(6), e13288. https://doi.org/10.1111/jar.13288

van Swieten, M., de Looff, P., VanDerNagel, J., Bouwmeester, S., & Didden, R. (2025). Listening to music is associated with reduced physiological and subjective stress in people with mild intellectual disabilities: A biofeedback study. *Research in Developmental Disabilities, 161,* 104976. https://doi.org/10.1016/j.ridd.2025.104976

van Swieten, M., Nijman, I., de Looff, P., VanDerNagel, J., & Didden, R. (2025). A systematic review of studies on the association between physiological parameters and self-harm. *Research in Developmental Disabilities, 162*, 105010. https://doi.org/10.1016/j.ridd.2025.105010

Submitted

Nijman, I.*, **Van Swieten, M.***, Bogaerts, S., Didden, R., Embregts, P., Hagoort, K., Hasselman, F., Koldijk, S., Konijn, C., Masthoff, E., VanDerNagel, J., Nijhof, K., Noordzij, N., Popma, A., Scheepers, F., De Schepper, F., Smeets, K., Strijbosch, E., & De Looff, P. (2025). Changes in heart rate and electrodermal activity after aggression in residential treatment facilities: A multicentre study

* Shared first author, these authors contributed equally to this work

Presentations

Nijman, I., **Van Swieten, M.** de Looff, P. (2024). Psychophysiology and wearable technology in Forensic Psychiatric Care: Potential, Pitfalls, and Progress. Presented at the *13th European Congress on Violence in Clinical Psychiatry*, Krakow, November 7, 2024.

trajectum

Radboud University

Behavioural
Science
Institute