Biocueing and ambulatory biofeedback to enhance emotion regulation: A review of studies investigating non-psychiatric and psychiatric populations

J.F. ter Harmsel¹,²,³, M.L. Noordzij⁴, A.E. Goudriaan⁵,⁶, J.J.M. Dekker⁵,⁶, L.T.A. Swinkels¹,², T.M. van der Pol¹,², A. Popma²

¹ Inforsa, Forensic Mental Health Care, Amsterdam, the Netherlands
² Amsterdam UMC, Department of Child and Adolescent Psychiatry and Psychology, Vrije Universiteit Amsterdam and Amsterdam Public Health Research Institute, Amsterdam, the Netherlands
³ Arkin, Department of Research and Quality of Care, Amsterdam, the Netherlands
⁴ University of Twente, Department of Psychology, Health and Technology, Enschede, the Netherlands
⁵ Amsterdam UMC, Department of Psychiatry, University of Amsterdam and Amsterdam Public Health Research Institute, Amsterdam, the Netherlands
⁶ Vrije Universiteit Amsterdam, Faculty of Behavioral and Movement Sciences, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Systematic review
Biocueing
Ambulatory biofeedback
Emotion regulation
Wearable technology

ABSTRACT

Over the last years, biofeedback applications are increasingly used to enhance interoceptive awareness and self-regulation, in psychiatry and beyond. These applications are used to strengthen emotion regulation skills by home training (ambulatory biofeedback) and real-time support in everyday life stressful situations (biocueing). Unfortunately, knowledge about the feasibility and effectivity of these applications is still scarce. Therefore, a systematic literature search was performed. In total, 30 studies (4 biocueing, 26 ambulatory biofeedback) were reviewed; 21 of these studies were conducted in non-psychiatric samples and 9 studies in psychiatric samples. Study characteristics, biofeedback characteristics, effectivity and feasibility outcomes were extracted. Despite the rapid advances in wearable technology, only a few biocueing studies were found. In the majority of the studies significant positive effects were found on self-reported (stress-related) psychological measures. Significant improvements on physiological measures were also reported, though these measures were used less frequently. Feasibility of the applications was often reported as sufficient, though not adequately assessed in most studies. Taken into account the small sample sizes and the limited quality of the majority of the studies in this recently emerging field, biocueing and ambulatory biofeedback interventions showed promising results. Future research is expected to be focusing on biocueing as a just-in-time adaptive intervention. To establish this research field, closer cooperation between research groups, use of more rigorous as well as individually tailored research designs and more valid feasibility and effectivity assessment are recommended.

1. Introduction

1.1. Emotion regulation and psychopathology

The human ability to regulate emotions, by monitoring, evaluating and modulating emotional responses, is essential in mental health (Gross and Jazaieri, 2014). Emotion regulation or self-regulation consists of the integrative functioning of psychological and physiological processes to adequately identify and respond to real and perceived threats (Taylor, 2006). Over the last years, the multifaceted process of emotion regulation gained attention as one of the core transdiagnostic concepts in treatment of psychiatric disorders (Hulsbergen et al., 2014). The large variety of psychopathology and maladaptive behaviors associated with emotion regulation difficulties, ranging from anxiety and mood disorders to substance abuse and aggressive behavior, stresses the clinical relevance of this concept (Gratz and Tull, 2010; Sheppes et al., 2015). Adequate assessment and treatment of problematic patterns of emotion regulation, is important in ameliorating psychiatric symptoms as well as general functioning (Gross and Jazaieri, 2014). Therefore, the development of technological applications aimed to enhance insight and to support self-regulation by real-time monitoring of physiological processes and delivering just-in-time-adaptive interventions (Riley et al., 2015) are of high relevance for therapy.

Emotional awareness is considered essential for adequate emotion regulation gained attention as one of the core transdiagnostic concepts in treatment of psychiatric disorders (Hulsbergen et al., 2014). The large variety of psychopathology and maladaptive behaviors associated with emotion regulation difficulties, ranging from anxiety and mood disorders to substance abuse and aggressive behavior, stresses the clinical relevance of this concept (Gratz and Tull, 2010; Sheppes et al., 2015). Adequate assessment and treatment of problematic patterns of emotion regulation, is important in ameliorating psychiatric symptoms as well as general functioning (Gross and Jazaieri, 2014). Therefore, the development of technological applications aimed to enhance insight and to support self-regulation by real-time monitoring of physiological processes and delivering just-in-time-adaptive interventions (Riley et al., 2015) are of high relevance for therapy.

Emotional awareness is considered essential for adequate emotion regulation, in psychiatry and beyond. These applications are used to strengthen emotion regulation skills by home training (ambulatory biofeedback) and real-time support in everyday life stressful situations (biocueing). Unfortunately, knowledge about the feasibility and effectivity of these applications is still scarce. Therefore, a systematic literature search was performed. In total, 30 studies (4 biocueing, 26 ambulatory biofeedback) were reviewed; 21 of these studies were conducted in non-psychiatric samples and 9 studies in psychiatric samples. Study characteristics, biofeedback characteristics, effectivity and feasibility outcomes were extracted. Despite the rapid advances in wearable technology, only a few biocueing studies were found. In the majority of the studies significant positive effects were found on self-reported (stress-related) psychological measures. Significant improvements on physiological measures were also reported, though these measures were used less frequently. Feasibility of the applications was often reported as sufficient, though not adequately assessed in most studies. Taken into account the small sample sizes and the limited quality of the majority of the studies in this recently emerging field, biocueing and ambulatory biofeedback interventions showed promising results. Future research is expected to be focusing on biocueing as a just-in-time adaptive intervention. To establish this research field, closer cooperation between research groups, use of more rigorous as well as individually tailored research designs and more valid feasibility and effectivity assessment are recommended.

* Corresponding author at: For the corresponding author: Annemieke ter Harmsel, Vlaardingenlaan 5, 1059 GL Amsterdam, the Netherlands.
E-mail address: annemieke.ter.harmsel@inforsa.nl (J.F. ter Harmsel).

https://doi.org/10.1016/j.ijpsycho.2020.11.009
Received 22 July 2020; Received in revised form 17 November 2020; Accepted 19 November 2020
Available online 25 November 2020
0167-8760/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
regulation (Füstös et al., 2013; Gross and Jazari, 2014). This ability is often divided into two main components: the ability to attend to the own emotional experiences (attention to emotion) and the extent to which the own emotional experiences are understood (emotional clarity) (Boden and Thompson, 2015). Emotional awareness, or lack thereof, has been found to be important in understanding psychopathology, both in internalizing problems (Berenbaum et al., 2012; Sendzik et al., 2017) and externalizing problems (Roberton et al., 2015; Velotti et al., 2017). For example, considering sadness, low emotional awareness has been identified as a significant predictor of anxiety and depression symptoms, especially in younger populations. Considering anger, a lack of emotional awareness can result in behavioral problems. For example, forensic psychiatric patients often have difficulty observe increased inner tension and timely detect rising arousal levels using physiological signals such as accelerating heartbeat or sweating, which have been shown to precede aggressive behavior (De Looff et al., 2019).

Emotional granularity, a more nascent concept in emotion theory, seems relevant for adequate emotion regulation as well. This term refers to the ability to differentiate between emotions based on a combination of the dimensions of valence (negative versus positive affectivity) and arousal (calming versus exciting) (Tugade et al., 2004). Lowered levels of emotional granularity are associated with psychopathology, such as schizophrenia, borderline personality disorder, depression and substance abuse (Smidt and Suvak, 2015).

Furthermore, adequate emotion regulation depends on the body’s ability to achieve and maintain homeostatic regulation (Crockett et al., 2017). One of the human regulatory systems is the autonomic nervous system (ANS), which regulates respiration, heartbeat and sweat secretion. The ANS is divided into two branches: the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). Typically, in case of emotional stress, SNS and PNS cooperate reciprocally: when the SNS is activated, creating a fight- or flight-response, the PNS is deactivated (Levenson, 2014). In case of coactivation of SNS and PNS, research results indicate that coping resources are overwhelmed by exaggerated arousal, leading to internal distress and anxious responses (Hastings et al., 2009) as well as reactive aggression (El-Sheikh et al., 2009; Raine, 2008). In case of co-inhibition of both systems, studies report passive attention to threat or dangerous activities to induce arousal, contributing to covert forms of externalizing behavior, such as proactive aggression (Boyce et al., 2001; El-Sheikh et al., 2009). Altogether, dysregulated responses of the ANS have been shown to hinder homeostatic and adaptive functioning, to enhance stress vulnerability and to give rise to both physical and psychological dysfunctions (Mandel, 2003). Therefore, training methods aimed at enhancing emotional awareness, such as mindfulness based interventions for patients with internalizing problems (Kranzer et al., 2016) and treatment interventions focusing on recognition of arousal for patients with externalizing, aggressive behavior (Novaco, 2007) are often mentioned as potentially beneficial additions to current treatment protocols. Nevertheless, it should be noted that no consistent or specific patterns of ANS changes were found in association with specific emotion categories (Siegel et al., 2018). In summary, ANS information may support emotional awareness (particularly the facet of attention to emotion) and may aid the detection of emotional intensity based on the arousal dimension, though is not expected to enhance emotional granularity by strengthening the determination of valence.

1.2. Biofeedback to enhance emotion regulation

The addition of real-time biofeedback in-and outside the treatment setting, using new technological applications, is an innovative attempt to enhance emotion regulation skills. In the process of biofeedback, instruments monitor physiological parameters (e.g. heart rate, skin conductance, respiration), transform these measurements into auditory and/or visual signals and present these signals directly to the user (McKee, 2008). Providing psychiatric patients with this physiological information is considered to be a helpful means to signal arousal changes in response to emotional events, to recognize emotions adequately and to influence emotional processes in a health-improving way (Lisetti and Nasoz, 2004; Yucha and Montgomery, 2008; Cornet et al., 2017).

In the past decades, several biofeedback approaches have been developed. During a standard biofeedback paradigm consisting of multiple sessions, patients are trained to influence a physiological parameter, for example Heart Rate Variability (HRV) (Lehrer, 2013) or a derivative thereof called coherence (McCraty and McCraty and Zayas, 2014), by consciously alternating their (breathing) responses to the feedback provided. The ability to regulate their physiological reactions in response to emotional stimuli is trained during the sessions and is supposed to be transferred to emotionally challenging situations outside the sessions, when biofeedback is not available (Peira et al., 2014). To differentiate this biofeedback training paradigm from more modern approaches, this type is called traditional biofeedback. To support patients to train at home, new devices were developed and studied (Whited et al., 2014). These portable devices can be used at home, in combination with a personal computer, when the users wants, preferably several times a day. In this study, we will refer to this biofeedback type as ambulatory biofeedback. Over the last years, rapid developments in the field of non-invasive devices (such as breath bands, wrist sensors, smart fibers and interactive textiles) opened up opportunities for real-time measurement in everyday life. These developments have enabled a new type of personalized biofeedback, delivered in everyday life, which will be referred to as biocueing in this review. Biocueing can be defined as cueing users when physiological values are in a specific ‘at risk’-range, and allowing just-in-time behavioral support (Paradiso et al., 2011; Riley et al., 2015). As such, whereas ambulatory biofeedback can mainly be seen as a method to encourage at home training, biocueing has an additional function to signal deviating arousal levels, real-time, in everyday life.

Based on multiple studies, mainly performed among patients with internalizing psychological problems, traditional biofeedback is considered efficacious in treatment of anxiety and chronic pain, and possibly efficacious in treatment of depression (Yucha and Montgomery, 2008). Considering cardiovascular feedback, the main focus of our study, two recent meta-analyses revealed significant reductions in (self-reported) stress and anxiety after HRV-biofeedback training (Goossel et al., 2017) and an overall small to medium effect size for HRV-biofeedback, with largest effect sizes for anxiety, depression, anger and performance measures (Lehrer et al., 2020). It is hypothesized that HRV-biofeedback stimulates activity in the prefrontal regions that are sensitive to interoceptive information, thereby strengthening the neuronal pathways underlying emotion regulation (Mather and Thayer, 2018). In ambulatory biofeedback studies, this potential to enhance self-regulation skills is also found, both among psychiatric and non-psychiatric populations with emotional problems (Eddie et al., 2015; Tolin et al., 2017; Yu et al., 2018). Biocueing is relatively new, but may be particularly useful for patients who lack insight in the physiological signals that precede dysregulated behavior in everyday life such as binge eating episodes (Godfrey, 2018), self-injurious behavior (Koenig et al., 2017) and psychotic experiences (Schlier et al., 2019). Focusing on aggression, several studies showed that physiological information can be used as a predictor of aggressive behavior, for example among youth with autism spectrum disorder (Cumpaassio et al., 2018; Ferguson et al., 2019; Nuke et al., 2019), patients with intellectual disabilities (Palix et al., 2017), patients with schizophrenia (Wang et al., 2019) and forensic patients (Kuijpers et al., 2012; De Looff et al., 2019). Considering the emotion of anger – in forensic populations often manifested as externalizing, aggressive behavior – biocueing could strengthen awareness of high risk situations and support patients to use learned behavioral skills to prevent aggressive incidents (Cornet et al., 2016).
2.2. Search strategy

3) background articles and reviews; 4) articles only accessible in a language different from English; 5) use of a non-human sample; 6) use of a sample with physical conditions; 7) background articles and reviews; 8) articles only accessible in a language other than English or Dutch; 9) conference, poster or presentation abstracts and 10) use of another intervention instead of biocueing or ambulatory biofeedback (e.g. pharmacotherapy, yoga, mindfulness training). Neurofeedback studies were excluded as well, since neurofeedback devices do not allow ambulatory measurement in everyday life.

2.3. Selection criteria

We decided to include published experimental and pilot-studies in which new biofeedback interventions, using ambulatory technology, were used for a wide range of emotion regulation difficulties in psychiatric populations (according to DSM-IV-TR or DSM-5 criteria) and non-psychiatric populations. Exclusion criteria for the studies were: 1) use of a non-human sample; 2) use of a sample with physical conditions; 3) background articles and reviews; 4) articles only accessible in a language other than English or Dutch; 5) conference, poster or presentation abstracts and 6) use of another intervention instead of biocueing or ambulatory biofeedback (e.g. pharmacotherapy, yoga, mindfulness training). Neurofeedback studies were excluded as well, since neurofeedback devices do not allow ambulatory measurement in everyday life.

2.4. Selection procedure

First, duplicate publications were excluded. Titles and abstracts of studies resulting from the search were independently screened for eligibility by authors AtH and MN. Discrepancies in judgement for eligibility were discussed until agreement between reviewers was achieved. The reviewers used the web application Rayyan (Ouzzani et al., 2016) for the selection process. Subsequently all selected papers were read in full to check for all in- and exclusion criteria. Only two articles met the most strict criteria for inclusion. Therefore, the remaining papers were read to include studies using biocueing among non-psychiatric populations and studies investigating ambulatory biofeedback in psychiatric and non-psychiatric populations as well. Among the ambulatory biofeedback studies, we found eight studies in which the same samples were used. In these cases, the most recent (3) or complete (1) studies were included.

2.5. Data collection

AtH extracted the following data from the included studies: the study characteristics (sample size, research design), participant characteristics (age, gender and psychopathology), used device or application, modality and duration of biofeedback, outcome measures, feasibility information and effectiveness outcomes. Regarding the feasibility and effectiveness outcomes, we extracted both qualitative and quantitative information (psychological and physiological measures, statistical significance). Data extraction was checked by the second author (MN). Disagreements were resolved by consensus.

2.6. Risk of bias in individual studies

Three master students, EN, NvE and LS, independently assessed and rated methodological rigor, selection and reporting bias of the included studies using the PEDro checklist (Maher et al., 2003). Based on this checklist, the quality of a paper was considered ‘high’ in case of 8–11 points, ‘medium’ in case of 4–7 points and ‘low’ in case of 1–3 points. In case of discrepancy between the raters, the final score was based on the mean.

3. Results

3.1. Identified studies

The literature search yielded a total of 4149 unique study records. After title and abstract screening, 119 studies remained. The most common reason for exclusion after preliminary screening was the population: a non-human sample or a human population with physical problems (e.g. heart failure, asthma). The full text of the remaining 119 papers were examined in detail. Based on the in- and exclusion criteria described above, only two biocueing studies using a psychiatric sample were found. Two more biocueing studies were added, performed among non-psychiatric populations. Expansion of our search strategy made us include 26 ambulatory biofeedback studies; seven studies performed among psychiatric populations and 19 studies among non-psychiatric populations. After final selection, information from 30 studies was extracted and summarized. An overview of the selection process is visualized in Fig. 1.

3.2. Study quality

Of the 30 studies we selected for inclusion in this review, two studies were rated as high quality, 19 were rated as medium quality and 9 were rated as low quality. Based on recent guidelines (Koo and Li, 2016), intrarater reliability between the three raters was excellent, ICC(3,k) = 0.91–0.97.

3.3. Study characteristics

In total, we included four biocueing and 26 ambulatory biofeedback
Of all studies, 21 used a non-psychiatric sample and 9 studies were performed among psychiatric populations. Targeted psychological problems varied widely: aggressive behavior, post-traumatic stress, depression, anxiety and pregnancy-, work- and performance-related stress. Participants were predominantly adults (24 studies); young adults (5) and children (1) formed a minority. Different study designs were used: randomized controlled trials (6 studies), quasi-experimental studies (12), case studies (5) and pilot studies (8). A summary of the study characteristics is presented in Table 1.

### 3.4. Biofeedback characteristics

Different devices and applications were used in the selected studies. Considering the ambulatory biofeedback studies, most studies used the emWave device (12 studies), followed by the StressEraser (6). Other applications were the Positive Technology app (2), the RELAX app (1), the BART app (1), the Freespira (1), the EliteHRV app (1), the Breath Pacer (1), the STRS app (1), the Ripple shirt (1), the MioFuse heart rate monitor (1), a not specified heart rate monitor (1) and a multi-biosensor platform (1). In most studies, Heart Rate Variability (HRV) was chosen as biofeedback modality (23). Other studies used Heart Rate (HR; 3), respiration measures (2), Skin Conductance Level (SCL; 1) or a combination of measures (1). The duration of training and actual use of the biofeedback devices varied widely. In the majority of the bio cuing studies, a brief instruction was given or a user manual was provided. Actual frequency of use depended on the user; except for one, more structured study. In the ambulatory biofeedback studies, training varied from a single and brief instruction up to 10 training sessions. Participants in these studies were encouraged to train once to three times a day, with a total duration ranging from one week to eight weeks. A summary of the biofeedback characteristics can be found in Table 1.

### 3.5. Effectivity outcomes

In most studies, effectivity of the intervention was evaluated using psychological or physiological measures (37%) or a combination of both (57%). In two studies (7%), no quantitative measures were administered. An overview of significant ($p < .05$) pre-post intervention changes (and group differences, if applicable) on at least one of the psychological or physiological outcome measures is presented in Fig. 2. The results of the bio cuing studies – our primary focus – will be reviewed first and in more detail, followed by the ambulatory biofeedback studies; both categories subdivided into non-psychiatric and psychiatric populations.

#### 3.5.1. Bio cuing

##### 3.5.1.1. Non-psychiatric populations

In general population samples, two studies were found in which devices enabled prompting of regulation strategies by physiological information. In one study, a systemic approach to bio cuing was piloted among three children with violent or disruptive behavior (though without a formal diagnosis) and their parents in a multi-family group setting (McHugh et al., 2010). All children...
### Table 1

Description of main characteristics of the selected studies.

| Study characteristics | Biofeedback characteristics | Outcomes |
|-----------------------|----------------------------|-----------|
| **Biofeedback among non-psychiatric populations** |
| Howell et al. (2016)  | Pilot; app evaluation study | Ripple SCL No traditional protocol used | No measures used; no measures used Qualitative feedback only; mainly critical |
| N = 17               | Duration (training; practice; intervention length): L: 8–20 h | T: brief instruction, one meeting P: daily | |
| Age: 19–41           |                          | 24.1–32.2                                           |
| F/M: 53/47           |                          | | |
| McHugh et al. (2010) | Case studies; evaluation study | Heart rate monitor HR No traditional protocol used | No measures used; only qualitative information No measures used Qualitative feedback only; positive |
| N = 3                | Duration (training; practice; intervention length): L: 8–20 h | T: brief instruction P: user dependent L: user dependent | |
| Age: 10–15           |                          | 19–41                                               |
| F/M: 33/67           |                          | | |
| Biocueing among psychiatric populations |
| Mackintosh et al. (2017) | RCT; TAU + BF vs. TAU | RELAX app HR No traditional protocol used | No measures used No measures used TFO, qualitative feedback; positive |
| N = 58               | Duration (training; practice; intervention length): L: user dependent | T: introduction and user manual P: user dependent L: 6 weeks | |
| Age: 24–71           |                          | 18–41                                               |
| F/M: 0/100           |                          | | |
| Savard (2017)        | Case studies; mixed methods, ABABAB-design | Mio Fuz heart rate monitor HR No traditional protocol used | No measures used No measures used Qualitative feedback only; positive yet critical |
| N = 5                | Duration (training; practice; intervention length): L: user dependent | T: daily during intervention phase P: 5 days a week L: 6 weeks | |
| Age: 15–18           |                          | 17–19                                               |
| F/M: 20/80           |                          | | |
| Ambulatory biofeedback among non-psychiatric populations |
| De Bruin et al. (2016) | RCT; MM vs. BF vs. PE | StressEraser HRV HRV-BF protocol not further specified | ACS, BRIEF-A, FFMQ, SCS, PSWQ; *pp, MM = BF – PE No measures used No measures used Impl. rate and author report; positive |
| N = 75               | Duration (training; practice; intervention length): L: 4 weeks | T: instruction and assisted practice P: daily L: 4 weeks | |
| Age: 18–40           |                          | 17.3–40.8                                           |
| F/M: 73/27           |                          | | |
| Gaggioli et al. (2014) | Pilot; app evaluation study | Positive Technology app HRV No traditional protocol used | SAM/PSS; *pp HR before and after exercise; n.s. Author report only; positive yet critical |
| N = 32 (7)           | Duration (training; practice; intervention length): L: 6 weeks | T: guided training P: user dependent, 5–15 min/session L: user dependent | |
| Age: n.r.            |                          | 19.1–27.9                                           |
| F/M: n.r.            |                          | | |
| Hasuo et al. (2019)  | Open-label RCT; BF vs. control group | Breath Pacer HRV No traditional protocol used, based on HRV-BF protocol | SF-12; *pp, BF = control HRV-indices; *pp, BF > control Impl. rate and author report; positive |
| N = 54               | Duration (training; practice; intervention length): L: 4 weeks | T: 1 session and self-training P: 5–20 min a day L: 4 weeks | |
| Age: n.r.; m = 63    |                          | 17.8–28.2                                           |
| F/M: 68/32           |                          | | |
| Kizakevich et al. (2019) | Pilot; pre-test vs. post-test | emWave HRV Coherence protocol | STAI-Y, PES; n.s. HRV-indices; *pp Qualitative feedback only; positive Adherence rate and author report; mainly critical |
| N = 7                | Duration (training; practice; intervention length): L: 4 weeks | T: 5–6 sessions P: 20 min a day L: ± 5 weeks | |
| Age: 18–45           |                          | 18.1–42.5                                           |
| F/M: 100/0           |                          | | |
| Keeney (2009)        | Pilot; app evaluation study | BART app HRV No traditional protocol used, based on HRV-BF principles | CD-RISC, BCS, PSS, PG, SDS(2); outcomes not reported HRV-indices; *pp in subset Adherence rate and author report; mainly critical |
| Military personnel, veterans and civilians |
| Kizakevich et al. (2019) | Pilot; pre-test vs. post-test | BART app HRV No traditional protocol used, based on HRV-BF principles | CD-RISC, BCS, PSS, PG, SDS(2); outcomes not reported HRV-indices; *pp in subset Adherence rate and author report; mainly critical |
| N = 328 (49)        | Duration (training; practice; intervention length): L: ± 4 weeks | T: 1 session and self-training P: 3 times a week L: (up to 12 months) | |
| Age: 20–69           |                          | 27.1±7.8                                           |
| F/M: 47/53           |                          | | |
| Kudo et al. (2014)   | Quasi-experiment; TAU + BF vs. TAU | StressEraser HRV HRV-BF protocol not further specified | EPDS; *pp, TAU+ > TAU HRV-indices; *pp, BF > control Author report only; positive |
| Mothers in early post-partum period with stress |
| N = 55               | Duration (training; practice; intervention length): L: ± 4 weeks | T: brief instruction P: daily L: ±4 weeks | |
| Age: n.r.; m = 33.4 vs. m = 30.5 |                          | 26.9–19.7                                           |
| F/M: 100/0           |                          | | |
| Lemaire et al. (2011) | RCT; TAU + BF vs. TAU | emWave HRV Coherence protocol | PSS, PQQA-R; *pp, TAU+ > TAU HR, BP, salivary and cortisol levels, n.s., BF = control Adherence scores and author report; positive |
| Physicians with work-related stress |
| N = 40              | Duration (training; practice; intervention length): L: 4 weeks | T: 1 session and support visits P: 5 min, 3 times a day L: 4 weeks | |
| Age: n.r.; m = 44.8 vs. m = 47.8 |                          | 23.1–39.1                                           |
| F/M: 42/58           |                          | | |

(continued on next page)
Table 1 (continued)

| Study characteristics | Biofeedback characteristics | Outcomes |
|-----------------------|-----------------------------|----------|
| **Application**       | **Modality**                | **Feasibility**               | **Psychological measure(s)** | **Physiological measure(s)** | **Feasibility measure(s)** |
| **Implemented protocol** | **Duration** (training; practice; intervention length) | **Feasibility** | **Psychological measure(s)** | **Physiological measure(s)** | **Feasibility measure(s)** |
| **Author(s)**         | **Sample**                  | **Design**                        | **Application** | **Modality** | **Duration** | **Outcomes** | **Feasibility** | **Psychological measure(s)** | **Physiological measure(s)** | **Feasibility measure(s)** |
| May et al. (2019)     | Students                    | N = 90                            | RCT; BF vs. HIIT | emWave        | Coherence protocol | T: 3 sessions a week n: 20 min sessions, and frequent self-practice | L: ±4 weeks | P: several times a week; ± one day | L: ±3 months | SBI, CES-D, STAI, serial subtraction task; *pp, BP, BF > HIIT vs control | HRV-indices, BP; *pp, BF > HIIT vs control | No measures used |
| McCraty et al. (2009) | Correctional officers with work-related stress | N = 75                            | Quasi-experiment; BF vs. wait-list control | emWave        | Coherence protocol | T: 5 sessions | L: ±4 weeks | P: several times a week; ± one day | L: ±3 months | PWP, JAS, BSI, POQA-R; *pp, BF > control | HRV-indices, BP, cortisol; *pp, BF > control | No measures used |
| McCraty and Atkinson (2012) | Police officers with work-related stress | N = 65                            | Quasi-experiment; BF vs. wait-list control | emWave        | Coherence protocol | T: 3 sessions | L: ±4 weeks | P: several times a week; ± one day | L: ±3 months | POQA-R, qualitative information; *pp; BF > control | No measures used to evaluate | No measures used |
| Narita et al. (2018)  | Pregnant woman with anxiety about labor | N = 97 (38)                        | Quasi-experiment; BF vs. no intervention | StressEraser HRV | HRV-BF protocol | T: brief instruction | L: ±4 weeks | P: daily | L: ±7 weeks | W-DEQ, PSQI, fatigue-VAS, qualitative information; *pp, BF > no intervention | HRV-indices; n.s. | Qualitative feedback only; positive |
| Poston (2018)         | Pregnant women with depressive symptoms | N = 41                            | Quasi-experiment; self-selected BF vs. no intervention | emWave        | Coherence protocol | T: video training and self-training | L: ±4 weeks | P: daily | L: ±4 weeks | EPDS, PDSS(2), qualitative information; *pp, BF > no intervention | No outcomes reported | Author report only; positive yet critical |
| Ramey et al. (2016)   | Police officers with work-related stress | N = 38                            | Quasi-experiment; baseline vs. post-tests | emWave        | Coherence protocol | T: 4 weeks | L: ±4 weeks | P: educational class; ± before and after stressful events | L: ±3 weeks | HRV-indices, BP, cortisol; *pp | Qualitative feedback only; positive |
| Serino et al. (2014)  | Society members interested in stress therapy | N = 68                            | Pilot; app evaluation study | Positive Technology app | No traditional protocol used | T: brief instruction | L: ±1 week | P: user dependent | L: ±15 minutes/session | SAM/PSS; *pp | HR before and after exercise; n.s. | Author report only; positive yet critical |
| Schumann et al. (2019) | Society members interested in biofeedback | N = 24                            | Quasi-experiment; BF vs. control | EliteHRV app HRV | No traditional protocol used, based on HRV-BF protocol | T: self-training after one session | L: ±1 week | P: 5 times a week | L: ±8 weeks | SST; n.s. | HRV, SCI, BP, BR, BRS, DIA; *pp, BF > control | No measures used |
| Tanis (2012)          | Female athletes with sports related mental stress | N = 13                            | Quasi-experiment; pre-test vs. post-test | emWave        | Coherence protocol | T: performance score | L: ±8 weeks | P: 6 sessions | L: ±6 weeks | Performance score, qualitative information; n.s. | HRV-indices; *pp, BF > control | No measures used |
| Thurber (2007)        | Student musicians with performance anxiety | N = 14                            | RCT; BF vs. no intervention | emWave        | Coherence protocol | T: 4 – 5 sessions, self-practice after 3 sessions | L: ±6 weeks | P: daily | L: ±6 weeks | STAI-Y, PAI, FSS, qualitative information; *pp, BF > no intervention | No measures used | Qualitative feedback only; positive yet critical |
| Weltman et al. (2014) | Police office personnel with work-related stress | N = 14                            | Case-study; app evaluation study | SRTS app HRV | No traditional protocol used, based on coherence principles | T: self-training after one session | L: ±1 week | P: 5 times a week | L: ±8 weeks | POQA-R, qualitative mentor and participant reports; *pp | No measures used | Qualitative feedback only; positive |
| Wu et al. (2012)      | Society members, half of them unemployed | N = 67 (33)                        | Quasi-experiment; pre-test vs. post-test evaluation | Multi-Biosensor Platform HRV | No traditional protocol used | T: lab-experiment, 3 training sessions | L: ±6 weeks | P: user dependent | L: ±1 week | HRV-indices, respiratory pattern; *pp | No measures used | No measures used |

(continued on next page)
| Author(s) | Sample | Participants (number; age range; gender) | Design | Biofeedback characteristics | Implemented protocol | Duration (training; practice; intervention length) | Outcomes | Psychological measure(s) | Physiological measure(s) | Feasibility measure(s) |
|----------|--------|----------------------------------------|--------|---------------------------|----------------------|-----------------------------------------------|----------|--------------------------------|-----------------------------|----------------------------|
| Eddie et al. (2015) | Young adults with substance use disorder | N = 41; Age: 20–25; F/M: 0[100] | Quasi-experiment; TAU + BF vs. TAU | emWave HRV | Abbreviated HRV-BF protocol | T: 3 sessions; P: daily; L: 3 weeks | PACS, PSS; *pp, TAU+ − TAU; TAU+ − TAU | Qualitative feedback only; positive | Author report only; positive |
| Kim et al. (2012) | Patients with severe brain injury | N = 13; Age: 23–63; F/M: 54/46 | Quasi-experiment; pre-test 1 vs. pre-test 2 vs. post-test | emWave HRV | Coherence protocol | T: 10 sessions, self-practice after 4 sessions; P: daily; L: ± 6 weeks; T: evaluation after 1 week of self-practice | Impairment index, HCT, FDA-CPT, BRIEF-A; n.s. | HRV-indices; *pp | Qualitative feedback only; positive | Author report only; positive |
| McFayden and Spira (2009) | Military officer with adjustment disorder | N = 1; Age: 39; F/M: 0[100] | Case-study | StressEraser HRV | No traditional protocol used, based on HRV-BF principles | T: 10-20 min a day, on weekdays; P: user dependent; L: 6 weeks. | No measures used | Author report only; positive | N/A |
| O’Neill and Findlay (2014) | Patients with traumatic brain injury and aggressive behavior | N = 2; Age: 18, 33; F/M: 0[100] | Case-studies, within subjects AB-design | emWave HRV | Coherence protocol | T: 1 min instruction; P: 20 min a day; L: 3–4 weeks | OASMNR; *pp (J[2]) | No measures used | Author report only; positive | N/A |
| Reiner (2008) | Patients with anxiety disorders and depression | N = 20; Age: 18–65; F/M: 50/50 | Pilot; RSA-BF device evaluation study | StressEraser HRV | HRV-BF protocol not further specified | T: 15 min instruction; P: 20 min a day; L: 24 days (1); 22 weeks (2) | STAI-Y, PSQI, STAXI, and qualitative information; *pp | No measures used | Qualitative feedback only; positive yet critical | N/A |
| Tolin et al. (2017) | Patients with panic disorder | N = 48; Age: 18–65; F/M: 59/41 | Quasi-experiment; pre-test vs. post-test vs. follow-up | Freespira Respiration | No traditional protocol used | T: briefing instruction; P: 17 min, twice a day; L: 4 weeks | DDS(1), SDS (1) and CGI-S; *pp | Adherence and patient satisfaction scores; positive | N/A |
| Zucker et al. (2009) | Patients with substance use disorder and PTSD | N = 38; Age: 18–60; F/M: 45/55 | Pilot; RSA-BF vs. PMR as adjunctive interventions | StressEraser HRV | HRV-BF protocol not further specified | T: briefing instruction; P: 20 min a day; L: 4 weeks | DAPS, PCL-C, BDI-II, ISI; *pp, BF > PMR | HRV-indices; *pp, BF > PMR | Qualitative feedback only; positive | N/A |

Notes: *pp, significant pre-post intervention change (p < .05) on at least one of the outcome measures; X > Y, significant difference (p < .05) between intervention and control group, favoring the intervention, on at least one of the outcome measures; X = Y, no significant difference (p < .05) between intervention and control group.

Abbreviations of terms and questionnaires: ACS, Attention Control Scale; BART, Biofeedback-Assisted Resilience Training; BCS, Brief Coping Scale; BF, Biofeedback; BDI-II; Beck Depression Inventory; B-IPF, Brief Inventory Psychosocial Functioning; BR, Breathing Rate; BRIEF-A, Behavior Rating Inventory of Executive Function-Adult Version; BRS, Baroreflex Sensitivity; BSI, Brief Symptom Inventory; CBCL, Child Behavior Checklist; CD-RISC, Connor-Davidson Resilience Scale; CES-D, Center for Epidemiologic Studies Depression Scale; CGI-S, Clinician Global Impression-Severity Scale; DAPS, Detailed Assessment of Posttraumatic States; DAR-5, Dimensions of Anger Reactions; DIA, Pupil Diameter; DOF, Direct Observation Form; EPDS, Edinburgh Postnatal Depression Scale; ESS, Epworth Sleepiness Scale; F/M, female-male ratio; FFMQ, Five Facet Mindfulness Questionnaire; FSS, Flow State Scale; GSS, Groningen Sleep Quality Scale; HCT, Halstead Category Test; HHT, High Intensity Interval Training; HR, Heart Rate; HRV, Heart Rate Variability; IES, Impact Events Scale; BI, Insomnia Severity Index; IVA-CPT, Integrated Visual and Auditory Continuous Performance Test; JAS, Jenkins Activity Survey; m, mean; MQ, Maastricht Questionnaire; N, number; n.r., not reported; n.s., not significant; OASMNR, Overt Aggression Scale Modified for Neurorehabilitation; PACS, Parental Account of Children’s Symptoms; PAI, Performance Anxiety Inventory; PCL-5, PTSD Checklist for DSM-5; PCL-C, PTSD Checklist-Civilian Version; PDSS(1), Panic Disorder Severity Scale; PDSS(2), Postpartum Depression Screening Scale; PE, Physical Exercise; PES, Pregnancy Experience Scale; PET, End-tidal CO2; PGI, Posttraumatic Growth Inventory; PHQ-9, Patient Health Questionnaire; PMR, Progressive Muscle Relaxation; POQA-R: Personal and Organizational Quality Assessment—Revised; PSSQI, Pittsburgh Sleep Quality Index; PSS, Perceived Stress Scale; PSWQ, Penn State Worry Questionnaire; PWP, Personal Wellness Profile; RELAX, Remote Exercises for Learning Anger and Excitation Management; RCT, Randomized Clinical Trial; RR, Respiration Rate; RSES, Response to Stressful Experiences Scale; SBI, School Burnout Inventory; SCS, Self-Compassion Scale; SCL, Skin Conductance Level; SDS(1), Sheehan Disability Scale; SDS(2), Sleep Disturbance Scale; SF-12, Short Form Health Survey; STA(−Y), State Trait Anxiety Inventory (Y-version); STAXI, State Trait Anger Expression Inventory; SUS, System Usability Scale; TAU, Treatment As Usual; TFQ, Teaching Feedback Questionnaire; TRF, Teacher’s Report Form; VAS, Visual Analog Scale; W-DEQ, Wijma Delivery Expectancy/Experience Questionnaire; YSR, Youth Self-Report.
used heart rate breast band monitors, wirelessly connected with watch-like wrist bands. The monitor emitted an auditory signal when a specific threshold (compared to a personal baseline) was reached. In the first case, the use of the biofeedback system helped to show the child’s difficulty to verbalize emotions. In the second case, the system was used for breathing control and focused relaxation to facilitate a faster return to resting heart rate levels after a stressful event. In the last case, the alarm function of the monitor supported the child to take time to ‘cool down’. Teachers and parents reported increased self-control and anger management among pupils. Overall, the authors deemed the intervention helpful to reflect on triggers and to initiate (self-)calming strategies. They state that feedback from children and parents was consistently positive. In another study, the Ripple, a biosensing shirt measuring and displaying skin conductance levels, was piloted (Howell et al., 2018). Participants were observed, interviewed and encouraged to critically review the design. The Ripple system increased emotional awareness and reflection, but also yielded new questions. Whereas some participants related the displayed data to their own feelings in a natural way, others seemed to adjust their emotional experiences to the biosensing information. Furthermore, the system seemed to value some emotions over others; the absence or presence of display changes fostered insecurities in some participants. Authors emphasize that ethical issues should be considered in the future given the perceived authority of biosensing interventions.

3.5.1.2. **Psychiatric populations.** Two more bioecuing studies were found among psychiatric populations. In one high-quality RCT, conducted among veterans with post-traumatic stress disorder, a mobile application system was added to regular anger management therapy (Mackintosh et al., 2017). This system consisted of an app, a wearable heart rate monitor, a remote server and a web-based therapist interface. The app could be used to increase awareness, to prompt and practice skills in daily life, to monitor symptoms and to record physiological data. Post-treatment, participants in both conditions demonstrated significant reductions in anger severity and reductions in PTSD-symptomatology; the experimental condition did not outperform treatment as usual. Participants who used the RELAX-app reported increased ease of use, increased helpfulness and decreased frustration over time. In another study, a multi-method AB-design was used to evaluate the therapeutic use of a personal heart rate monitor among adolescents with anger-related diagnosis and disruptive behavior (Savard, 2017). Adolescents were instructed repetitively in the use of the monitoring device and deployment of their personalized relaxation strategy when their heart rate elevated. In three out of five cases, a significant decline in anger and aggressive behavior was shown; the other cases showed a non-significant decline or mixed results. Linear regression showed that aggressive behaviors declined in self-report as well as in all parents’ and teachers’ reports, except for one teacher. Overall, participants were satisfied with the simplicity of the method. Some reported mechanical or tactile issues with the device or difficulty following the directions. The majority of the adolescents, parents and teachers reported that the intervention contributed to (small) behavioral changes.

3.5.2. **Ambulatory biofeedback**

3.5.2.1. **Non-psychiatric populations.** Our search yielded 21 ambulatory biofeedback studies conducted in general population samples that are known to be prone to stress, due to work- or performance related pressure or pregnancy. Four studies were conducted among police personnel and correctional officers. In a quasi-experimental study among correctional officers significant improvements in HRV and blood pressure and significant reductions in psychological stress measures were found in the biofeedback condition compared to the wait-list control condition (McCraty et al., 2009). A case-study among police staff, evaluating the STRSS biofeedback application, showed significant improvement on personal and organizational quality, including stress, post-intervention (Weltman et al., 2014). In a study among police officers significantly greater reductions in distress, negative emotions and depression were found for the experimental condition compared to wait-list control (McCraty and Atkinson, 2012). In another study among police officers, the addition of ambulatory biofeedback was associated with significant improvement in HRV and positive, although non-significant, changes on other psychological and physiological outcomes (Ramey et al., 2016).

Ambulatory biofeedback was also tested among other groups in which high (work- and performance related) stress levels are often reported. In a subset of a study among (former) military personnel and civilians HRV-values significantly decreased in a stressor task and significantly increased during biofeedback-assisted slow paced breathing (Kizakevich et al., 2019). In a RCT among physicians the perceived stress score in the biofeedback condition decreased significantly compared to support visits only (Lemaire et al., 2011). In an usability study evaluating a portable resonant breathing device among family caregivers of cancer patients, quality of life scores significantly improved post-intervention, though not better compared to the control condition (Hasuo et al., 2019). Low-frequency HRV-values improved significantly better in the biofeedback condition. Furthermore, a quasi-experimental study among athletes with sports-related mental stress showed no significant changes on individual (athletic) performance.

![Fig. 2. Effectivity outcomes on psychological and physiological measures.](image-url)
(Tanis, 2012). Qualitatively, reductions in mental and physical stress and enhancement of mental and physical state were reported. In a study among student musicians performance anxiety and HRV improved with large effect sizes in the biofeedback condition compared to the control group (Thurber, 2007). Furthermore, a RCT among stressed young adults revealed that three stress regulation methods (mindfulness meditations, ambulatory biofeedback and physical exercises) were all effective in improving stress-related outcomes and did not differ significantly from each other (De Bruin et al., 2016). In a similar study, two interventions to reduce school burnout (ambulatory biofeedback and physical workout) were compared to a wait-list control condition (May et al., 2019). Significant reductions in burnout symptomatology and blood pressure improvements were found in the biofeedback condition only. HR decreased significantly from pre- to post-intervention in both experimental conditions; not in the control condition.

Several other studies were conducted among women at different stages of pregnancy. In a study piloting ambulatory biofeedback among pregnant woman with stress and anxiety, no significant reductions on anxiety or pregnancy-specific stress measures were found (Keeney, 2009). However, HRV-results indicated improved autonomic function post-intervention. In a quasi-experimental study among women with labor anxiety, self-reported fear of childbirth scores were significantly reduced in the experimental condition, whereas no significant changes were reported for the control group (Narita et al., 2018). In another quasi-experimental study, in which women with elevated prenatal depression scores self-selected their condition, a significant drop in depression scores was found in the experimental condition only (Poston, 2018). Lastly, in a quasi-experimental study among mothers in the early part-post-partum period, significant decreases in depression scores were found post-intervention for the biofeedback group compared to the control condition, as well as significant decreases in heart rate and increases in the standard deviation of the normal heartbeat interval (Kudo et al., 2014).

Furthermore, four studies were conducted in convenience samples of society members interested in stress research. In one study, HRV and baroreflex function significantly increased in the biofeedback condition; whereas performance on an impulsivity task did not improve (Schumann et al., 2019). Furthermore, the authors of a quasi-experimental study state that their multi-biosensor platform is appropriate for decreasing sympathetic arousal, increasing parasympathetic activity and enhancing overall capability of ANS modulation, although the quantitative results of the ambulatory part of the intervention remained unclear (Wu et al., 2012). At last, two pilot-studies investigating a mobile biofeedback application were found, both part of the INTERSTRESS project (Wiederhold and Riva, 2012). In one study, 196 sessions (out of which 63 sessions with biofeedback, not separately analyzed) from 38 participants were used for analysis (Serino et al., 2014). In the other study 182 relaxation sessions from 7 participants were analyzed (Gaggioli et al., 2014); potential overlap in these samples remained unclear. Results of both studies indicated significant decreases in stress level and arousal, and an increase in valence after a stress management exercise. In both studies a non-significant pre-post decrease in mean heart rate values was found.

3.5.2.2. Psychiatric populations. Finally, nine ambulatory biofeedback studies were found among psychiatric populations. Two quasi-experimental studies were performed among young adults with substance use disorder. In one study the biofeedback condition demonstrated a greater, though not significantly different, reduction in substance craving compared to treatment as usual (Eddie et al., 2015). Although significant pre-post session increases in HRV levels were found in the experimental group, no chronic HRV changes were found for both groups after the entire intervention. In a similar study comparing Progressive Muscle Relaxation (PMR) and ambulatory biofeedback, both interventions showed significant reductions in comorbid symptoms of insomnia and PTSD (Zucker et al., 2009). Compared to PMR, significantly greater reductions in depressive symptoms, increased HRV-indexes and a trend for reduced substance craving were found in the biofeedback condition.

Concerning depressed mood, anxiety and depression, one case-study evaluated the addition of ambulatory biofeedback to regular treatment of a military officer with adjustment disorder (McLay and Spira, 2009). Sleep improved significantly, though the intervention did not prevent the patient from developing more serious (PTSD-)symptoms later. Furthermore, a quasi-experimental study investigated a brief respiratory biofeedback program among patients with panic disorder. Significant decreases in panic severity, global illness severity and functional impairment, as well as a normalized amount of carbon dioxide in the exhaled air, were found post-intervention (Tolin et al., 2017). In another pilot-study among outpatients with anxiety disorder and comorbid depression, significant reductions in state anxiety, trait anxiety and trait anger were found post-intervention (Reiner, 2008). Qualitatively, reduction of stress and increased relaxation were often reported. Overall, higher compliance was associated with greater stress reductions.

Finally, two studies were conducted among patients with severe traumatic brain injury. In a quasi-experimental study, HRV indexes increased significantly compared to baseline after an ambulatory biofeedback intervention (Kim et al., 2012). Although no significant improvements were found on psychological measures, self-rating scores of the participants became more closely aligned to the perception of their family members, probably indicating enhanced emotional insight. Enlarged ability to recognize emotions was also found in two case-studies evaluating ambulatory biofeedback (O’Neill and Findlay, 2014). An increased ability to recognize frustration and to use relaxation strategies to prevent self-harm was reported in case 1; significantly enlarged self-efficacy and control to maintain a calm state in case 2.

3.6. Feasibility outcomes

Feasibility assessment varied widely, ranging from a combination of quantitative measures and extensive qualitative reports (7%), to qualitative information only (33%) or a brief report of the author’s opinion (37%). In seven studies (23%) no information regarding feasibility of the used devices was reported. In studies that (to a greater or lesser extent) assessed feasibility, the outcomes were predominantly positive (61%), positive yet critical, needing further improvement (30%) or mainly critical (9%). Most reports were positive, describing participants’ enthusiasm, perceived helpfulness and continued use of the devices. Insufficient adherence to the device, side-effects or disadvantages (e.g. sleepiness, frustration) and ethical questions were mentioned as requiring further attention.

4. Discussion

In this review, we aimed to provide an overview on effectiveness and feasibility of biofeedback and ambulatory biofeedback in (addition to) treatment of emotion regulation difficulties, in non-psychiatric and psychiatric populations. Therefore, we performed a systematic literature search, which yielded 30 relevant studies. In the majority of the studies, the addition of biofeedback or ambulatory biofeedback yielded significant improvements on one or more self-reported (stress-related) psychological measures. In half of the studies using (less frequently assessed) physiological outcome measures, significant improvements were reported as well. Adherence to the biofeedback and ambulatory biofeedback devices was assessed in only a few studies. Although the feasibility of the devices was often reported to be sufficient, most studies lacked adequate feasibility assessment. Before we discuss these effectiveness and feasibility results in more detail, we will focus on the differences between the two subsets of the included studies.

To create a broad overview of biofeedback interventions supporting daily practice aimed at enhancing emotional regulation skills, we included
both biocueing and ambulatory biofeedback studies. Although both types of interventions use ambulatory devices, biocueing is the only intervention that can actually prompt users in everyday life stressful situations to practice regulation skills. Therefore, biocueing can be seen as a further developed extension of ambulatory biofeedback. Matching the relatively new status of this type of biofeedback, our review only identified four studies that could be categorized as biocueing. Two of these studies actually ‘cued’ participants in everyday life, when the physiological measures reached a specific range, possibly indicating emotional stress. (McHugh et al., 2016; Mackintosh et al., 2017). In another study, the possibility to prompt regulation skills was one of several options; actual use of this function was not reported (Mackintosh et al., 2017). In these studies, the biocueing system was used to initiate just-in-time (self-)calming strategies and practice emotion regulation skills to prevent disruptive or aggressive behavior in everyday life. In one study, the system was also used to reflect on triggers at a later time and to discuss the connection between physiological values and actual behavior (McHugh et al., 2016). Taken into account the limitations of the study designs, these pivotal studies provide preliminary evidence that interoperative awareness, emotion regulation and self-control can be facilitated by monitoring physiological measures and biocueing in everyday life, in line with the results of previous studies (Füstös et al., 2013; Gros et al., 2016; Novaco, 2007). In studies investigating ambulatory biofeedback, portable or wearable devices were provided to support out-of-session practice, often within a more traditional biofeedback training paradigm. In most studies, applications were used in combination with a personal computer, at home. In a few studies, applications enabled the participants to practice in everyday life, using their smartphone, though without the possibility to be cued in ‘at risk’-situations. Although biocueing and ambulatory biofeedback differ in their applicability in everyday life stressful situations, both interventions showed significant results. Based on this review, biocueing seemed particularly useful for participants with low emotional awareness and externalizing behavior, whereas ambulatory biofeedback was mainly used among participants with internalizing, stress-related problems. Concerning effectiveness, pre-post results demonstrate (predominantly) positive effects of biocueing and ambulatory biofeedback. The distribution of significant and non-significant results seemed not to differ between non-psychiatric and psychiatric populations. Significant improvements were mainly found on psychological measures, such as decreased perceived emotional, social or organizational stress and reduction of anxiety symptoms. In the smaller number of studies that included physiological measures, a slightly smaller percentage of significantly improved physiological functions, known to be related to adequate emotion regulation (Crockett et al., 2017), was found. Whether this difference might be due to, for example, the duration of the intervention or social desirability, could not be derived from these studies. Furthermore, the majority of the (quasi-)experimental ambulatory biofeedback studies showed significantly better results compared to control conditions in non-psychiatric samples. Compared to other mind-body interventions, such as mindfulness meditation and physical exercise, ambulatory biofeedback was equally effective (De Bruin et al., 2016) or showed interchangeable results (May et al., 2019). In psychiatric samples, only three (quasi-)experimental studies included control conditions. In one study, ambulatory biofeedback was proven more effective than progressive muscle relaxation (Zucker et al., 2009); in two other studies, biofeedback interventions did show effective results, but did not perform better than standard treatment, as was reported (May et al., 2015; McNutt et al., 2017). Based on the results of this review, biocueing and ambulatory biofeedback are considered useful treatment supplements, especially in case of low emotional awareness or insufficient response to other interventions. Future studies should further evaluate their added value compared to other interventions, preferably using (randomized) controlled designs. Looking at the practical features and feasibility of the used devices, the StressEraser and emWave, designed to improve HRV (Lehrer, 2013) and coherence (McCraty and Zayas, 2014) respectively, were applied most in the ambulatory biofeedback studies. In the biocueing studies, HR was used instead of HRV. This choice might have been made since HRV-measurement using unobtrusive wrist worn sensors is sensitive to motion artefacts and may therefore be less reliable for biocueing in everyday life (Van Lier et al., 2019), whereas HR-measurement can be corrected for movement using currently available wrist worn wearables and appropriately designed applications (Derks et al., 2019). Furthermore, among the studies that were categorized as biocueing, patients more often self-determined the time, frequency and place of practice. The method and duration of training was less structured and often shorter than in the ambulatory biofeedback studies. In most of these studies, training consisted of several hours or even multiple days and participants were encouraged to practice daily. These differences in duration of training and actual use complicate interpretation of the results. For example, if controlled for actual (frequency of) use, the effects of the interventions could have been larger. Furthermore, the usability of the devices was considered sufficient to good based on author reports and (concise) qualitative user feedback. However, only in a few studies feasibility and adherence were adequately measured, using usability measures and qualitative interviewing. This is a clear shortcoming of these studies, since it is known that many users stop using a health application after two weeks, especially when their preferences and goals are not met (Torous et al., 2019). The results of this systematic review should be viewed in the light of several limitations. First, our systematic search yielded only a small number of biocueing studies, in which we had a specific interest given the focus of this review on recent advances using new, wearable technology. Second, we did not limit the search strategy to specific study designs. Although this allowed us to provide a complete overview, the effectiveness outcomes should be interpreted with caution due to the small sample sizes and low to medium quality of the majority of the papers. Third, due to the variation in designs, procedures and outcome measures of the selected studies, a reliable meta-analysis could not be performed. Therefore, studies were systematically and narratively reviewed. Fourth, some studies, especially those investigating new approaches and devices, lack a clear description of implementation procedures, complicating future replication. Furthermore, in several studies using more established devices, such as the StressEraser and emWave, protocols were not well specified; therefore, not all procedures might be implemented according to the corresponding theoretical principles. Fifth, some of the papers received commercial funding, which might have influenced selective reporting of (positive) study results. Sixth, it is important to note that outcomes of studies performed among non-psychiatric populations cannot be generalized to psychiatric populations. Given the small number of studies using a psychiatric sample, no firm conclusions can be drawn regarding psychiatric patients with (severe) emotion regulation difficulties. 4.1. Methodological considerations To prevent an unwanted proliferation of biofeedback applications without evidence of their effectiveness, more and better designed studies should be performed to validate future interventions (Bakker et al., 2016; Leigh and Platt, 2015). Among non-psychiatric samples independent, larger and (randomly) controlled designed studies should be performed. Meanwhile, case series and single case experimental designs should be considered to investigate effectiveness in small psychiatric samples in which more robust research designs are not feasible (Smith, 2012). These individually tailored designs can also be used to evaluate (small) changes in applications or to detect the optimal duration of training and practice of interventions (Maric and Van der Werff, 2020). Furthermore, to validate the effects of biofeedback interventions, self-reported psychological measures should preferably be combined with observed behavioral measures and physiological measures. Agreement between research groups on validated outcome measures and
implementation procedures is also recommended. Regarding feasibility, adherence to and usability of the applications should be more thoroughly assessed by tracking actual use of the devices, administering system usability questionnaires and qualitative interviewing. Finally, although beyond the scope of this article, ethical questions about surveillance and privacy, agency and autonomy should be carefully considered (Sharon, 2017; Torous et al., 2019).

4.2. Conclusions and future directions

Despite the rapid developments in mobile health technology and the increasing number of wearable, non-invasive devices that can be used for everyday stress regulation as an addition to psychotherapy, it seems that these applications have only sparsely found their way into clinical research and practice (Derks et al., 2019). Although modern biosensors showed their value in monitoring and predicting behavior (Goodwin et al., 2019; Kuipers et al., 2012; Roushan et al., 2019) hardly any studies were found investigating biofeedback in an everyday life context. Still, the limited but promising results of the included studies seem to suggest that biofeedback and ambulatory biofeedback interventions can aid emotion or stress regulation, both on a psychological and physiological level. As such, the results support the potential of biofeedback and ambulatory biofeedback to enhance emotion regulation by delivering individually tailored feedback. Fitting within the actual trend of personalized medicine, individual differences should be taken into account: whereas biofeedback may help some patients to pay attention to physiological sensations in everyday life situations; others might get easily overwhelmed or irritated by this information (Cornet et al., 2017; Howell et al., 2018; Owens and Cribb, 2019). In the near future, we expect biofeedback research to be focusing more on biofeedback as a just-in-time adaptive intervention (Riley et al., 2015; Wang and Miller, 2019) supporting users to practice regulation skills in everyday stressful situations.

Declaration of competing interest

This systematic review was performed within a larger research project, financially supported by the Dutch Ministry of Justice and Security. The subsidizing party had no involvement in the creation of this review article. All authors have declared that they have no competing or potential conflicts of interest.

Acknowledgements

The authors are grateful to Tale Evenhuis for conducting the literature search. We also would like to thank Eline Middelhoven, Nine van Eerde and Lisette Stegeman, at the time of this research project master students from different universities in the Netherlands, for performing independent quality assessment of the selected studies.

Appendix A

Search strategy PubMed:

("Biofeedback, Psychology*[MeSH] OR "Feedback, Psychological*[MeSH] OR Biofeedback.tw OR Biosensor*tw OR (Sensor* [tw AND Integrat*[tw]) OR Feedback[tw] OR "Real time" [tw OR Realtime[tw] OR Monitoring[tw] OR Monitored[tw] OR Wearable* [tw] AND ("Heart Rate*[MeSH] OR "Heart Rate*[tw] OR Heart Rates*[tw] OR "Heart Rating*[tw] OR "Heart Ratings*[tw] OR "Heart frequency*[tw] OR "Heart frequencies*[tw] OR "Heart beat*[tw] OR "Heart beats*[tw] OR Heart beating*[tw] OR "Heart beatings*[tw] OR "Cardiac frequency*[tw] OR "Cardiac frequenc* [tw] OR "Ventricle Rate*[tw] OR "Ventricle Rates*[tw] OR "Ventricle Ratings*[tw] OR "Pulse Rate*[tw] OR "Pulse Rates*[tw] OR "Pulse Rating*[tw] OR Pulse Ratings*[tw] OR "Cardiac Chronotrop*[tw] OR "Cardiac Chronotropies*[tw] OR "Cardiac Chronotropism*[tw] OR "Cardiac rate*[tw] OR "Cardiac rates*[tw] OR "Cardiac rating*[tw] OR "Cardiac ratings*[tw] OR HRV*[tw] OR "(Galvanic Skin Response*[MeSH] OR Dermogalvanic*[tw] OR ((Skin [tw] OR Cutaneous*[tw]) AND (Conduct*[tw] OR Resist* [tw] OR Impedanc*[tw] OR Galvanic*[tw])))) OR ("Respiration*[MeSH] OR Breathing[tw] OR Respirat*[tw] OR Inhalat*[tw] OR Exhalar*[tw])) AND ("Aggression*[MeSH] OR Psychomotor Agitation*[MeSH] OR "Inhibition (Psychology)*[MeSH] OR "Impulsive Behavior*[MeSH] OR Aggression*[tw] OR Aggressiv*[tw] OR Inhibition*[tw] OR "Disinhibition*[tw] OR "Emotionality*[tw] OR "Emotion Regulation*[tw] OR "Emotion Regulations*[tw] OR "Emotion Regulating*[tw] OR "Emotion Regulating*[tw] OR "Emotion Management*[tw] OR Impulsiv*[tw] OR "Impulse Control*[tw] OR "Self Regulation*[tw] OR "Self Regulating*[tw] OR "Self Control*[tw] OR "Self Controlling*[tw] OR Agitat*[tw]) AND ("2007*[PDAT]: "2019*[PDAT])"

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpsycho.2020.11.009.

References

Bakker, D., Kazarinov, N., Rickwood, D., Rickard, N., 2016. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment. Health 3 (1), e7. https://doi.org/10.2196/mental.4984.
Berenbaum, H., Bredemeier, K., Thompson, R.J., Boden, M.T., 2012. Worry, anxious depression, and emotional styles. Cognit. Ther. Res. 36, 72–80. https://doi.org/10.1007/s10608-010-9226-6.
Boden, M.T., Thompson, R.J., 2015. Facets of emotional awareness and associations with emotion regulation and depression. Emotion 15 (3), 399. https://doi.org/10.1037/emo0000597.
Boyce, W.T., Quas, J., Alkon, A., Smider, N.A., Essex, M.J., Kupfer, D.J., 2001. Autonomic reactivity and psychopathology in middle childhood. Br. J. Psychiatry 179, 144–150. https://doi.org/10.1192/bjp.179.2.144.
Cornet, L.J.M., Mandersloot, M., Pool, R., De Kogel, C., 2016. Neurowetenschappelijke toepassingen in de jeugdstraftochten. (Onderzoek en beleid: No. 318). Boom Criminologie, Den Haag, The Netherlands, pp. 148–152.
Cornet, L.J.M., Mandersloot, M., Pool, R., De Kogel, C.H., 2017. De ‘zelfmetende justitiabele: Een verkennend onderzoek naar technologische zelfmeetmethoden binnen justitiële context. (Cahiers 2017; No. 17). Wetenschappelijk Onderzoeks- en Documentatie Centrum van het Ministerie van Justitie en Veiligheid, Den Haag, The Netherlands, pp. 25–94.
Crockett, J.E., Gill, D.I., Cashwell, T.H., Myers, J.E., 2017. Integrating non-technical and technological peripheral biofeedback in counseling. J. Spiritual Ment. Health 39 (2), 163–179. https://doi.org/10.17744/mehc.39.2.06.
Cumpaniasio, D., Mazefsky, C., Stedman, A., Peura, C., Tian, P., Guo, Y., Isamidis, S., Birmaher, D., Siegel, M., 2018. Pleth, Goodwin biomarkers for prediction of imminent aggression in minimally verbal children with autism spectrum disorder. J. Am. Acad. Child Adolesc. Psychiatry 56 (10), S256–S257. https://doi.org/10.1016/j.jaac.2017.09.295.
De Bruin, E.L, Van der Zwan, J.E., Bogels, S.M., 2016. A RCT comparing daily mindfulness meditations, biofeedback exercises, and daily physical exercise on attention control, executive functioning, mindful awareness, self-compassion, and worrying in stressed young adults. Mindfulness 7, 1182–1192. https://doi.org/10.1007/s12671-016-0561-5.
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), e13420. https://doi.org/10.1111/pop.13420.
Derks, Y.P., Klaassen, R., Westerhof, G.J., Bohlmeijer, E.T., Noordzij, M.L., 2019. Development of a feasibility study for emotion regulation training in children with borderline personality disorder: multicycle usability testing study. JMIR mHealth uHealth 7 (10), e13479. https://doi.org/10.2196/13479.
Erdogmus, D., Siegel, M., Goodwin, M.S., 2018. Physiological biomarkers for emotion regulation and depression, emotion regulation and depression, Emotion 15 (3), 399. https://doi.org/10.1037/emo0000597.
Foust, J., Gramman, K., Herbert, B.M., Pollatos, O., 2013. On the embodiment of emotion regulation: interoceptive awareness facilitates reappraisal. Soc. Cogn. Affect. Neurosci. 8 (8), 911–917. https://doi.org/10.1093/scan/nos089.

International Journal of Psychophysiology 159 (2021) 94–106

104
