A Multidimensional Perspective on Social Anxiety Disorder and Its Treatment Through Internet-Based Cognitive Behavioral Therapy

Inauguraldissertation der Philosophisch-humanwissenschaftlichen Fakultät der Universität Bern

zur Erlangung der Doktorwürde

vorgelegt von

Dajana Šipka

Baden (AG)

Begutachtung:

Prof. Dr. phil. Thomas Berger (Hauptgutachter)

PD Dr. phil. Tobias Krieger (Zweitgutachter)

Mentorat:

PD Dr. phil. Anja Gysin-Maillart

Selbstverlag, Bern

Januar 2025

Von der Philosophisch-Humanwissenschaftlichen Fakultät der Universität Bern
auf Antrag von Prof. Dr. phil. Thomas Berger und PD Dr. phil. Tobias Krieger angenommen.

Bern, den 18. Dezember 2024 Der Dekan Prof. Dr. Elmar Anhalt

Acknowledgment

First and foremost, I want to thank my supervisor, Thomas Berger. Thank you for supporting me from the very start and giving me the creative space and time to try out all my ideas. This freedom and your trust gave me the opportunity to try out different areas of research, to independently collaborate outside our main project, OPTIMIZE, to gain expertise through numerous courses, and finally to be able to do my psychotherapy training along my PhD. Thank you for all your time and patience through the numerous meetings we had, throwing ideas back and forth and discussing statistical methods to oblivion. Additionally, to Tobias Krieger, thank you for being my second advisor. Another special thank you goes out to Yosuke Morishima. Thank you for taking me in as an intern back in 2020, for giving me the opportunity to implement my very first research question on my own, and for supporting me throughout my whole research internship and beyond. You are one of the main reasons why I wanted to pursue a career in research in the first place; with your knowledge and enthusiasm, I started to have fun and found passion in programming and research. A big thank you goes out to Anja Gysin-Maillart. You are one of my greatest role models as a researcher. Thank you for mentoring me through the years and for the fact that I could always ask you for advice. I would also like to thank all my co-authors. It was a pleasure to collaborate with you all and to learn from your expertise; thank you for all the time and effort you invested in our projects. A wholehearted thank you goes out to our whole lovely institute. I hereby especially want to thank Laura Bielinski, Gwen Wälchli, and Andrej Skoko. Laura, to me, you are a great role model as an exceptional researcher, and, most importantly, a dear friend. Thank you for making work feel more than just a workplace. Gwen, thank you for all the fun and exuberance we had during the countless weekends in our psychotherapy training; with you, it made the hard work so much easier, and it was a pleasure to have you by my side at all our different stages in our careers. Andrej, thank you for all your small and big visits to my office, all our coffee breaks, and the hilarious hours we spent together at uni. A big shoutout goes to all our master's students and trainees from the OPTIMIZE project, who put so much effort into the data acquisition process. Furthermore, I want to thank all my friends and family who went along on this journey. Especially, I want to express my sincere gratitude to my wife, Faye Dittrich. Thank you for always believing in me and lending me your ear, for sharing my passion and enthusiasm for my research, and for making even the darkest hours light. To my sister, Jovana Šipka, who has always been and is my best friend in the whole world. Thank you for being with me; no matter how many kilometers we lived apart, I never felt alone. To my oldest heartfelt friends, Salome Häuselmann and Rubén Fuentes Artiles, whom I have known for almost my whole life, thank you for being my chosen family. A special shoutout goes to Taylor Swift's music, which accompanied me through all my PhD years and brought me so much joy. Lastly, I want to thank all our brave participants for all their patience and effort, and without whom my dissertation would not have been possible.

Abstract

Throughout the day, we gather and talk with other people and need to perform in different social situations, which can be as small as only signing a document in front of another person. In all those social situations, it is a common human need to come across in a positive way. If we don't act, behave, or perform in an expected way, social anxiety (SA) and embarrassment can arise. If a person fears being (even potentially) evaluated in a negative way or that others might see their anxiety symptoms in social situations, we speak of a social anxiety disorder (SAD). The aim of this dissertation was to provide a multidimensional perspective on SAD and its treatment through internet-based cognitive behavioral therapy (ICBT). The multidimensional perspective consisted of an emotion-focused perspective and a structural perspective. From the emotion-focused perspective, the emotion of embarrassment was thematized (Article I), as well as its connection to SA and SAD. From the structural perspective, SA and SAD were analyzed with a focus on the two questionnaires Social Phobia Scale (SPS) and Social Interaction Anxiety Scale (SIAS) (Article II), by using factor analyses. The OPTIMIZE project was the main project of this dissertation, where the main components of an ICBT for SAD were examined (Articles III & IV). Concluding, the discussion summarized the outcomes from the emotional and structural perspective and merged the results with the ICBT for SAD, suggesting how ICBT can be expanded with the multidimensional perspective. The particular focus lay on how the knowledge of embarrassment, which is strongly intertwined with SAD, can be used to create a more holistic view of the disorder and its treatment in research as well as clinical practice.

List of Articles

Article I:

Šipka, D., Vlasenko, B., Stein, M., Dierks, T., Magimai-Doss, M., & Morishima, Y. (2025). Multidisciplinary Characterization of Embarrassment through Behavioral and Acoustic Modeling. [Manuscript submitted for publication].

Article II:

Šipka, D., Brodbeck, J., Schulz, A., Stolz, T., & Berger, T. (2023). Factor structure of the Social Phobia Scale (SPS) and the Social Interaction Anxiety Scale (SIAS) in a clinical sample recruited from the community. *BMC Psychiatry*, 23(646), 1–11. https://doi.org/10.1186/s12888-023-05142-8

Article III:

Lopes, R. C. T., Šipka, D., Krieger, T., Klein, J. P., & Berger, T. (2021). Optimizing cognitive behavioral therapy for social anxiety disorder and understanding the mechanisms of change: Study protocol for a randomized factorial trial. *Internet Interventions, 26,* 100480. https://doi.org/10.1016/j.invent.2021.100480

Article IV:

Šipka, D., Lopes, R., Krieger, T., Klein, J. P., & Berger, T. (2025). Active Components in Internet-Based Cognitive Behavioral Therapy for Social Anxiety Disorder: A Randomized Full Factorial Trial. *Psychotherapy and Psychosomatics,* 1-20. https://doi.org/10.1159/000542425

Table of contents

1. Introduction

Social interactions and encounters are a common part of most people's daily lives. We greet our neighbors, have a coffee break with a co-worker, or meet friends after work. According to the basic needs definition by Grawe (1998), the need for attachment is one of the four basic needs. It is a human need to connect. Since human connection is so important, it can also be accompanied by fear and anxiety. Social anxiety (SA) is a well-known phenomenon that occurs in daily life. One of the most common forms of SA occurs in public speaking situations (e.g., holding a presentation in front of a crowd) (Gallego et al., 2022; Leary & Kowalski, 1997). SA arises from the fear of possibly being negatively evaluated by others in social situations. One of the emotions associated with SA is embarrassment, which occurs when one feels one's desired social image is threatened (Miller & Tangney, 1994; Tangney et al., 1996). Both SA and embarrassment share a core aspect in their common definitions: the fear of being negatively evaluated by others. While people perceive these kinds of experiences as unpleasant and negative, they are, to a certain degree, socially functional. Bystanders tend to react helpfully in embarrassing situations and tend to rate the person who shows embarrassment favorably (Miller, 2012). But when the fear or anxiety of negative evaluation and with that, more generally, the fear of one or multiple social situations causes impairments in daily life (among other symptoms) and becomes persistent, it is called a social anxiety disorder (SAD) (American Psychiatric Association [APA], 2013a). SAD is one of the most common anxiety disorders (Stein et al., 2017) and tends to progress chronically, causing impairments across different areas of life (Fehm et al., 2005; Ruscio et al., 2008). There are different self-report questionnaires that measure SA or, within a diagnosis process measure, SAD symptoms (cf. Caballo et al., 2013). Some of the most common self-report questionnaires are the Social Phobia Scale (SPS) and Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998), which are often presented together (Thompson et al., 2019). These questionnaires differentiate between performance-related situations (measured with the SPS) and interaction-related situations (measured with the SIAS). Even though the questionnaires are widely used, there is no consent about the underlying factor structure, and therefore, it is unclear how the questionnaires should actually be used and interpreted. Depending on the factor structure, the questionnaires could be used independently, only their total sum, or a combination of both.

There are different effective pharmacological and psychotherapeutic treatment options for SAD (Mayo-Wilson et al., 2014). Cognitive behavioral therapy (CBT) is considered the first-line treatment for SAD (National Institute for Health and Care Excellence, 2013). Over the years, a rising number of internet-based cognitive behavioral therapy (ICBT) treatments have appeared as an alternative to face-to-face (f2f) CBT. Many randomized controlled trials (RCTs), reviews, and meta-analyses have shown ICBT to be as effective as f2f treatments (Andersson et al., 2014; Boettcher et al., 2013; Carlbring et al., 2018; Clark et al., 2023; Guo et al., 2021; Olthuis et al., 2016) and, in

general, to be effective in significantly reducing SAD symptoms (e.g., Andrews et al., 2018; Schulz et al., 2016; Stolz et al., 2018). Despite the positive results of CBT and ICBT for SAD, there is still a substantial proportion of patients left who do not profit from the treatment (Boettcher et al., 2013). One of the reasons for limited efficacy is the lack of understanding of how the ICBT treatment for SAD works (Kazdin, 2017). Therefore, one of the goals of this dissertation's main project (OPTIMIZE) was to find the active components of ICBT for SAD.

The main aim of this dissertation was to provide a multidimensional perspective on SAD and its treatment through ICBT. This dissertation consists of four Articles (I – IV, see List of Articles). Article I revolved around embarrassment and how it can be modeled with the help of subjective measures and machine learning techniques. Article II examined the underlying joint factor structure of the SPS and SIAS. Article III is the study protocol of the OPTIMIZE study. Article IV presented the main outcomes of the OPTIMIZE study, where the active components of ICBT for SAD were investigated.

2. Theoretical Background

2.1. Definition of Embarrassment

Embarrassment is a self-conscious and prosocial emotion that many people encounter in their daily lives (Maire & Agnoletti, 2020). Tracy & Robins (2004) proposed in their theoretical model of self-conscious emotions three things that need to be given to feel embarrassed: attention to public self, an appraisal that identity goals are relevant in a situation, and being incongruent (i.e., how a person wants to appear publicly vs. how they are perceived), and lastly internal attributions. In order to be able to feel embarrassment, one has to be able to imagine what others think of us; therefore, broadly speaking, embarrassment can be seen as a social experience (Miller, 2010). According to the social evaluation model, embarrassment occurs when a person's desired social image is threatened through an unwanted event (Miller, 1995; Miller & Tangney, 1994). In these situations, the affected people anticipate negative consequences through cognitive appraisal of this situation (Harris, 2006). In order to feel embarrassment, an imagined or real audience must be present (Goffman, 1963; Tangney et al., 1996). The dramaturgic model, on the other hand, proposes that embarrassment can occur when the script of a social encounter is simply disrupted, and the affected person is left uncertain about their role in the situation (Miller, 1995). When people sing *happy birthday* to a person, or someone receives too many compliments, there is no social transgression, and this person is in the middle of positive attention but still feels embarrassed (Harris, 2006; Miller, 2010). Such situations can be explained through the dramaturgic model, where the affected person does not fear negative evaluation but rather does not know what to do. Since most studies that were discussed in this dissertation examined embarrassment in the context of social transgression and since this is one of the core fears of SAD, this dissertation thematized embarrassment mainly from the social evaluation perspective. Another important distinction for this dissertation is between personal vs. empathic embarrassment. Personal embarrassment is experienced for oneself, while empathic embarrassment is experienced for someone else (Stocks et al., 2011). Since the focus of this dissertation was SAD, only personal embarrassment was considered.

In the past, embarrassment has been regarded as a part of shame. Today, embarrassment is considered an emotion on its own (Miller & Tangney, 1994; Tangney et al., 1996; Wan & Wyer, 2020).[1] Nevertheless, it is important to establish the differences between these two emotions since they are closely related. The difference between shame and embarrassment can be summarized in the following way, according to the popular paper from Tangney et al. (1996): Shame is a more intense emotion that occurs tendentially with more serious and moral transgressions. It is tied to the perception of one's core self, respectively, the core self is perceived as deficient and negative. This

[1] Aside from embarrassment and shame, guilt is an additional emotion that shares similarities with those two emotions (Bastin et al., 2016; Tangney et al., 1996). However, since guilt was less relevant to the overall topic of this dissertation, it was not discussed any further.

results in a broader and more enduring negative attribution of oneself. Shame can also be experienced alone, while embarrassment needs an audience (either present or imagined). Embarrassment is less intense, shorter, and occurs in trivial transgressions. Furthermore, embarrassment is more likely to occur in the presence of strangers or acquaintances, while shame occurs more with familiar people whom we share an affective connection with.

Miller (2012) showed that embarrassment occurs most often in so-called "normative public deficiencies" (p. 187) situations where people unintentionally behave against consensual social norms of behavior, by, for example, acting in a clumsy way (e.g., spilling something). This was also supported by Keltner (1996), who found that physical mishaps (e.g., slipping) were most frequently recalled by their participants for eliciting embarrassment. Furthermore, if people feel embarrassed, they most likely act like nothing happened, and second-most likely try to remediate the mishap or take it with humor (Miller, 2012). It is important to note here, that the samples used in the study of Miller (2012) are from the US and that the results cannot simply be applied to non-western cultures.

2.1.1. Embarrassment and the Impact on the Voice

Even though embarrassment is such a common emotion in our daily lives, it is grossly neglected in emotion research (Simon-Thomas et al., 2009). One of the reasons might be that embarrassment is a rather complex emotion in comparison to other basic emotions and differs from them in numerous ways (e.g., embarrassment requires mental self- and other representation, one has to identify a whole set of facial expressions, body posture, and head movement) (Caillaud et al., 2020). Additionally, Ekman & Cordaro (2011) noted that embarrassment shows some of the 12 criteria for basic emotions, but not all. They mentioned that embarrassment showed some overall typical actions, facial expressions, hand movements, gaze, and posture but that more research was needed. The traditionally used methods are probably not sufficient enough to capture embarrassment. Cowen & Keltner (2021) used a computational approach to provide a more differentiated emotion taxonomy, where they used different perspectives (e.g., facial-bodily expressions, neurological research, mammalian behavior). One of their points of criticism of the basic emotion theory was that there are only a limited number of emotions assumed with defined borders and measured with particular methods (e.g., the emotion recognition task), which cannot differentiate between the expression and the actual experience a person has. Therefore, they suggested capturing emotion in a broader and more complex way by including, for example, the whole body, face, and voice. With the emergence of new and complex technology, this has become possible.

Interestingly, the above-mentioned articles do not mention the change in the voice but rather visible changes, like the change in posture and facial expression. Keltner et al. (2019) provided an overview of evidence for the recognition of different emotions for four modalities (i.e., facial, head or bodily action, voice, touch, and music). While facial, head, or bodily actions were strongly

represented, there was only one study listed which examined the voice. The same was shown in the literature research of Article I, that voice analyses are seldomly used in comparison to, for example, neurological (see Bastin et al., 2016, for a system review of neurobiological studies) and physiological measures (e.g., Müller-Pinzler et al., 2012), self-report measures, and rating methods. Low et al. (2020) showed in their systematic review of automated speech processing in psychopathology that there are already many studies on schizophrenia, depression, and bipolar personality disorder, but studies on anxiety disorders are still lacking. The studies in their review are just a glimpse of the possibilities of what such automated speech processing can achieve. After all, the voice is rich in information about the emotional state, mental health, and many other aspects (Kadali & Mittal, 2020; Low et al., 2020).

As with basic emotion research, voice analysis studies have been criticized for using acted emotions instead of natural speech (Drahota et al., 2008). Therefore, in Article I, natural speech sections were used, in order to provide as naturalistic a setting as possible. With the help of machine-learning-based voice analysis, one does not rely on subjective evaluations (e.g., ratings, self-report measures) (Weeks et al., 2012), nor, in some cases, on even the spoken content (Burkhardt et al., 2005). Additionally, the voice can be analyzed on different levels (e.g., in Article I, embarrassment was described on a dimensional as well as categorical level). Since there is very little literature on the impact of embarrassment on the voice, Article I attempted to describe embarrassment exploratorily from different perspectives, using acoustic data to model embarrassment.

2.2. Definition of SAD

The International Classification of Diseases 11th revision (ICD-11; World Health Organization, 2024) and the Diagnostic and Statistical Manual of Mental Disorders 5th edition (text rev.) (DSM-5-TR; APA, 2022) define SAD as an excessive and marked fear or anxiety that occurs in one or more social situations where individuals perform, interact, or are observed doing something. People with SAD fear that they will be evaluated negatively by either behaving in a certain way or by showing anxiety symptoms. They then try to avoid the feared social situation(s) and/or endure them. The fear or anxiety is persistent for at least several months (DSM-5-TR defines this precisely with at least six months). Finally, fear, anxiety, or avoidance must lead to significant distress or impairment in one or more important areas of functioning (e.g., family or social life, occupation, etc.). The lifetime prevalences for SAD vary across the globe (Stein et al., 2017). Concerning the Western European region, where the samples of Articles I, II, and IV are from (specifically German-speaking countries), Stein et al. (2017) found a lifetime prevalence of 3%. A large epidemiological study from the US showed a lifetime prevalence of 12.1% for SAD (Ruscio et al., 2008). SAD is highly persistent and usually appears with comorbidities (Fehm et al., 2005; Spence & Rapee, 2016). The

great majority of affected people develop the disorder before reaching 18 years of age (Solmi et al., 2022).

One important difference between the DSM-5-TR and ICD-11 is that the DSM-5-TR (since the DSM-5, APA 2013a) provides the following specification: "Performance only: If the fear is restricted to speaking or performing in public" (p. 231). This is just one attempt of many to classify SAD. Previously, in the DSM-IV (APA, 1994), there was the specification of generalized SAD, where people with generalized SAD feared most social situations. In the end, the generalized SAD specification was removed from the DSM-IV since there was not enough evidence to support this specification. However, there is still an ongoing debate about the current performance-only specification and its clinical utility (D'Avanzato & Dalrymple, 2016; Hyett & McEvoy, 2018). Fuentes-Rodriguez et al. (2018) suggested, among other researchers (e.g., Ruscio, 2010; Tei et al., 2020), that SAD should be looked at dimensionally (e.g., according to Fuentes-Rodriguez et al., from mild SAD (where performance-only would be) to severe SAD) or a hybrid form (Skocic et al., 2015) instead of categorically. Chapter 2.5. will describe the dimensionality of SAD in more depth.

2.3. Embarrassment, SA, and SAD

Showing embarrassment can elicit a favorable impression from others (Miller, 2012). Dijk et al. (2009), for example, found that participants reacted more sympathetically towards actors who blushed after a mishap (e.g., spilling coffee on someone) or transgression (e.g., jumping a queue) in comparison to non-blushing actors. These examples show that displaying embarrassment promotes harmonious relationships, acting in favor of the affected person and rectifying the situation. Generally speaking, embarrassment serves an affiliating social function (Fischer & Manstead, 2016; Miller, 2009, 2010). Although embarrassment has a prosocial function, it also leads to negative self-evaluation (Bas-Hoogendam et al., 2018). This might be one reason why many people fear embarrassing circumstances, even though most embarrassing events end positively.

As was shown in the previous chapter, when the fear of negative evaluation by others becomes clinically relevant, among other symptoms, it is considered an SAD. Hofmann et al. (2006) already proposed almost 20 years ago that embarrassment was part of SAD. In a later paper Hofmann et al. (2010) described the fear of violating social norms in SAD as closely related to these violations in embarrassment. As with SAD, embarrassed people feel like they are being the center of attention after a violation (Miller, 2010). Several studies have shown that people who score higher on SA get more easily and more intensely embarrassed than people with lower SA scores (Leary & Hoyle, 2013; Rozen & Aderka, 2023). The same was found in Article I of this dissertation (see Chapter 3.1.). Embarrassability was also found to be connected to high sensitivity to social norms (Miller, 1995). In summary, embarrassment, SA, and SAD share the same core: the fear of being negatively evaluated by others.

However, even though they share many similarities, there are also distinct differences between embarrassment and SAD. Three differences by Miller (2010) that are relevant to this dissertation will be discussed further, namely: phenomenology, timing, and behavioral sequels. It is important to note that these differences refer to SAD and SA within SAD and not SA in healthy people. This is an important differentiation since SA has numerous benefits and prosocial functions, but these get lost with the disorder. Concerning phenomenology, SAD and embarrassment feel differently; SAD is primarily linked to fear, whereas embarrassment is linked to awkwardness. Additionally, SAD launches different processes entering a dreaded social situation, for example, anticipation of negative outcomes or reactions, heightened attention to oneself, and post-processing (see Chapter 2.6.2. for more information on these processes). Concerning timing, SAD is foreseeable. The affected people anticipate the feared situations, and this state lasts longer, whereas embarrassment occurs mostly surprisingly after, for example, a transgression and lasts shorter than SA in SAD. Concerning the behavioral sequels, the experience of high SA leads to behavioral depletion, where people with SAD try to avoid the feared situations or use safety behavior that might even lead to the feared outcome (e.g., trying to hide sweat by wearing covering clothes, which leads to more sweating), sometimes avoid eye contact, speak less fluently, etc. On the other hand, embarrassment leads most often to socially repairing actions, like the use of humor or an apology. While people react mostly positively to the display of embarrassment, the display of SA in SAD leads to less positive evaluation. In conclusion, SA within SAD is mostly destructive, whereas embarrassment, in a moderate amount, usually has a constructive, prosocial effect.

A study from Čolić et al. (2020) showed that embarrassment led to repeated thinking about an embarrassing event (i.e., post-event processing) and difficulties in forgetting the event. However, participants with SAD and major depression disorder (MDD) interpreted important social interactions more often as subjectively embarrassing in comparison to a healthy control group. One of the authors' interpretations was that people with higher SA and suffering from either SAD or MDD perceived social situations more often wrongly since people with higher SA underestimated their appearance and interpreted ambiguous social situations negatively; hence, the perception of embarrassment. Furthermore, they proposed that the feeling of embarrassment might mediated the relationship between SA and the probability of post-event processing. Another interpretation was that people with SAD and MDD actually had more embarrassing encounters due to the lack of social skills, but this explanation was ruled as less probable according to the literature. The discard of this interpretation is also supported by the study of Miller (2009), where embarrassability did not correlate with a person's social skills but was highly correlated with the fear of negative evaluation (Miller, 2009).

Bas-Hoogendam et al. (2018) found that social norm violations were rated as more inappropriate and embarrassing by participants with higher SA scores. Interestingly, while intentional violations were rated as more embarrassing in comparison to unintentional in participants with low SA, this difference was no longer significant for participants with high SA. Concerning the ratings of

the inappropriateness of the action, participants with low and high SA rated intentional violations as more inappropriate in comparison to unintentional. According to Bas-Hoogendam et al. (2018), these results support the discrepancy between cognitive and affective evaluation in people with high SA and with SAD in general: While the affected people can cognitively differ between intentional (e.g., a conscious transgression) and unintentional violation (e.g., a mishap), they fail to do so emotionally. They further suggested that embarrassment may contribute to the development and maintenance of SAD, where often repeated and intensive experiences of embarrassment lead to a negative self-evaluation, which then leads to the overestimation of the importance of mishaps, dysfunctional concerns about others' judgment, and passive behavior.

In conclusion, while embarrassment is connected to a rather prosocial effect in social situations, SAD elicits destructive effects. Nevertheless, embarrassment and SAD share numerous similarities, and embarrassment seems to be an important part of SAD. Embarrassment is part of some vicious cycles in SAD (e.g., repeated thinking about the situations, post-event processing) and can be a causes of negative self-evaluation. On the other hand, a high SA leads to a distorted perception of the self and the social situation, which then can fuel the feeling of embarrassment. These results imply that SAD and embarrassment are highly intertwined and interact with each other in different ways in social situations.

2.4. The Dimensions and Measurement of SAD

As was shown in Chapter 2.3., the assumptions about the possible structure and dimension of SAD varied over the years. The research in the above-mentioned discussion revolved around clinical usefulness and how, for example, the performance-only specification was related to SAD (cf. D'Avanzato & Dalrymple, 2016). Article II of this dissertation addressed this topic from a different point of view: using questionnaires, the structural perspective on SAD was considered. There are numerous self-report measures to assess SAD symptoms (cf. Caballo et al., 2013). These questionnaires, which are used in research and clinical routine, represent assumed concepts of SAD. In Article II, the method of factor analyses was used to assess the underlying joint factor structure of two questionnaires and, with that, the assumed structure of SAD. Since the factor structure of SPS and SIAS were examined in Article II and were used as the main outcome for the OPTIMIZE study (Article III & IV) as well as their short form (i.e., SPS-6 & SIAS-6; Peters et al., 2012) in the embarrassment study (Article I), only these two questionnaires were further considered in this chapter. Additionally, out of five popular SAD self-report questionnaires, SPS and SIAS were considered differently from the other three in the sense that they measured more cognitive aspects of SAD, according to Caballo et al. (2013).

2.4.1. SPS & SIAS

SPS and SIAS (Mattick & Clarke, 1998; German version, which was used in Article IV: Stangier et al., 1999) are self-report questionnaires that are used to assess SA or as part of the diagnosis process of SAD. The questionnaires are often used together. SPS assesses SA in performance-related situations (e.g., walking down the street, holding a presentation), while SIAS assesses SA in interaction-related situations (e.g., talking with other people, talking about oneself). Each questionnaire consists of 20 items that can be rated from 0 ("not at all") to 4 ("extremely"). Internal consistency for both German versions was $\alpha = 0.94$ (Stangier et al., 1999).

SPS and SIAS are two of the most widely used questionnaires in the assessment of SA in clinical and non-clinical populations (Thompson et al., 2019). Many studies have examined the factor structure of the SPS and SIAS, either the factor structure for each questionnaire separately or of their joint factor structure with either clinical or non-clinical samples, resulting in a wide range of different models (see Wong et al., 2019, for an overview of previous studies which conducted factor analyses on SPS and SIAS). Article II examined the joint factor structure (see Chapter 3.2. for more information) and evaluated the one-, two-, and bifactor models. Depending on the model, there are different implications for the usage and interpretability of the SPS and SIAS and, with that, of the interpretation of SAD (if SPS and SIAS are considered as part of an SAD diagnosis) as a construct in general. The one-factor model proposes a combined usage of SPS and SIAS, where only their sum would be interpretable and would be an indicator for the global construct SAD. The two-factor model proposes a separate usage only, where each questionnaire would indicate a different kind of SAD (SPS proposes performance-related and SIAS interaction-related SA). The bifactor model is the combination of the two previous models. This model proposes a combined usage as well as the usage of the SPS and SIAS as separate valid questionnaires. It consists of one general factor, which explains a broad factor and shared variance, in this case, SAD. The corresponding group factors represent subdomain constructs that are conceptually specific and provide additional variance explanation (Rodriguez et al., 2016; Sun et al., 2022); in the case of the group factors SPS and SIAS, these would be performance-related SA and interaction-related SA, respectively. All factor analyses assume that the measured construct is dimensional. The one-factor model assumes that the construct is unidimensional, whereas the bifactor model assumes multidimensionality (Reise et al., 2013).

The bifactor model, in general, fits the growing popular perspective of SAD being a dimensional rather than categorical construct (Bornovalova et al., 2020; Gomez & Watson, 2017). Since the bifactor model assumes a multidimensional construct, where the group factors and the general factors are orthogonal, it is possible to determine how much common variance is explained besides the general factor. Additionally, with the help of factor analyses, it is possible to evaluate the reliability of SA(D) questionnaires and to draw conclusions about the SAD construct within the

dimensional perspective. What clinical consequences the research around the construct may yield will be discussed in Chapter 4.

2.5. Treatment of SAD

2.5.1. CBT and ICBT

CBT is regarded as a first-line treatment for SAD (National Institute for Health and Care Excellence, 2013) and was consistently shown in different meta-analyses to be an effective treatment for SAD (Acarturk et al., 2009; de Ponti et al., 2024; Hofmann & Smits, 2008; Kindred et al., 2022; Mayo-Wilson et al., 2014). An equally effective treatment is the ICBT. Many meta-analyses (Andrews et al., 2010, 2018; Carlbring et al., 2018) and RCTs (Andersson et al., 2012; Andrews et al., 2011; Berger et al., 2009, 2011; Boettcher et al., 2012; Clark et al., 2023) have shown that ICBT for SAD is effective and as effective as f2f treatments (Andersson et al., 2014). Additionally, Nordgreen et al. (2018) found their ICBT to be effective in their naturalistic within-group design.

Apart from the general advantages of ICBT in comparison to CBT (e.g., less costs, accessibility independent of residence, autonomous work on therapy content), there are some advantages particularly relevant for SAD (e.g., lower inhibition threshold, no social barriers, repetition of the therapy material) (Berger & Caspar, 2011; Guo et al., 2021). Nevertheless, only a fraction of the affected people seek help (Bruffaerts et al., 2022; Gross et al., 2005; Issakidis & Andrews, 2002), and if they do, there is still a significant amount of people who do not improve in a clinically significant way (Boettcher et al., 2013). One of the main goals of the OPTIMIZE study (Article III, IV) was, therefore, to determine the active components of the ICBT for SAD, in order to find a possibility more efficacious and briefer SAD treatment. The ICBT used in OPTIMIZE was based on the cognitive SAD model by Clark & Wells (1995), which will be presented in the next chapter, as well as on the books by Rapee (1998) and Stangier et al. (2003) (which were based on Clark & Wells (1995)).

2.5.2. The Clark & Wells (1995) Cognitive SAD Model

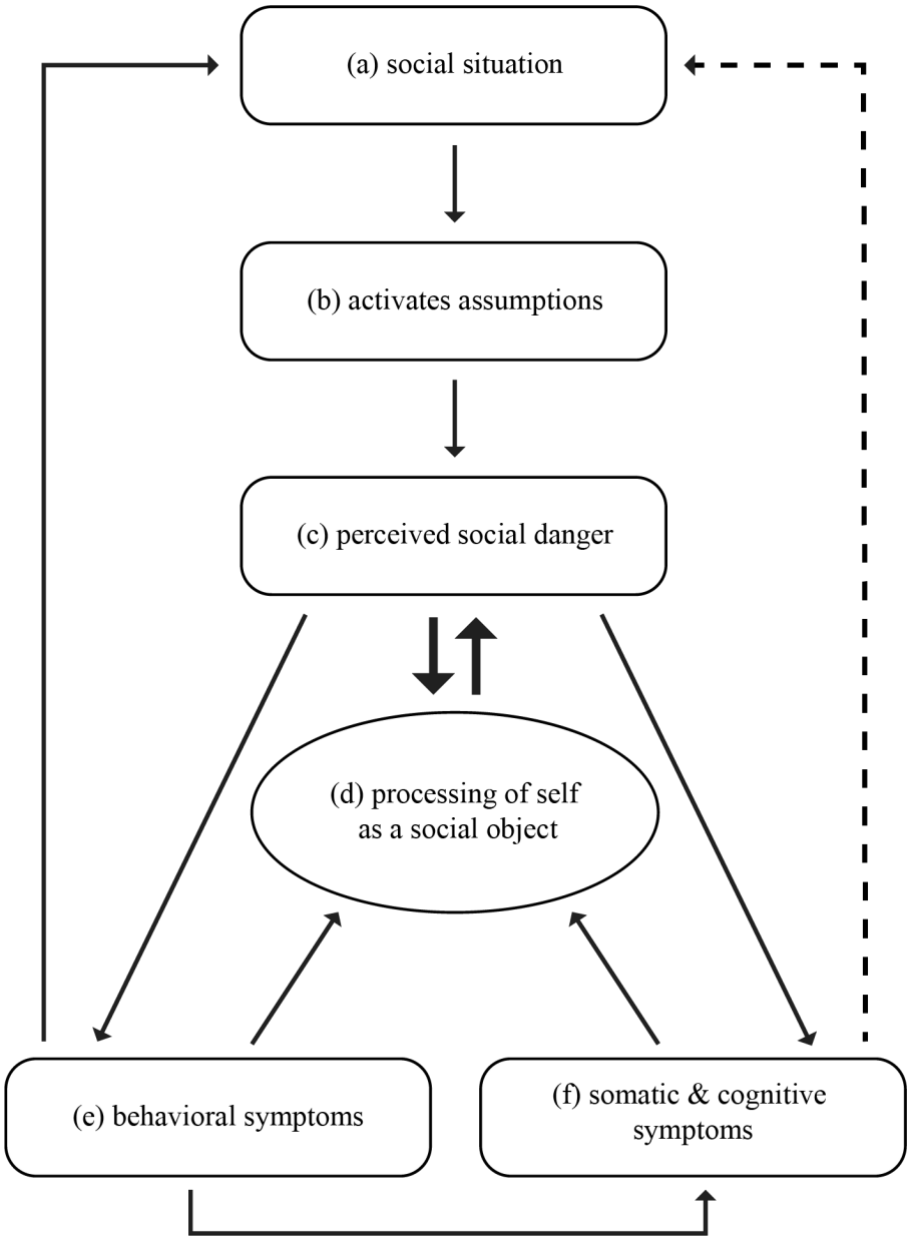
The cognitive SAD model by Clark and Wells (1995) is depicted in Figure 1. The authors describe in this model the internal and maintaining process of a person with SAD when entering a feared social situation: The (a) social situation is already biased and feared due to past (negative) experiences and the person's predispositions. This, on the other hand, (b) activates assumptions about themselves and others (e.g., I am going to be rejected by others, I am going to embarrass myself, etc.). According to Skocic et al. (2016), the negative pre-assumptions lead to two interpretation biases: cost-evaluation bias (i.e., catastrophic interpretations of social events) and likelihood evaluation bias (i.e., catastrophe is likely to occur). These two biases lead to the (c) perceived social danger, respectively,

cost likelihood evaluation bias, which suggests to people with SAD that a negative event is very likely to happen and will lead to tremendous negative consequences. This causes an attention shift to themselves (i.e., becoming self-conscious) and to their interoceptive information, leading to emotional reasoning (e.g., I feel anxious, therefore, others must notice that I am anxious). This is what the model means by (d) processing themselves as a social object: People with SAD imagine or feel how others must perceive them (e.g., a person with SAD is feeling hot and imagines that others see their presumably crimson face), basically constructing an image of how others must perceive them (negatively) and they do not perceive other external, opposing information. The process between (c) and (d) is highlighted with bold arrows, and it shows, according to Clark and Wells (1995), how the evidence for negative beliefs of people with SAD stems from their own impression rather than from actual observations around them. The (c) perception of danger leads to further (e) behavioral (i.e., safety and avoidance behaviors) as well as (f) somatic and cognitive symptoms, which cause different vicious cycles, eventually leading to the maintenance of SAD.

There are four (I)CBT components that are widely used in (I)CBT for SAD: psychoeducation, cognitive restructuring, attention training, and exposure. These were the four main components that were examined in OPTIMIZE. The study and the four main components will be described in the next chapter.

Figure 1

Cognitive SAD model by Clark & Wells (1995)



Note. Cognitive SAD model adapted from *A cognitive model of social phobia* by Clarke & Wells (1995) (p.72), in *Social Phobia: Diagnosis, Assessment, and Treatment* by Heimberg, Liebowitz, Hope, & Schneider (Eds.), Guilford Press.

2.5.3. OPTIMIZE

This chapter focuses solely on the clinical content of OPTIMIZE and the current corresponding literature. The ICBT program utilized in the OPTIMIZE study, Shyne, was based on the previously presented model by Clark and Well (1995), as well as on Rapee (1998) and Stangier et al. (2003). Due to the full factorial design of OPTIMIZE (for information on the design of the study, see Chapter 3.3.)., there were 16 possible conditions (four treatment components, respectively, main components with two possible levels (the component was either absent or present), resulting in a permutation of 16 versions). There was one wait-list control group (WL) and 15 active conditions with one to four main components, depending on the version (see Chapter 3.3., Table 1 for an overview of all possible conditions). Participants were randomized to one of the 16 conditions, where they used the Shyne version adjusted to their particular condition for eight weeks. During this time, they received guidance from trained master students. The study by Andersson et al. (2012) showed that the guides' expertise did not make a difference in the outcome. Guidance was provided to potentially enhance adherence and to facilitate the same dosage over all active conditions (i.e., all participants were told to work on the program for 50-60 minutes, independent of their condition). Before the reception of the program, after four weeks, and after eight weeks (time points: pre, mid, post), different data were ascertained from the participants. Four months after post, the last data were ascertained (time point: follow-up) (see Article IV for more information on the study procedure of OPTIMIZE).

Independent of the condition, all 15 versions of Shyne contained an introduction and a closing chapter. The goals of the introduction were, firstly, to introduce the participants to Shyne and to provide the necessary information for the upcoming eight weeks (e.g., show where the chat function is and how guidance works). Secondly, the motivation (and indirectly the adherence) was expected to be increased by asking the participants to write about the current constraints caused by their SAD and how their lives might look without the disorder. Additionally, a short chapter was provided about the effectiveness of ICBT for SAD.

The goal of the closing chapter was to provide relapse prevention. For that purpose, for every version of Shyne, an individualized closing chapter was provided based on the main components that the version contained, with tips and reminders on how to continue with the treatment (e.g., continue exposure while avoiding safety and avoidance behavior). In the end, participants were asked to write about the following two topics: 1.) Which are the most important things that you learned and that you want to keep in mind? and 2.) Which techniques/exercises/thoughts, etc., have helped you the most in the event of relapses or slow progress? How do you want to proceed in the event of relapses? What is important to you to do/to think, etc., in such moments?

The following subchapter (i.e., 2.5.4.1. – 2.5.4.4.) will shortly present the content of each main component, followed by the current research on its effectiveness. It is important to present the

content of each component first because, even though the components are common, it does not mean that they contain the same information or techniques across different studies.

2.5.4. (I) CBT Treatment Components in OPTIMIZE

2.5.4.1. Psychoeducation. This component was realized as a text-based chapter called "Understanding Anxiety and Anxiety Disorder" and one protocol (see Article IV, Supplementary Material for the protocol). The first few subchapters provided information on SAD (e.g., potential causes, prevalences, maintenance processes, etc.). The maintenance processes of SAD were explained based on the Clark and Wells cognitive SAD model (1995; see Chapter 2.5.2.). Based on the learned information, participants were then asked to provide their own SAD model and with that, boosting memory and tailoring the information to the individual participant. In the end, short information was provided on the treatment of SAD. At the end of each subchapter were comprehension questions to facilitate information consolidation and provide interactive elements. The corresponding protocol was a template for the participants, where they were instructed to systematically describe and rate their personally feared situations. This was the only protocol that all 15 versions of Shyne contained. Even though it belonged thematically to psychoeducation, it was also a basis for all versions since participants needed to know their fears in order to work with them.

There is not much research that tests psychoeducation as an intervention component for SAD specifically. Nordmo et al. (2015) tested whether adding a 90-minute f2f psychoeducation session before starting an ICBT for SAD would add an additional value. They showed that there were no significant differences concerning the SAD symptom outcome, treatment satisfaction, or the amount of program completion. Dijk et al. (2012) tested a psychoeducation course on the fear of blushing, where they used the SIAS as one of the outcome measures. They found significant SAD symptom reduction, as well as a reduction in the fear of negative evaluation and social avoidance. As the symptom reduction grew significantly over the time points (comparing pre to six weeks (i.e., after the course), three months, and one year later), they concluded that the effect of psychoeducation might take time, since participants need to practice with the learned tools.

2.5.4.2. Cognitive Restructuring. This component was realized as one text-based chapter called "Thinking realistically" (including the basic protocol from psychoeducation) and one protocol. This chapter focused on dysfunctional and negative cognitions. In the beginning, participants received information on automatic cognitions and typical faulty reasoning in SAD (e.g., jumping to conclusions in ambivalent social situations, catastrophizing, etc.). They then received support on how to identify and question automatic negative and dysfunctional thoughts (e.g., What would you tell a friend in this situation?). With the help of the corresponding protocol, the participants could put the

learned cognitive restructuring principles into practice by using thoughts of situations from their own daily lives.

The research on cognitive restructuring in SAD is rather meager. However, there is literature on the cognitive bias modification (CBM). A meta-analysis from Liu et al. (2017) showed small and significant effects for the CBM. Interestingly, compared to the attention bias modification (ABM), the CBM was superior in regard to the reduction of SAD symptoms. However, they compared different techniques, like the dot-probe task with cognitive restructuring training (which is closer to our component, e.g., working on interpretation bias within ambiguous social situations). Trainings like the CBM target rather implicit biases in comparison to the OPTIMIZE cognitive restructuring. The following two studies used cognitive restructuring training close to the one of OPTIMIZE: Mattick et al. (1989) examined the effect of cognitive restructuring, exposure, and their combination in a group therapy setting. The combination, as well as cognitive restructuring alone, showed superiority when looking at all measures. However, in comparison to exposure, the positive effects of cognitive restructuring were further apparent from post to follow-up. Cougle et al. (2020) examined the interpretation bias modification with progressive muscle relaxation. Interestingly, both techniques showed significant SAD symptom reduction for post and follow-up, but there was no significant between-group effect.

2.5.4.3. Attention Training. This component was realized as one text-, audio-, and video-based chapter called “Attention Training” (including the basic protocol from psychoeducation) and exercise sections. In the first two subchapters, participants received information on attention bias, the ways of directing attention (either inward or outward), and how self-focused attention (inward) is an important part of SAD maintenance. For practice, three different exercises were provided. In the first audio exercise, participants were instructed to listen to an audio file and, in each segment, to focus on only one thing (e.g., birdsong) while ignoring the rest (e.g., indistinguishable chatter, footsteps, etc.). The goal was to train conscious attention control. In the second video exercise, participants were presented with a short story. They were instructed to read through a short story within a minute at most and then tell the story in front of a recorded audience. The audience could be either a woman, a man, or a group of people (i.e., a video was presented with a person/a number of people listening with changing expressions). At the very beginning, there were two rounds: In the first round, participants were instructed to focus solely on themselves. In the second round, they were instructed to focus on the task. After each round, participants were instructed to rate their anxiety and performance and how much focus there was on themselves. After that, they could practice the attention focus away from them with different stories and audiences and could track their progress and ratings graphically. This exercise was the biggest exercise and most important of all three. The third exercise was a short text on how to practice mindfulness in feared situations (i.e., shift attention away from feared cues in the situation and focus on the environment).

As with cognitive restructuring, the research on attention training is meager. Heeren et al., (2015) found in their meta-analysis a small effect for ABM, but they also mentioned that online training may be less effective than in the laboratory because the same level of anxiety might not be reached in the online training. Liu et al. (2017) found in their meta-analysis that the effects from the laboratory were significantly bigger than those from online-training. While Amir et al. (2009) found a significant reduction of SAD symptoms with their ABM treatment, Boettcher et al. (2014) on the other hand, did not find advantageous outcomes by adding ABM in comparison to ICBT alone. Interestingly, in both studies, the cues in the ABM tasks showed the emotion of disgust. Schmidt et al. (2009) used ABM as well, but due to vague conclusions from the results, their study was not considered here.

Attention training and cognitive restructuring are special cases within the four treatment components, as very different cognitive restructuring and attention training techniques are conducted across literature in comparison to the other two components. For example, people within the component exposure expose themselves to their feared situation, while attention training can be a modified dot-probe task (Amir et al., 2009; Boettcher et al., 2014) or, in the case of OPTIMIZE, exercising telling stories in front of an audience while practicing attention shift. Consequently, Amir et al. (2009) and Boettcher et al. (2014) reported their results as a change in response time to specific cues, while the OPTIMIZE study used the Self-Consciousness Scale (SCS; Lopes et al., 2021) as the corresponding outcome measure. Therefore, even though there is literature on the effect of attention and cognitive training, there is almost no research on the particular techniques used in OPTIMIZE.

2.5.4.4. Exposure. This component was realized as a text-based chapter called “Testing the Reality” (including the basic protocol from psychoeducation) and one protocol (see Appendix D, Supplementary Material A for the protocol). In the beginning, participants were informed about different processes before, during, and after exposure (e.g., pre-and post-processing, habituation, avoidance, and safety behavior) and why exposure was important for the treatment of SAD. The protocol from psychoeducation was the basis that they could use to develop a personal anxiety hierarchy. With the exposure protocol, participants could, on one hand, plan exposures (i.e., document expected anxiety levels, physical symptoms, and which safety behavior the person wants to avoid, and other variables). On the other hand, the protocol had the function of a diary, where they could document the important aspects after the exposure (i.e., the actual anxiety level, physical symptoms, etc.) in order to avoid the typical SAD post-processing, respectively, post-mortem after exposure.

Exposure techniques are one of the most studied and widely implemented components in CBT and are considered a first-line treatment for numerous anxiety disorders (Kaczkurkin & Foa, 2015). Hofmann (2004) compared exposure group therapy to CBT and did not find any significant differences at post, implying that the component exposure might be enough to reduce SAD symptoms effectively. The same results have been found by multiple meta-analyses (Kaczkurkin & Foa, 2015;

Powers et al., 2008). Mattick et al. (1989) showed that the combination of exposure and cognitive restructuring was superior to exposure alone when measured with SA(D) scales. But as it was mentioned in the second to last subchapter, when compared to cognitive restructuring, this component continued to improve from post to follow-up in comparison to exposure, while exposure alone even deteriorated. Gil et al. (2001) showed that exposure was not more effective when combined with social skills training or cognitive restructuring. One explanation is that there is a common therapeutic element shared by these techniques. For example, social skills training contains confrontation with the feared situation (i.e., exposure in vivo). This Chapter 2.6.4. shows, in general, how necessary it is to try to compare pure components in order to truly understand if and how the components actually work.

3. Key Results of the Dissertation Articles

This chapter provides the key results of the Articles I – IV. Article I focused mainly on embarrassment and how it can be modeled by using a mixture of subjective and objective measures. Article II focused on the measurement and dimensional conceptualization of SAD by examining the joint factor structure of the questionnaires SPS and SIAS. Articles III and IV thematized the treatment of SAD with ICBT within the OPTIMIZE study. Article III is the study protocol and will be used to explain the factorial design used in OPTIMIZE. Article IV is the main outcome paper of OPTIMIZE and will be used to present the most important results from the study. The contextualization and embedding of the results from Articles I – IV will be discussed in Chapter 4.

3.1. Article I – Embarrassment

- **Title:** Multidisciplinary Characterization of Embarrassment through Behavioral and Acoustic Modeling
- **Authors:** Dajana Šipka (first author), Bogdan Vlasenko, Maria Stein, Thomas Dierks, Mathew Magimai-Doss, Yosuke Morishima.
- **Journal (re-submission date):** Nature Scientific Reports (21.01.2025)

In this article, embarrassment was described and examined from different points of view. The four goals of this article were (1) to induce embarrassment in participants, test the success of the induction, and examine how embarrassment was related to SA, (2) to test the prediction performance of a trained model in the sample data in pre-induction, embarrassment, and post-induction and to show the robustness of the embarrassment dataset, (3) to map embarrassment on the three-dimensional emotion space with the axes valence, arousal, and dominance (VAD) in a dimensional approach, and finally (4) to compare embarrassment to other emotions in a categorical approach.

The sample consisted of N = 33 undergraduate psychology students. Concerning the study procedure, before the day of the assessment, participants filled out online questionnaires consisting of demographic questions and the questionnaires Short Form Social Phobia Scale (SPS-6) and Short Form Social Interaction Anxiety Scale (SIAS-6) (Peters et al., 2012). On the day of the assessment, participants had to write about an embarrassing experience without knowing that they would later need to read it out loud to the conductor (i.e., embarrassment induction). Before and after this embarrassment induction, they had to tell stories based on neutral pictures. During the whole assessment, participants' voices were recorded. The participants marked how embarrassed they felt at the moment on a Visual Analog Scale (VAS) before writing the embarrassing story (VAS 1), after writing (VAS 2), and after telling the embarrassing story (VAS 3).

Concerning the methods, a mixture of subjective measures, objective engineering and machine learning approaches was chosen. The subjective measures consisted of the VAS as well as the SPS-6 and SIAS-6. The objective approaches consisted of machine learning models. For goal 2, a classification model was trained based on our sample with the leave-one-speaker-out approach. For goals 3 and 4, the models were based on publicly available emotional speech corpora (i.e., acoustic samples with emotion labels) with a cross-corpora approach.

The following key results were found:

- Verify embarrassment induction (goal 1): Embarrassment induction was successful. There was a significant difference between VAS 2 and VAS 3 ($V = 551$, $p < .001$), meaning that the participants got significantly more embarrassed after reading the story out loud in comparison to just writing it down. Additionally, there was a significant negative correlation between the total sum of SPS-6 and SIAS-6 and the difference score between VAS 2 and VAS 3 ($rs = – 0.36$, $p = .037$), meaning that the higher the prior SA was, the more embarrassed participants became after reading the story.
- Prediction performance (goal 2): The best prediction performance was found for pre vs. embarrassment with support vector machine, where in 86.4% of the cases, the right prediction was made. The predictions for post vs. embarrassment (81.8%) and for pre + post vs. embarrassment (81.8%) were satisfactory as well. No satisfactory discrimination was possible between pre vs. post (59.1%, slightly above chance), indicating these two states to be similar.
- Dimensional modeling of embarrassment (goal 3): The density map showed that the dimensions of arousal and dominance showed some predictive value, while valence did not have any at all. Yet, none of those differentiated enough between pre vs. embarrassment and post vs. embarrassment.
- Categorical modeling of embarrassment (goal 4): Mean posterior values showed significant differences between pre vs. embarrassment and post vs. embarrassment for the emotions sadness, boredom, and neutral. In pre vs. embarrassment, participants' voices shifted significantly to more boredom ($p < .001$), more neutral state ($p < .001$), and less sadness ($p < .001$) related characteristics. The opposite for post vs. embarrassment was found, specifically, the voices shifting back to less boredom ($p = .001$), less neutral state ($p < .001$), and more sadness ($p < .001$) related characteristics. As was shown in the prediction performance already, there were no significant differences for pre vs. post for any of the seven possible emotions (i.e., happiness, sadness, disgust, fear, boredom, anger, and neutral state).

Concerning the basic research on embarrassment, the results based on the subjective approach support the assumption that the feeling of embarrassment requires an audience. Participants only got significantly more embarrassed after reading the story to the conductor. Furthermore, it shows once again the link between embarrassment and SA, where higher SA scores correlated with higher embarrassment scores. The results based on the machine learning approach showed that the VAD dimensions could not model embarrassment. The categorical approach, on the other hand, found a shift toward more neutral, more bored, and less sad voice characteristics when embarrassed.

3.2. Article II – Factor Analyses of SPS & SIAS

- **Title:** Factor structure of the Social Phobia Scale (SPS) and the Social Interaction Anxiety Scale (SIAS) in a clinical sample recruited from the community
- **Authors:** Dajana Šipka (first author), Jeannette Brodbeck, Ava Schulz, Timo Stolz, Thomas Berger
- **Journal (doi):** BMC Psychiatry (https://doi.org/10.1186/s12888-023-05142-8)

In this article, the joint factor structure (i.e., factor analyses were based on all 40 items of the SPS and SIAS) of the questionnaires SPS and SIAS was examined. The two goals of this article were (1) to evaluate the underlying joint factor structure of the SPS and SIAS and (2) to test whether SPS and SIAS are reliable scales to assess two different aspects of SA(D). The study sample consisted of N = 298 people with a diagnosed SAD from the community. Confirmatory factor analyses (CFA) and exploratory factor analyses (EFA) were conducted. The CFA tested a one-factor, two-factor, and bifactor model. The one-factor model assumed that all 40 items would load on a general factor (GF), representing SA(D). The two-factor model assumed that the 20 SPS items would load on an SPS factor and the 20 SIAS items on an SIAS factor, meaning that SA(D) would be represented through social performance (SPS) and interaction anxiety (SIAS). The bifactor model was a combination of the two previous models and, therefore, assumed that SA(D) consisted of a GF as well as performance and interaction anxiety, assuming that SA(D) would be a multidimensional construct. For each model, the errors of the three reversed SIAS items were correlated.

The following key results were found:

- The one-factor and two-factor models were rejected. The one-factor model did not show any acceptable indices, and the two-factor model showed only an acceptable RMSEA (0.073, 90% Cl [0.069, 0.077]).
- The bifactor model showed only two acceptable indices, RMSEA (0.059, 90% Cl [0.054, 0.063]) and SRMR (0.067). Therefore, a refined bifactor model without the seven non-significant items and one error correlation was calculated. Due to a model suggestion, an additional error correlation between two SIAS items was calculated.

- The refined bifactor model showed not good but acceptable indices (RMSEA: 0.052, 90% Cl [0.047, 0.056]; SRMR: 0.067; CFI: 0.917, TLI: 0.908). Therefore, an item-driven EFA was conducted, but it did not show any new models with a better fit.
- The bifactor-specific indices for the refined bifactor model showed, on one hand, that the majority of the reliable variance was attributed to GF (74%), and half of the common variance was explained by GF (ECV = .51). On the other hand, SPS and SIAS were shown to be reliable subscales (SPS ωs = .91, SIAS ωs = .94) and explained equally the remaining common variance (SPS ωhs = .52, SIAS ωhs = .47). The remaining indices (i.e., PUC, FD, H) have further indicated that the refined bifactor model is tendentially multidimensional and that SPS and SIAS measure different aspects of SA(D).
- The content analysis showed, among others, that social performance anxiety was less well represented by the refined bifactor model (six SPS items were not significant) compared to social interaction anxiety (two SIAS items were not significant).

In conclusion, Article II has shown that the underlying joint factor structure of SPS and SIAS was best represented by a refined bifactor model and that his model was tendentially multidimensional. This means that SPS and SIAS assess specific aspects of SA(D) beyond the general SA(D). In practice, these results imply that SPS and SIAS could be reliably used in combination as well as separately. Nevertheless, not all items in the model were significant and especially SIAS seemed to be more fitting to the refined model than SPS.

3.3. Article III – OPTIMIZE - Study protocol

- **Title:** Optimizing cognitive-behavioral therapy for social anxiety disorder and understanding the mechanisms of change: Study protocol for a randomized factorial trial
- **Authors:** Rodrigo C.T. Lopes (first author), Dajana Šipka, Tobias Krieger, Jan Philipp Klein, Thomas Berger
- **Journal (doi):** Internet Interventions (https://doi.org/10.1016/j.invent.2021.100480)

Since Article III is the study protocol of the OPTIMIZE study, this chapter will be used to explain the design of the study instead of showing any key findings. As was already mentioned in the introduction, up to this day, it is known that ICBT for SAD works, but it is not clear how it works. Therefore, one of the main goals of OPTIMIZE was to find the active components of ICBT for SAD. For this, the full factorial design was chosen. This decision was made because an RCT is not an efficient method to examine the performance of individual components, nor is it able to examine how the components influence each other, depending on their absence or presence (Collins, 2018). In other words, with a factorial design, the main effect of each component and the interaction effects of two or

more components can be examined. Table 1 shows the factorial design of the OPTIMIZE study. In the case of OPTIMIZE, there were four treatment components (i.e., psychoeducation, cognitive restructuring, attention training, and exposure) that could either be absent or present, leading to a 2x2x2x2 design with 16 possible conditions. Per condition, n = 29 participants were assigned so that a total of N = 464 participants were included in the OPTIMIZE study (for the exact calculation of the sample size, see Article III, Appendix C).

Another advantage of a factorial experiment in comparison to an RCT is that a factorial experiment is a more economical design and more powerful method. This can be shown with one example from our study: In Table 1, condition 10 comprises psychoeducation and exposure. This means that from the perspective of psychoeducation and exposure, a participant from condition 10 is in an active condition (these two components are present), while from the perspective of cognitive restructuring and attention training (these two components are absent), the same participant is in a control condition. This example first shows that depending on which main component is examined, the same participant can be either in an active or control condition. Secondly, it shows how fewer participants are needed than in an RCT since the four main components are not four separate arms, and the power is not calculated per *n* of a condition but per *n* of a factor level (see for more information Collins, 2018). The next chapter will present the statistical methods of OPTIMIZE and the key findings.

Table 1

Full factorial design of the OPTIMIZE study

condition	psychoeducation	cognitive restructuring	attention training	exposure
1 (WL)				
2				x
3			x	
4			x	x
5		x		
6		x		x
7		x	x	
8		x	x	x
9	x			
10	x			x
11	x		x	
12	x		x	x
13	x	x		
14	x	x		x
15	x	x	x	
16 (FV)	x	x	x	x

Note. x = the component is present; blank space = the component is absent. WL = wait-list control group. FV = full version.

This table is intentionally not according to the APA 7 guidelines for better readability.

3.4. Article IV – OPTIMIZE - Main Outcome Paper

- **Title:** Active Components in Internet-Based Cognitive Behavioral Therapy for Social Anxiety Disorder: A Randomized Full Factorial Trial.
- **Authors:** Dajana Šipka (first author), Rodrigo Lopes, Tobias Krieger, Jan Philipp Klein, Thomas Berger
- **Journal (doi):** Psychotherapy and Psychosomatics (https://doi.org/10.1159/000542425)

The goal of Article IV was to examine the main effects of the four main treatment components in ICBT for SAD. For the statistical analysis of the main outcomes, linear mixed models (LMMs) were calculated. Time (i.e., pre, mid, post, follow-up) was the within-subject factor, and the four main components (either present (1) or absent (-1)) were the between-subject factors. The LMMs were built using a content-driven approach, where the fixed effects were time points, the main component, and their interaction term. The random effect was the intercept of each participant. For the main results, the intention-to-treat sample was used. The main outcome measure was the composite score (CS) of the SPS and SIAS.

The following key results were found:

- Interaction effect between time (i.e., pre, mid, post) and main component measured with the CS: A significantly higher reduction in SAD symptoms was found for the versions containing psychoeducation in comparison to versions without psychoeducation (F(2, 690.39) = 12.56, p < .001; Cohen's d = 0.39, 95% CI [0.21, 0.58]) and for versions containing exposure in comparison to versions without exposure (F(2, 691.50) = 9.60, p < .001; Cohen's d = 0.41, 95% CI [0.22, 0.59]).
- Between-group comparison from pre to post on condition level measured with the CS: All active conditions (i.e., 2-16), independent of the component's presence or absence, reduced the SAD symptoms significantly in comparison to the WL (Fs (2, $df(res)_{range}$ [85.33, 102.71]) ≤ 36.99, ps < .001); Cohen's d range: from d = 0.67, 95% Cl [0.07, 1.26] to d = 1.56, 95% Cl [0.87, 2.25]).
- Secondary measures from pre to post: Participants with the versions containing psychoeducation were significantly more satisfied with Shyne (t(272) = -3.40, p < .001), used the program significantly longer (t(426) = -4.78, p < .001; d = 0.46, 95% CI [0.27, 0.66]), and had significantly fewer SAD diagnosis (χ^2(1, n = 250) = 10.03, p = .002, OR = 1.08) in comparison to the versions without psychoeducation. Independent of the versions, participants were significantly less depressed and generally anxious and showed a higher mental quality of life (QoL) from pre to post (see Article IV).
- Time (i.e., post, follow-up) effect for main component measured with the CS: The versions without psychoeducation and without exposure caught up to the versions with the

respective component so that no significant difference was found anymore (post hoc tests with all time points (i.e., pre, mid, post, follow-up) included: psychoeducation: $t(916.57) = -1.36$, $p = .175$; exposure: $t(917.62) = -1.74$, $p = .082$). There was a significant time effect for the versions with cognitive restructuring ($F(1, 105.26) = 10.35$, $p = .002$) and attention training ($F(1, 95.68) = 9.05$, $p = .003$) in comparison to those without the respective component. However, there was no significant interaction effect.

- Dropouts at post: Participants with versions containing exposure were more likely to drop out in comparison to those without exposure ($\chi^2(1, N = 464) = 5.49$, $p = .019$). Measured with the CS, participants with higher SAD symptom scores were more likely to drop out ($t(462) = -1.98$, $p = 0.049$).
- Dropouts at follow-up: Participants without the component psychoeducation were more likely to drop out in comparison to those with psychoeducation ($\chi^2(1, N = 464) = 7.37$, $p = .007$).

In conclusion, from pre to post, the versions with the component psychoeducation and exposure were superior in terms of SAD symptom reduction in comparison to the versions without the respective component. Furthermore, on the condition level, it was shown that all active conditions, independent of a main component, contributed to a significant reduction of SAD symptoms. The secondary outcomes from pre to post further highlighted the superiority of psychoeducation, namely the higher satisfaction with the program, more time spent on the program, and the lower number of SAD diagnoses. From post to follow-up, the versions without psychoeducation and exposure caught up so that there was no significant difference between the versions with and without the respective component. The versions with cognitive restructuring and attention training showed a significant SAD symptom reduction but without a significant interaction effect between time and the component.

4. Discussion

The aim of this dissertation was to provide a multidimensional perspective on SAD and its treatment through ICBT. Chapter 4.1. will discuss the multidimensional perspective, while Chapter 4.2. will discuss the treatment. Chapter 4.3. will present the limitations of the dissertation. At the end, Chapter 4.4. will reveal future implications.

4.1. The Multidimensional Perspective on SAD

SAD can be observed from different perspectives. Firstly, SAD can be observed from a non-clinical, emotion-focused perspective through the lens of embarrassment and SA. As the theoretical background showed, embarrassment, SA, and SAD are strongly intertwined. One of the core aspects that all these concepts share is the fear of being negatively evaluated and the feeling of being the center of attention (especially after a social norm-related violation). Some studies, as well as Article I, have shown that people with higher SA get more easily embarrassed. Furthermore, people who get more easily embarrassed are more sensitive to social norms. More concretely, Čolić et al. (2020) showed how embarrassment among people with SAD possibly affected their perception of ambiguous social situations and post-processing. Bas-Hoogendam et al. (2018) explained the connection between SA(D) and embarrassment with the difficulty for people with high SA to differentiate emotionally between intentional and unintentional violations and how repeated and intensive embarrassment experiences led to negative self-evaluation and dysfunctional cognition, among others. Article I characterized embarrassment through voice analysis and, using the categorical approach, found that when participants got embarrassed, their voices shifted to less bored, less sad, and more neutral voice characteristics. One interpretation was that the articulation of the voice changes with the feeling of embarrassment, and embarrassed people might, for example, try to express themselves more clearly. It would go along with the literature, which has shown how embarrassed people try to rectify the situation. Another interpretation was that embarrassed people actually distance themselves emotionally from the situation and that this would be reflected in the voice. Both interpretations would need to be tested in future studies. The dimensional approach could not characterize embarrassment as satisfactory on the VAD dimensions. One hypothesis was that these dimensions were too simplistic and that embarrassment as an emotion is too complex to be captured by these three dimensions, as was also shown in Chapter 2.1.1. In conclusion, it was shown that embarrassment and SA(D) are intertwined and interact with each other in different ways. It is, therefore, indicated to pay more attention to the emotional perspective of SAD in research as well as clinical practice.

Secondly, SAD can be observed from a structural perspective concerning its assumed underlying factor structure, as was done in Article II. While the DSM-5 (APA, 2013a) defines SAD as a categorical construct (i.e., one either has an SAD diagnosis or not), a more dimensional perspective

(i.e., no threshold between a healthy state and a disorder) seems to appear. Although the DSM-5 has added a continuum-based assessment to some disorders (e.g., substance use disorder; APA, 2013b), SAD is still a categorical diagnosis but, it now includes the possibility for a specification of performance-only SAD. Article II examined the underlying joint factor structure of the SPS and SIAS; by using factor analysis, a dimensional construct was assumed. The results showed a refined bifactor model, with the bifactor-specific indices showing a multidimensional construct. Nevertheless, the refined bifactor model showed only an acceptable fit and was reduced by a number of items that did not fit the model. With the appearance of the DSM-5, a generally more dimensional instead of a categorical view of psychopathology was encouraged (LeBeau et al., 2015). According to Rice et al. (2021), most SAD questionnaires are limited to a range of symptoms and are not up-to-date with the new DSM-5 criteria. This is particularly evident with the SPS and SIAS. As was mentioned in the theoretical background of this dissertation, the SPS and SIAS are considered to take a predominantly cognitive view of SAD, and they were created more than 20 years ago. In the literature, there is no consent for the underlying joint factor structure. Concerning the content, the SIAS item 14 ("I have difficulty talking to attractive persons of the opposite sex") has already been avoided by researchers more than 10 years ago due to its political incorrectness, according to Caballo et al. (2013). The three reverse-scored SIAS items were found to not fit into the factor models numerous times (Eidecker et al., 2010; Gomez & Watson, 2017; Rodebaugh et al., 2006, 2007). An example of a dimensional addition to research as well as clinical practice assessments is the Social Anxiety Disorder Dimensional Scale (SAD-D; Craske et al., 2013) from the APA. The SAD-D assesses SA(D) as a dimension based on symptom severity and has already been assessed in numerous studies (e.g., Knappe et al., 2014; LeBeau et al., 2012; Rice et al., 2021). In conclusion, regarding the presented results this dissertation, it would be indicated to at least add a dimensional assessment, like the SAD-D, to the classical self-report questionnaires like the SPS and the SIAS.

4.2. The Treatment of SAD

In the main project of this dissertation (OPTIMIZE) the main effects of the four treatment components (i.e., psychoeducation, cognitive restructuring, attention training, and exposure) were examined. Concerning psychoeducation, Nordmo et al. (2015) did not find any advantages by adding psychoeducation. However, they only did one additional session before starting ICBT. In contrast, psychoeducation was one out of four treatment components in OPTIMIZE, and the versions with psychoeducation showed a significant reduction in SAD symptoms and fewer SAD diagnoses in comparison to the versions without the component from pre to post. Furthermore, participants were more satisfied with the program and spent more time on the program when receiving versions with psychoeducation in comparison to those without psychoeducation. So, it may be that more time is needed than just one session. Another explanation could be that the content of Nordmo et al. (2015)

did not cover the topics important to the participants. In comparison to the study of Dijk et al. (2009), there was no significant improvement from post to follow-up for the versions with psychoeducation in comparison to those without in Article IV. However, the OPTIMIZE CS consisted of the SPS and SIAS combined. It was not tested whether SPS or SIAS alone yielded different results, respectively, if participants improved differently on the separate questionnaires since Dijk et al. (2012) used the SIAS only.

Concerning cognitive restructuring, there are mixed results. Measured with the Behavioral Approach Test, Mattick et al. (1989) found exposure to be superior to cognitive restructuring at post, but cognitive restructuring caught up at follow-up, and exposure even deteriorated. They hypothesized that cognitive restructuring may take longer to affect behavioral measures. The same delay for the effect of cognitive restructuring was found in Article IV (measured with the CS), but there was still no significant difference between the versions with cognitive restructuring in comparison to those without. However, it was not clear from the Mattick et al. (1989) study how much time had passed between post and follow-up. On the other hand, Cougle et al. (2020) found significant large effects for SAD symptom reduction at post after only a 4-week training session. In comparison to Mattick et al. (1989) and Article IV, Cougle et al. (2020) provided sentences for the cognitive restructuring instead of participants having to come up with examples on their own. Therefore, one explanation might be that it is more difficult for participants to come up with their own anxiety-relevant cognitions and that they need more time. Concerning attention training, there was no previous literature on the particular technique used in Article IV. Therefore, it is not possible to evaluate how the results from Article IV can be interpreted within the literature. However, as with cognitive restructuring, attention training showed a delayed effect from post to follow-up, where the version with attention training significantly reduced SAD symptoms, but there was also no significant interaction effect between the component and time. Therefore, in research, it may be important to leave enough time between the post and follow-up measures in order to find effects. In practice, it may be important to stick to these techniques long enough. However, clearly, more research is needed, especially for cognitive restructuring and attention training in the way these components were implemented in Article IV.

Concerning exposure, most of the literature suggested that exposure alone might be enough to reduce SAD symptoms; also, additional CBT techniques did not show any significant benefits concerning SAD symptom reduction (Hofmann, 2004; Kaczkurkin & Foa, 2015; Powers et al., 2008), except for the study by Mattick et al. (1989). In Article IV, as with psychoeducation, the versions with exposure showed significantly higher SAD symptom reduction in comparison to the versions without exposure from pre to post. Additionally, exposure showed the only significant interaction effect between a secondary outcome and the CS from pre to post, where the versions with exposure had a significantly lower depression score than the versions without exposure.

Looking at the condition level from pre to post, it was shown that all active conditions were significantly better than the WL in terms of SAD symptom reduction. There was no significant SAD

symptom reduction for WL. Additionally, the effect of the number of components from pre to post was examined exploratorily. Again, there was a significant difference between the WL and one, two, three, and four components. However, there was no significant difference between one to four components, implying that it is better to provide affected people with anything from the SAD treatment, but it does not seem to matter how much. Concerning the secondary outcomes from pre to post, there was a significantly lower number of SAD diagnoses for the versions with psychoeducation in comparison to those without, further underpinning the superiority of psychoeducation. Independently of the component, there was a significant reduction from pre to post for depressive symptoms and general anxiety symptoms and an increase in mental QoL for the versions that contained a component in comparison to those that did not. These results revealed that although psychoeducation and exposure showed significantly higher SAD symptom reduction from pre to post, the other two components showed improvement in secondary outcomes as well. The results from the condition level, the number of components, and the significant time effects from post to follow-up for cognitive restructuring and attention training showed that the inclusion of each component leads to improvement on some level.

The dropout analysis at post showed a higher drop-out rate for the versions with exposure. One hypothesis in Article IV was that participants with exposure might need more support in order to expose themselves to their feared situations than participants with other components. Additionally, the dropout rate was higher for participants with more severe SAD symptoms. This might imply that participants with higher symptom severity would need more support as well in order to continue using the ICBT. At follow-up, participants with the component psychoeducation were more likely to drop out. The hypothesis in Article IV was that participants might have been satisfied enough and, therefore, dropped out due to the positive outcomes at post (Lawler et al., 2021).

In conclusion, psychoeducation and exposure were superior to cognitive restructuring and attention training in the sense that the versions with psychoeducation and exposure reduced the SAD symptoms more quickly and that there were additional significant effects for psychoeducation in particular. Article IV examined the main effects of the four components, one of the next steps would be to test the interaction effects. At the moment, we know that there is a significant effect for the versions with psychoeducation and exposure in comparison to the versions without the respective component. However, it is not clear how the effects behave if, for example, the interaction of psychoeducation and exposure is compared to the versions with a single component. After that, the components psychoeducation, exposure, and their combination should be tested in an RCT, as OPTIMIZE chose an exploratory approach to select potential components to optimize ICBT for SAD (cf. Collins, 2018). Another study within OPTIMIZE is planned, where the mediation and moderation effects of the main components are tested. This is a further step towards understanding how ICBT actually works and what the potential mechanisms of the changes are (Kazdin, 2007).

4.3. Limitations

This chapter will first present the overall limitations of this dissertation. Secondly, it will present the limitations of Articles I, II, and IV that are relevant to the overall topic of this dissertation. For detailed limitations on each specific Article, consult the limitations section of the respective Article.

This dissertation took the social evaluation model as a framework for the definition of embarrassment, where a social transgression evokes embarrassment. While the fear of negative evaluation by showing anxiety symptoms or behaving in a certain way is part of the SAD diagnosis (DSM-5 TR, diagnostic criterion B), the mere possibility of scrutiny in social situations should be enough to evoke anxiety (DSM-5 TR, diagnostic criterion A) (APA, 2022). This is also shown with the items of the SPS and SIAS, where writing in front of others or becoming aware of one's voice might evoke SA. As Clark & Wells (1995) described in their model that the mere anticipation of possible negative events or consequences leads to SA and further vicious cycles. In contrast, most embarrassment and SA(D) studies do actually induce embarrassment/SA or have tasks that would evoke embarrassment/SA in many people (e.g., in the study of Hofmann et al. (2006), participants had to sing; in the study of Laukka et al. (2008), participants did a public speaking task). There, the feared social transgression often actually happens. However, there are not enough studies that examine people with high SA or SAD and embarrassment in situations where there is no mishap or social norm transgression, where people are, for example, observed when they write or when they are simply the center of attention, without having, for example, to hold a presentation or perform. Therefore, this dissertation has shown how embarrassment is connected to SA and SAD in embarrassing or, by a majority of people perceived, socially challenging situations, but there is still a lack of knowledge about how these three concepts are connected in non-embarrassing situations or personal socially challenging situations that people without SAD would mostly not fear. This is particularly important for the daily lives of the affected people since mundane tasks, like having to make a phone call, obviously occur more often than holding a presentation in front of an audience.

Furthermore, shyness was not taken into account in the theoretical background, even though SA, embarrassment, and shyness share certain qualities (Hofmann et al., 2006) and influence different aspects of social situations. The same applies to the emotion of shame (cf. Wang et al., 2020). Shame is particularly important from an emotion-focused therapeutic perspective, where shame is one of the most frequent types of maladaptive emotional schemata, with the core belief of being inherently bad or defective (Auszra et al., 2017). People talk less about shame in comparison to embarrassment (Keltner, 1996) and this emotion is more tied to one's core self in comparison to the public self, which is the focus of embarrassment (Tangney et al., 1996). While embarrassment evokes concerns about one's self-presentation, shame threatens one's self-esteem (Wan & Wyer, 2020). This additional perspective on the treatment of SAD would be important to take into account since the emotion-

focused perspective might help to understand more inner processes of participants. The emotion-focus perspective is additionally important since people with SA and SAD were consistently shown to use expressive suppression (i.e., minimizing emotion expression or changing expression to another outward emotion; Rozen & Aderka, 2023).

Article I was mainly an exploratory study with a small N that used psychology students as participants. Even though SA was taken into account, the results cannot be transferred to people with SAD. A future study would need to examine the voice parameters for a clinical sample. However, concerning the small number of participants, the machine-learning approaches showed nevertheless that the dataset was robust and yielded a satisfying prediction performance.

Article II only included the questionnaires SPS and SIAS, and the possible conclusion that can be drawn from the bifactor model depends on the input (Bornovalova et al., 2020). Although Caballo et al. (2013) showed that the five most common SA(D) questionnaires (including SPS and SIAS) revealed medium to high correlations, it is not possible to draw a conclusion about SA(D) questionnaires in general and how other questionnaires conceptualize SAD.

The results from Article IV can only be generalized to ICBT for SAD with guidance since it was not possible to test how guidance influenced the effects of the components. It was assumed in Article IV that especially participants in exposure might need more help and guidance, since there was a significantly higher dropout rate for versions with exposure in comparison to versions without exposure. However, the literature on guidance in ICBT in general shows mixed results, where some studies do not find a significant difference between self-guided and guided ICBT concerning symptom reduction of different mental disorders (SAD: Berger et al., 2011; Chen et al., 2020; Kishimoto et al., 2016; Meta-analysis SAD: Guo et al., 2021), and others do find significant differences in favor of the guided ICBT (depression: Bur et al., 2022; SAD: Wang et al., 2020). Furthermore, it was not controlled for the time spent in the program from pre to post. Even though Nordgreen et al. (2018), for example, did not find a dose-response relationship, it would be necessary to control for the influence of adherence, especially since psychoeducation showed a significant difference in the time spent on the program. One explanation for the significant difference could be that participants were more motivated to work on psychoeducation than other components, especially since psychoeducation provides a sense of control by making sense of the symptoms, which might be an additional motivator (cf. Rodrigues et al., 2018). On the other hand, exposure did not show any differences for the time spent on the program, while being at an advantage at post (as was psychoeducation), which would argue against a possible dose-response relationship. The same applies to the activities between post and follow-up: Participants had access to the program but did not receive guidance anymore. Therefore, it is not clear if and how activities after the official and guided 8-week period influenced the results from post to follow-up.

4.4. Future Implications

This subchapter will first present the overall implications for future research (4.4.1.) and for the treatment of SAD (4.4.2.) based on the findings presented in this dissertation. For detailed future implications on each specific Article, consult the future implications section of the respective Article.

4.4.1. Implications for Future Research

Concerning Article I, it would be interesting to further explore more technologically advanced methods and combine those methods with the classic self-report and rating methods. Many psychology studies solely rely on self-reporting and subjective measures like questionnaires. However, Article I has shown the additional benefits of adding objective methods like voice analysis. One of the advantages of more objective measures is the additional perspective which is less biased and can even be independent of the content that the participants deliver. In the case of voice analysis, machine learning techniques are required to process the data, but the data collection is very simple and requires only a recording device (in comparison to other objective measures like skin conduction or MRI). The combination of methods would additionally foster more cross-disciplinary collaborations and diversify perspectives on the same topic.

In line with the results from Article II, it would be interesting, particularly for Article IV, to examine the results separately for the SPS and SIAS. As Chen et al. (2020) pointed out, they did find different results for the SPS and SIAS. They argued that although the questionnaires correlate, the questionnaires claim to examine different aspects of SAD.

As was mentioned in the previous chapter, it is necessary to conduct more basic research with tasks that are not generally SA- and embarrassment-inducing (e.g., singing, holding a presentation) and that are closer to the affected people's lives. To ensure that participants still get anxious or embarrassed, a more individualized approach is suggested. Article I has shown how easily a task can be tailored to the individual participant without much effort or a big setup. Another example is the study from Reuman et al. (2015): They used vignettes where the level of threat and uncertainty varied in situations that were mundane to the students who participated (e.g., riding an elevator or submitting a term paper). A highly naturalistic and individualized approach is using ecological momentary assessment as, for example, Geyer et al., (2018) did. More specifically, the anxiety to be observed by others (which was also found as one SAD factor besides the performance and interaction fears by D'Avanzato & Dalrymple, 2016), is another aspect of embarrassment and SAD, which was deliberately left out in this dissertation. With this type of situation, one does not have to transgress a social norm in order to feel embarrassment or SA. A lot of studies presented in the theoretical background tested embarrassment, SA, and SAD with performance-related situations, where embarrassment from the dramaturgic model and, with it, the fear of simply being in the center of attention was not addressed enough. This is in line with the narrow perspective of SAD itself

according to the DSM-5, and in current research in general. A broader perspective of SAD should lead away from solely the fear of negative evaluation and consider the fear of positive evaluation as well (Skocic et al., 2015). Skocic et al. (2015) even concluded that the fear of any evaluation would be more appropriate than the fear of negative evaluation. As a result of the current definition, where the emphasis is put on the potential of negative evaluation, research tends to examine only one part of SAD (Skocic et al., 2015).

4.4.2. Implications for SAD Treatment

Regarding the treatment of SAD, psychoeducation could be expanded with information on embarrassment (especially on the usefulness of this emotion) besides the negative consequences (cf. Bas-Hoogendam et al., 2018). Embarrassment is perceived as negative, and people with SAD fear that others might see how embarrassed they are. It could help the affected people to know about the possible positive consequences of embarrassment by openly displaying it to others. Furthermore, it might facilitate exposure and help to reduce the safety behavior that aims to cover typical physiological embarrassment signs like blushing or sweating. It might also provide functional cognitions concerning the display of embarrassment and possibly facilitate the attention shift away from physiological processes. This would also make sense in light of the study results from Hope et al. (2010) where one of the main concerns of people with SAD was to experience anxiety and negative emotions, and even more interesting, therapists less often chose this as a topic for cognitive restructuring.

Article IV examined the main effects of the components. It would be necessary, as a next step, to examine their interaction effects to see how psychoeducation and exposure interact with each other and to check if and how they influence each other. However, on a larger scale, the main components could be extended with an emotion regulation module since there are studies that found emotion regulation deficits in people with SAD (Blalock et al., 2016; Dryman & Heimberg, 2018). Hawley et al. (2016) assumed in their study a potential mutual change mechanism in cognitive restructuring and exposure to be emotion regulation. It would be interesting to test possible mutual mechanisms besides the separate and component-specific mechanisms of change that are often considered (e.g., SAD knowledge increase as a change mechanism for psychoeducation). Nevertheless, in the here mentioned studies, none of those considered specific emotions. As embarrassment and SAD are strongly connected, considering and addressing embarrassment in SAD therapy could be another access point to the disorder and potentially raise the clinically relevant response rate to psychotherapy, providing a more holistic view of the disorder.

Embarrassment has been shown to be closely connected to shyness and shame, which are all important emotions regarding SAD. In order to understand the internal processes of SAD, it would be necessary to examine how people with SAD react emotionally in different and, especially, in their

personally dreaded situations. However, embarrassment may play an even more important role in Taijin-kyofusho (TKS), an SAD-related disorder, which occurs especially in Japan and Korea but was also found in other non-eastern Asian cultures. TKS is a broader concept than SAD and denotes interpersonal fear, where affected people fear making other people uncomfortable or embarrassed (APA, 2022; cf. Tei et al., 2020). Embarrassment is not only an important part of an SAD diagnosis but also of Agoraphobia or specific phobias, where affected people fear, among other things, embarrassing symptoms. Further, embarrassment also plays a role as a consequence of certain disorders, like trichotillomania and excoriation disorder, where embarrassment is counted under clinically significant distress (DSM-5 TR, criterion C; APA, 2022). This shows that embarrassment plays a crucial role across different disorders and is worth exploring further in the clinical context.

5. References

Acarturk, C., Cuijpers, P., van Straten, A., & de Graaf, R. (2009). Psychological treatment of social anxiety disorder: A meta-analysis. *Psychological Medicine*, *39*(2), 241–254. https://doi.org/10.1017/S0033291708003590

American Psychiatric Association (Ed.). (1994). *Diagnostic and Statistical Manual of Mental Disorders (4th ed.)*.

American Psychiatric Association (Ed.). (2013a). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)* (5th ed). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2013b). *DSM-5's Integrated Approach to Diagnosis and Classifications*. Psychiatry.Org. https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/DSM/APA_DSM-5-Integrated-Approach.pdf

American Psychiatric Association (Ed.). (2022). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR)*. https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425787

Amir, N., Beard, C., Taylor, C. T., Klumpp, H., Elias, J., Burns, M., & Chen, X. (2009). Attention training in individuals with generalized social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *77*(5), 961–973. https://doi.org/10.1037/a0016685

Andersson, G., Carlbring, P., Furmark, T., & on behalf of the S. O. F. I. E. Research Group. (2012). Therapist Experience and Knowledge Acquisition in Internet-Delivered CBT for Social Anxiety Disorder: A Randomized Controlled Trial. *PLoS ONE*, *7*(5), e37411. https://doi.org/10.1371/journal.pone.0037411

Andersson, G., Cuijpers, P., Carlbring, P., Riper, H., & Hedman, E. (2014). Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: A systematic review and meta-analysis. *World Psychiatry*, *13*(3), 288–295. https://doi.org/10.1002/wps.20151

Andrews, G., Basu, A., Cuijpers, P., Craske, M. G., McEvoy, P., English, C. L., & Newby, J. M. (2018). Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. *Journal of Anxiety Disorders*, *55*, 70–78. https://doi.org/10.1016/j.janxdis.2018.01.001

Andrews, G., Cuijpers, P., Craske, M. G., McEvoy, P., & Titov, N. (2010). Computer Therapy for the Anxiety and Depressive Disorders Is Effective, Acceptable and Practical Health Care: A Meta-Analysis. *PLoS ONE*, *5*(10), e13196. https://doi.org/10.1371/journal.pone.0013196

Andrews, G., Davies, M., & Titov, N. (2011). Effectiveness Randomized Controlled Trial of Face to Face Versus Internet Cognitive Behaviour Therapy for Social Phobia. *Australian & New Zealand Journal of Psychiatry*, *45*(4), 337–340. https://doi.org/10.3109/00048674.2010.538840

Auszra, L., Herrmann, I., & Greenberg, L. (2017). *Emotionsfokussierte Therapie. Ein Praxismanual*. Hogrefe.

Bas-Hoogendam, J. M., Steenbergen, H. van, Wee, N. J. A. van der, & Westenberg, P. M. (2018). Not intended, still embarrassed: Social anxiety is related to increased levels of embarrassment in response to unintentional social norm violations. *European Psychiatry*, *52*, 15–21. https://doi.org/10.1016/j.eurpsy.2018.03.002

Bastin, C., Harrison, B. J., Davey, C. G., Moll, J., & Whittle, S. (2016). Feelings of shame, embarrassment and guilt and their neural correlates: A systematic review. *Neuroscience & Biobehavioral Reviews*, *71*, 455–471. https://doi.org/10.1016/j.neubiorev.2016.09.019

Berger, T., & Caspar, F. (2011). Internetbasierte Psychotherapien. *Psychiatrie und Psychotherapie up2date*, *5*(01), 29–43. https://doi.org/10.1055/s-0030-1265934

Berger, T., Caspar, F., Richardson, R., Kneubühler, B., Sutter, D., & Andersson, G. (2011). Internet-based treatment of social phobia: A randomized controlled trial comparing unguided with two types of guided self-help. *Behaviour Research and Therapy*, *49*(3), 158–169. https://doi.org/10.1016/j.brat.2010.12.007

Berger, T., Hohl, E., & Caspar, F. (2009). Internet-based treatment for social phobia: A randomized controlled trial. *Journal of Clinical Psychology*, *65*(10), 1021–1035. https://doi.org/10.1002/jclp.20603

Blalock, D. V., Kashdan, T. B., & Farmer, A. S. (2016). Trait and Daily Emotion Regulation in Social Anxiety Disorder. *Cognitive Therapy and Research*, *40*(3), 416–425. https://doi.org/10.1007/s10608-015-9739-8

Boettcher, J., Berger, T., & Renneberg, B. (2012). Does a Pre-Treatment Diagnostic Interview Affect the Outcome of Internet-Based Self-Help for Social Anxiety Disorder? A Randomized Controlled Trial. *Behavioural and Cognitive Psychotherapy*, *40*(5), 513–528. https://doi.org/10.1017/S1352465812000501

Boettcher, J., Carlbring, P., Renneberg, B., & Berger, T. (2013). Internet-Based Interventions for Social Anxiety Disorder—An Overview. *Verhaltenstherapie*, *23*(3), 160–168. https://doi.org/10.1159/000354747

Boettcher, J., Hasselrot, J., Sund, E., Andersson, G., & Carlbring, P. (2014). Combining Attention Training with Internet-Based Cognitive-Behavioural Self-Help for Social Anxiety: A Randomised Controlled Trial. *Cognitive Behaviour Therapy*, *43*(1), 34–48. https://doi.org/10.1080/16506073.2013.809141

Bornovalova, M. A., Choate, A. M., Fatimah, H., Petersen, K. J., & Wiernik, B. M. (2020). Appropriate Use of Bifactor Analysis in Psychopathology Research: Appreciating Benefits and Limitations. *Biological Psychiatry*, *88*(1), 18–27. https://doi.org/10.1016/j.biopsych.2020.01.013

Bruffaerts, R., Harris, M. G., Kazdin, A. E., Vigo, D. V., Sampson, N. A., Chiu, W. T., Al-Hamzawi, A., Alonso, J., Altwaijri, Y. A., Andrade, L., Benjet, C., de Girolamo, G., Florescu, S., Haro, J. M., Hu, C., Karam, A., Karam, E. G., Kovess-Masfety, V., Lee, S., … The WHO World Mental Health Survey Collaborators. (2022). Perceived helpfulness of treatment for social anxiety disorder: Findings from the WHO World Mental Health Surveys. *Social Psychiatry and Psychiatric Epidemiology*, *57*(10), 2079–2095. https://doi.org/10.1007/s00127-022-02249-3

Bur, O. T., Krieger, T., Moritz, S., Klein, J. P., & Berger, T. (2022). Optimizing the context of support of web-based self-help in individuals with mild to moderate depressive symptoms: A randomized full factorial trial. *Behaviour Research and Therapy*, *152*, 104070. https://doi.org/10.1016/j.brat.2022.104070

Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., & Weiss, B. (2005). A database of German emotional speech. *Interspeech 2005*, 1517–1520. https://doi.org/10.21437/Interspeech.2005-446

Caballo, V. E., Salazar, I. C., Irurtia, M. J., Arias, B., & Nobre, L. (2013). The assessment of social anxiety through five self-report measures, LSAS-SR, SPAI, SPIN, SPS, and SIAS: A critical analysis of their factor structure. *Behavioral Psychology*, *21*(3), 423–448.

Caillaud, M., Bejanin, A., Laisney, M., Gagnepain, P., Gaubert, M., Viard, A., Clochon, P., de La Sayette, V., Allain, P., Eustache, F., & Desgranges, B. (2020). Influence of emotional complexity on the neural substrates of affective theory of mind. *Human Brain Mapping*, *41*(1), 139–149. https://doi.org/10.1002/hbm.24794

Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., & Hedman-Lagerlöf, E. (2018). Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. *Cognitive Behaviour Therapy*, *47*(1), 1–18. https://doi.org/10.1080/16506073.2017.1401115

Chen, H., Rodriguez, M. A., Qian, M., Kishimoto, T., Lin, M., & Berger, T. (2020). Predictors of treatment outcomes and adherence in internet-based cognitive behavioral therapy for social anxiety in China. *Behavioural and Cognitive Psychotherapy*, *48*(3), 291–303. https://doi.org/10.1017/S1352465819000730

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. Hope, & F. Schneider (Eds.), *Social Phobia: Diagnosis, Assessment, and Treatment* (pp. 69–93). Guilford Press.

Clark, D. M., Wild, J., Warnock-Parkes, E., Stott, R., Grey, N., Thew, G., & Ehlers, A. (2023). More than doubling the clinical benefit of each hour of therapist time: A randomised controlled trial

of internet cognitive therapy for social anxiety disorder. *Psychological Medicine*, *53*(11), Article 11. https://doi.org/10.1017/S0033291722002008

Čolić, J., Latysheva, A., Bassett, T. R., Imboden, C., Bader, K., Hatzinger, M., Mikoteit, T., Meyer, A. H., Lieb, R., Gloster, A. T., & Hoyer, J. (2020). Post-event processing after embarrassing situations: Comparing experience sampling data of depressed and socially anxious individuals. *Clinical Psychology in Europe*, *2*(4), e2867. https://doi.org/10.32872/cpe.v2i4.2867

Collins, L. M. (2018). *Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST)*. Springer International Publishing. https://doi.org/10.1007/978-3-319-72206-1

Cougle, J. R., Wilver, N. L., Day, T. N., Summers, B. J., Okey, S. A., & Carlton, C. N. (2020). Interpretation Bias Modification Versus Progressive Muscle Relaxation for Social Anxiety Disorder: A Web-Based Controlled Trial. *Behavior Therapy*, *51*(1), 99–112. https://doi.org/10.1016/j.beth.2019.05.009

Cowen, A. S., & Keltner, D. (2021). Semantic Space Theory: A Computational Approach to Emotion. *Trends in Cognitive Sciences*, *25*(2), 124–136. https://doi.org/10.1016/j.tics.2020.11.004

Craske, M., Wittchen, U., Bogels, S., Stein, M., Andrews, G., & LeBeau, R. (2013). *Severity Measure for Social Anxiety Disorder (Social Phobia)—Adult*.

D'Avanzato, C., & Dalrymple, K. L. (2016). Recent Insight Into the Subtypes of Social Anxiety Disorder. *Current Psychiatry Reports*, *18*(5), 50. https://doi.org/10.1007/s11920-016-0688-z

de Ponti, N., Matbouriahi, M., Franco, P., Harrer, M., Miguel, C., Papola, D., Sicimoğlu, A., Cuijpers, P., & Karyotaki, E. (2024). The efficacy of psychotherapy for social anxiety disorder, a systematic review and meta-analysis. *Journal of Anxiety Disorders*, *104*, 102881. https://doi.org/10.1016/j.janxdis.2024.102881

Dijk, C., Buwalda, F. M., & de Jong, P. J. (2012). Dealing with Fear of Blushing: A Psychoeducational Group Intervention for Fear of Blushing. *Clinical Psychology & Psychotherapy*, *19*(6), 481–487. https://doi.org/10.1002/cpp.764

Dijk, C., de Jong, P. J., & Peters, M. L. (2009). The remedial value of blushing in the context of transgressions and mishaps. *Emotion*, *9*(2), 287–291. https://doi.org/10.1037/a0015081

Drahota, A., Costall, A., & Reddy, V. (2008). The vocal communication of different kinds of smile. *Speech Communication*, *50*(4), 278–287. https://doi.org/10.1016/j.specom.2007.10.001

Dryman, M. T., & Heimberg, R. G. (2018). Emotion regulation in social anxiety and depression: A systematic review of expressive suppression and cognitive reappraisal. *Clinical Psychology Review*, *65*, 17–42. https://doi.org/10.1016/j.cpr.2018.07.004

Eidecker, J., Glöckner-Rist, A., & Gerlach, A. L. (2010). Dimensional structure of the Social Interaction Anxiety Scale according to the analysis of data obtained with a German version. *Journal of Anxiety Disorders*, *24*(6), 596–605. https://doi.org/10.1016/j.janxdis.2010.03.020

Ekman, P., & Cordaro, D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review*, *3*(4), 364–370. https://doi.org/10.1177/1754073911410740

Fehm, L., Pelissolo, A., Furmark, T., & Wittchen, H.-U. (2005). Size and burden of social phobia in Europe. *European Neuropsychopharmacology*, *15*(4), 453–462. https://doi.org/10.1016/j.euroneuro.2005.04.002

Fischer, A., & Manstead, A. (2016). Social Functions of Emotion and Emotion Regulation. In M. Lewis, J. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions (4th edition)* (pp. 1–27). Guilford.

Fuentes-Rodriguez, G., Garcia-Lopez, L.-J., & Garcia-Trujillo, V. (2018). Exploring the role of the DSM-5 performance-only specifier in adolescents with social anxiety disorder. *Psychiatry Research*, *270*, 1033–1038. https://doi.org/10.1016/j.psychres.2018.03.052

Gallego, A., McHugh, L., Penttonen, M., & Lappalainen, R. (2022). Measuring Public Speaking Anxiety: Self-report, behavioral, and physiological. *Behavior Modification*, *46*(4), 782–798. https://doi.org/10.1177/0145445521994308

Geyer, E. C., Fua, K. C., Daniel, K. E., Chow, P. I., Bonelli, W., Huang, Y., Barnes, L. E., & Teachman, B. A. (2018). I Did OK, but Did I Like It? Using Ecological Momentary Assessment to Examine Perceptions of Social Interactions Associated With Severity of Social

Anxiety and Depression. *Behavior Therapy*, *49*(6), 866–880. https://doi.org/10.1016/j.beth.2018.07.009

Gil, P. J. M., Carrillo, F. X. M., & Meca, J. S. (2001). Effectiveness of cognitive-behavioral treatment in social phobia: A meta-analytic review. *Psychology in Spain*, *5*(1), 17–25.

Goffman, E. (1963). *Embarrassment and Social Organization* (p. 548). John Wiley & Sons, Inc. https://doi.org/10.1037/11302-050

Gomez, R., & Watson, S. D. (2017). Confirmatory Factor Analysis of the Combined Social Phobia Scale and Social Interaction Anxiety Scale: Support for a Bifactor Model. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00070

Grawe, K. (1998). *Psychologische Therapie*. Hogrefe Verlag.

Gross, R., Olfson, M., Gameroff, M. J., Shea, S., Feder, A., Lantigua, R., Fuentes, M., & Weissman, M. M. (2005). Social anxiety disorder in primary care. *General Hospital Psychiatry*, *27*(3), 161–168. https://doi.org/10.1016/j.genhosppsych.2005.01.006

Guo, S., Deng, W., Wang, H., Liu, J., Liu, X., Yang, X., He, C., Zhang, Q., Liu, B., Dong, X., Yang, Z., Li, Z., & Li, X. (2021). The efficacy of internet-based cognitive behavioural therapy for social anxiety disorder: A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy*, *28*(3), Article 3. https://doi.org/10.1002/cpp.2528

Harris, C. R. (2006). Embarrassment: A Form of Social Pain: This enigmatic emotion likely evolved to smooth social interations, but it can have less desirable consequences in the modern world. *American Scientist, 94*(6), 524–533.

Hawley, L. L., Rector, N. A., & Laposa, J. M. (2016). Examining the dynamic relationships between exposure tasks and cognitive restructuring in CBT for SAD: Outcomes and moderating influences. *Journal of Anxiety Disorders*, *39*, 10–20. https://doi.org/10.1016/j.janxdis.2016.01.010

Heeren, A., Mogoaşe, C., Philippot, P., & McNally, R. J. (2015). Attention bias modification for social anxiety: A systematic review and meta-analysis. *Clinical Psychology Review*, *40*, 76–90. https://doi.org/10.1016/j.cpr.2015.06.001

Hofmann, S. G. (2004). Cognitive Mediation of Treatment Change in Social Phobia. *Journal of Consulting and Clinical Psychology*, *72*(3), 392–399. https://doi.org/10.1037/0022-006X.72.3.392

Hofmann, S. G., Anu Asnaani, M. A., & Hinton, D. E. (2010). Cultural aspects in social anxiety and social anxiety disorder. *Depression and Anxiety*, *27*(12), 1117–1127. https://doi.org/10.1002/da.20759

Hofmann, S. G., Moscovitch, D. A., & Kim, H.-J. (2006). Autonomic correlates of social anxiety and embarrassment in shy and non-shy individuals. *International Journal of Psychophysiology*, *61*(2), 134–142. https://doi.org/10.1016/j.ijpsycho.2005.09.003

Hofmann, S. G., & Smits, J. A. J. (2008). Cognitive-behavioral therapy for adult anxiety disorders: A meta-analysis of randomized placebo-controlled trials. *The Journal of Clinical Psychiatry*, *69*(4), 621–632. https://doi.org/10.4088/jcp.v69n0415

Hope, D. A., Burns, J. A., Hayes, S. A., Herbert, J. D., & Warner, M. D. (2010). Automatic Thoughts and Cognitive Restructuring in Cognitive Behavioral Group Therapy for Social Anxiety Disorder. *Cognitive Therapy and Research*, *34*(1), 1–12. https://doi.org/10.1007/s10608-007-9147-9

Hyett, M. P., & McEvoy, P. M. (2018). Social anxiety disorder: Looking back and moving forward. *Psychological Medicine*, *48*(12), 1937–1944. https://doi.org/10.1017/S0033291717003816

Issakidis, C., & Andrews, G. (2002). Service utilisation for anxiety in an Australian community sample. *Social Psychiatry and Psychiatric Epidemiology*, *37*(4), 153–163. https://doi.org/10.1007/s001270200009

Kaczkurkin, A. N., & Foa, E. B. (2015). Cognitive-behavioral therapy for anxiety disorders: An update on the empirical evidence. *Dialogues in Clinical Neuroscience*, *17*(3), 337–346. https://doi.org/10.31887/DCNS.2015.17.3/akaczkurkin

Kadali, D. B., & Mittal, V. K. (2020). Studies on Paralinguistic Sounds, Emotional Speech and Expressive Voices. *Workshop on Speech, Music and Mind (SMM 2020)*, 11–15. https://doi.org/10.21437/SMM.2020-3

Kazdin, A. E. (2007). Mediators and Mechanisms of Change in Psychotherapy Research. *Annual Review of Clinical Psychology*, *3*(1), 1–27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432

Kazdin, A. E. (2017). Addressing the treatment gap: A key challenge for extending evidence-based psychosocial interventions. *Behaviour Research and Therapy*, *88*, 7–18. https://doi.org/10.1016/j.brat.2016.06.004

Keltner, D. (1996). Evidence for the Distinctness of Embarrassment, Shame, and Guilt: A Study of Recalled Antecedents and Facial Expressions of Emotion. *Cognition & Emotion*, *10*(2), 155–172. https://doi.org/10.1080/026999396380312

Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional Expression: Advances in Basic Emotion Theory. *Journal of Nonverbal Behavior*, *43*(2), 133–160. https://doi.org/10.1007/s10919-019-00293-3

Kindred, R., Bates, G. W., & McBride, N. L. (2022). Long-term outcomes of cognitive behavioural therapy for social anxiety disorder: A meta-analysis of randomised controlled trials. *Journal of Anxiety Disorders*, *92*, 102640. https://doi.org/10.1016/j.janxdis.2022.102640

Kishimoto, T., Krieger, T., Berger, T., Qian, M., Chen, H., & Yang, Y. (2016). Internet-Based Cognitive Behavioral Therapy for Social Anxiety with and without Guidance Compared to a Wait List in China: A Propensity Score Study. *Psychotherapy and Psychosomatics*, *85*(5), 317–319. https://doi.org/10.1159/000446584

Knappe, S., Klotsche, J., Heyde, F., Hiob, S., Siegert, J., Hoyer, J., Strobel, A., LeBeau, R. T., Craske, M. G., Wittchen, H.-U., & Beesdo-Baum, K. (2014). Test–retest reliability and sensitivity to change of the dimensional anxiety scales for DSM-5. *CNS Spectrums*, *19*(3), 256–267. https://doi.org/10.1017/S1092852913000710

Laukka, P., Linnman, C., Åhs, F., Pissiota, A., Frans, Ö., Faria, V., Michelgård, Å., Appel, L., Fredrikson, M., & Furmark, T. (2008). In a Nervous Voice: Acoustic Analysis and Perception of Anxiety in Social Phobics' Speech. *Journal of Nonverbal Behavior*, *32*(4), 195–214. https://doi.org/10.1007/s10919-008-0055-9

Lawler, K., Earley, C., Timulak, L., Enrique, A., & Richards, D. (2021). Dropout From an Internet-Delivered Cognitive Behavioral Therapy Intervention for Adults With Depression and Anxiety: Qualitative Study. *JMIR Formative Research*, *5*(11), e26221. https://doi.org/10.2196/26221

Leary, M. R., & Hoyle, R. H. (2013). *Handbook of Individual Differences in Social Behavior*. Guilford Publications.

Leary, M. R., & Kowalski, R. M. (1997). *Social Anxiety*. Guilford Press.

LeBeau, R. T., Bögels, S., Möller, E., & Craske, M. (2015). Integrating dimensional assessment and categorical diagnosis in DSM-5: The benefits and challenges of the paradigm shift for the anxiety disorders. *Psychopathology Review*, *a2*(1), 83–99. https://doi.org/10.5127/pr.036414

LeBeau, R. T., Glenn, D. E., Hanover, L. N., Beesdo-Baum, K., Wittchen, H., & Craske, M. G. (2012). A dimensional approach to measuring anxiety for DSM-5. *International Journal of Methods in Psychiatric Research*, *21*(4), 258–272. https://doi.org/10.1002/mpr.1369

Liu, H., Li, X., Han, B., & Liu, X. (2017). Effects of cognitive bias modification on social anxiety: A meta-analysis. *PLOS ONE*, *12*(4), e0175107. https://doi.org/10.1371/journal.pone.0175107

Lopes, R. C. T., Šipka, D., Krieger, T., Klein, J. P., & Berger, T. (2021). Optimizing cognitive-behavioral therapy for social anxiety disorder and understanding the mechanisms of change: Study protocol for a randomized factorial trial. *Internet Interventions*, *26*, 100480. https://doi.org/10.1016/j.invent.2021.100480

Low, D. M., Bentley, K. H., & Ghosh, S. S. (2020). Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investigative Otolaryngology*, *5*(1), 96–116. https://doi.org/10.1002/lio2.354

Maire, H., & Agnoletti, M.-F. (2020). Expressing embarrassment (or not): Which effects on produced impression on others and on attributed social value? *European Review of Applied Psychology*, *70*(4), 100525. https://doi.org/10.1016/j.erap.2020.100525

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, *36*(4), 455–470. https://doi.org/10.1016/S0005-7967(97)10031-6

Mattick, R. P., Peters, L., & Clarke, J. C. (1989). Exposure and cognitive restructuring for social phobia: A controlled study. *Behavior Therapy*, *20*(1), 3–23. https://doi.org/10.1016/S0005-7894(89)80115-7

Mayo-Wilson, E., Dias, S., Mavranezouli, I., Kew, K., Clark, D. M., Ades, A. E., & Pilling, S. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: A systematic review and network meta-analysis. *The Lancet Psychiatry*, *1*(5), 368–376. https://doi.org/10.1016/S2215-0366(14)70329-3

Miller, R. S. (1995). On the Nature of Embarrassabllity: Shyness, Social Evaluation, and Social Skill. *Journal of Personality*, *63*(2), 315–339. https://doi.org/10.1111/j.1467-6494.1995.tb00812.x

Miller, R. S. (2009). Social anxiousness, shyness, and embarrassability. In *Handbook of individual differences in social behavior* (pp. 176–191). The Guilford Press.

Miller, R. S. (2010). Chapter 4—Are Embarrassment and Social Anxiety Disorder Merely Distant Cousins, or Are They Closer Kin? In S. G. Hofmann & P. M. DiBartolo (Eds.), *Social Anxiety (Second Edition)* (pp. 93–118). Academic Press. https://doi.org/10.1016/B978-0-12-375096-9.00004-3

Miller, R. S. (2012). The interactive origins and outcomes of embarrassment. In W. R. Crozier & P. J. de Jong (Eds.), *The Psychological Significance of the Blush* (1st ed., pp. 185–202). Cambridge University Press. https://doi.org/10.1017/CBO9781139012850.013

Miller, R. S., & Tangney, J. P. (1994). Differentiating Embarrassment and Shame. *Journal of Social and Clinical Psychology*, *13*(3). https://www.proquest.com/docview/1292186743/citation/9F6BBC5A7B70447BPQ/1

Müller-Pinzler, L., Paulus, F. M., Stemmler, G., & Krach, S. (2012). Increased autonomic activation in vicarious embarrassment. *International Journal of Psychophysiology*, *86*(1), 74–82. https://doi.org/10.1016/j.ijpsycho.2012.07.183

National Institute for Health and Care Excellence. (2013). *Social anxiety: Recognition, assessment and treatment: Clinical guideline*. https://www.nice.org.uk/guidance/cg159

Nordgreen, T., Gjestad, R., Andersson, G., Carlbring, P., & Havik, O. E. (2018). The effectiveness of guided internet-based cognitive behavioral therapy for social anxiety disorder in a routine care setting. *Internet Interventions, 13,* 24–29. https://doi.org/10.1016/j.invent.2018.05.003

Nordmo, M., Sinding, A. I., Carlbring, P., Andersson, G., Havik, O. E., & Nordgreen, T. (2015). Internet-delivered cognitive behavioural therapy with and without an initial face-to-face psychoeducation session for social anxiety disorder: A pilot randomized controlled trial. *Internet Interventions*, *2*(4), 429–436. https://doi.org/10.1016/j.invent.2015.10.003

Olthuis, J. V., Watt, M. C., Bailey, K., Hayden, J. A., & Stewart, S. H. (2016). Therapist-supported Internet cognitive behavioural therapy for anxiety disorders in adults. *Cochrane Database of Systematic Reviews*, *2016*(3), Article 3. https://doi.org/10.1002/14651858.CD011565.pub2

Peters, L., Sunderland, M., Andrews, G., Rapee, R. M., & Mattick, R. P. (2012). Development of a short form Social Interaction Anxiety (SIAS) and Social Phobia Scale (SPS) using nonparametric item response theory: The SIAS-6 and the SPS-6. *Psychological Assessment*, *24*(1), 66–76. https://doi.org/10.1037/a0024544

Powers, M. B., Sigmarsson, S. R., & Emmelkamp, P. M. G. (2008). A meta-analytic review of psychological treatments for social anxiety disorder. *International Journal of Cognitive Therapy*, *1*(2), 94–113. https://doi.org/10.1521/ijct.2008.1.2.94

Rapee, R. M. (1998). *Overcoming Shyness and Social Phobia: A Step-by-Step Guide*. Jason Aronson.

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and Structural Coefficient Bias in Structural Equation Modeling: A Bifactor Perspective. *Educational and Psychological Measurement*, *73*(1), 5–26. https://doi.org/10.1177/0013164412449831

Reuman, L., Jacoby, R. J., Fabricant, L. E., Herring, B., & Abramowitz, J. S. (2015). Uncertainty as an anxiety cue at high and low levels of threat. *Journal of Behavior Therapy and Experimental Psychiatry*, *47*, 111–119. https://doi.org/10.1016/j.jbtep.2014.12.002

Rice, K., Schutte, N., Rock, A., & Murray, C. (2021). Structure, validity and cut-off scores for the APA emerging measure of DSM-5 Social Anxiety Disorder Severity Scale (SAD-D). *Journal of Anxiety & Depression*, *10*, 406.

Rodebaugh, T. L., Woods, C. M., & Heimberg, R. G. (2007). The Reverse of Social Anxiety Is Not Always the Opposite: The Reverse-Scored Items of the Social Interaction Anxiety Scale Do Not Belong. *Behavior Therapy*, *38*(2), 192–206. https://doi.org/10.1016/j.beth.2006.08.001

Rodebaugh, T. L., Woods, C. M., Heimberg, R. G., Liebowitz, M. R., & Schneier, F. R. (2006). The factor structure and screening utility of the Social Interaction Anxiety Scale. *Psychological Assessment*, *18*(2), 231–237. https://doi.org/10.1037/1040-3590.18.2.231

Rodrigues, F., Bartolo, A., Pacheco, E., Pereira, A., Silva, C. F., & Oliveira, C. (2018). Psycho-Education for Anxiety Disorders in Adults: A Systematic Review of its Effectiveness. *Journal of Forensic Psychology*, *03*(02). https://doi.org/10.4172/2475-319X.1000142

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. https://doi.org/10.1037/met0000045

Rozen, N., & Aderka, I. M. (2023). Emotions in social anxiety disorder: A review. *Journal of Anxiety Disorders*, *95*, 102696. https://doi.org/10.1016/j.janxdis.2023.102696

Ruscio, A. M. (2010). The latent structure of social anxiety disorder: Consequences of shifting to a dimensional diagnosis. *Journal of Abnormal Psychology*, *119*(4), 662–671. https://doi.org/10.1037/a0019341

Ruscio, A. M., Brown, T. A., Chiu, W. T., Sareen, J., Stein, M. B., & Kessler, R. C. (2008). Social fears and social phobia in the USA: Results from the National Comorbidity Survey Replication. *Psychological Medicine*, *38*(01), 15–28. https://doi.org/10.1017/S0033291707001699

Schmidt, N. B., Richey, J. A., Buckner, J. D., & Timpano, K. R. (2009). Attention training for generalized social anxiety disorder. *Journal of Abnormal Psychology*, *118*(1), 5–14. https://doi.org/10.1037/a0013643

Schulz, A., Stolz, T., Vincent, A., Krieger, T., Andersson, G., & Berger, T. (2016). A sorrow shared is a sorrow halved? A three-arm randomized controlled trial comparing internet-based clinician-guided individual versus group treatment for social anxiety disorder. *Behaviour Research and Therapy*, *84*, 14–26. https://doi.org/10.1016/j.brat.2016.07.001

Simon-Thomas, E. R., Keltner, D. J., Sauter, D., Sinicropi-Yao, L., & Abramson, A. (2009). The voice conveys specific emotions: Evidence from vocal burst displays. *Emotion*, *9*(6), 838–846. https://doi.org/10.1037/a0017810

Skocic, S., Jackson, H., & Hulbert, C. (2015). Beyond DSM-5: An alternative approach to assessing Social Anxiety Disorder. *Journal of Anxiety Disorders*, *30*, 8–15. https://doi.org/10.1016/j.janxdis.2014.12.006

Skocic, S., Jackson, H., Hulbert, C., & Faber, C. (2016). The Maintaining Factors of Social Anxiety: A Three-Group Comparison of a Clinical Sample with Highly Socially Anxious Students and Non-Anxious Students. *Behavioural and Cognitive Psychotherapy*, *44*(4), 385–396. https://doi.org/10.1017/S1352465815000685

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar de Pablo, G., Il Shin, J., Kirkbride, J. B., Jones, P., Kim, J. H., Kim, J. Y., Carvalho, A. F., Seeman, M. V., Correll, C. U., & Fusar-Poli, P. (2022). Age at onset of mental disorders worldwide: Large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*, *27*(1), 281–295. https://doi.org/10.1038/s41380-021-01161-7

Spence, S. H., & Rapee, R. M. (2016). The etiology of social anxiety disorder: An evidence-based model. *Behaviour Research and Therapy*, *86*, 50–67. https://doi.org/10.1016/j.brat.2016.06.007

Stangier, U., Heidenreich, T., Berardi, A., Golbs, U., & Hoyer, J. (1999). Die Erfassung sozialer Phobie durch die Social Interaction Anxiety Scale (SIAS) und die Social Phobia Scale (SPS). *Zeitschrift Für Klinische Psychologie Und Psychotherapie*, *28*(1), 28–36. https://doi.org/10.1026//0084-5345.28.1.28

Stangier, U., Heidenreich, T., & Peitz, M. (2003). *Soziale Phobien: Ein kognitiv-verhaltenstherapeutisches Behandlungsmanual*. Beltz, PVU.

Stein, D. J., Lim, C. C. W., Roest, A. M., de Jonge, P., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Benjet, C., Bromet, E. J., Bruffaerts, R., de Girolamo, G., Florescu, S., Gureje, O., Haro, J. M., Harris, M. G., He, Y., Hinkov, H., Horiguchi, I., Hu, C., … WHO World Mental Health Survey Collaborators. (2017). The cross-national epidemiology of social anxiety disorder:

Data from the World Mental Health Survey Initiative. *BMC Medicine*, *15*(1), 143. https://doi.org/10.1186/s12916-017-0889-2

Stocks, E. L., Lishner, D. A., Waits, B. L., & Downum, E. M. (2011). I'm Embarrassed for You: The Effect of Valuing and Perspective Taking on Empathic Embarrassment and Empathic Concern. *Journal of Applied Social Psychology*, *41*(1), 1–26. https://doi.org/10.1111/j.1559-1816.2010.00699.x

Stolz, T., Schulz, A., Krieger, T., Vincent, A., Urech, A., Moser, C., Westermann, S., & Berger, T. (2018). A mobile app for social anxiety disorder: A three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. *Journal of Consulting and Clinical Psychology*, *86*(6), 493–504. https://doi.org/10.1037/ccp0000301

Sun, X., So, S. H. W., Chung, L. K. H., Chiu, C.-D., Chan, R. C. K., & Leung, P. W. L. (2022). Longitudinal bifactor modeling of anxiety, depression and schizotypy—The role of rumination as a shared mechanism. *Schizophrenia Research*, *240*, 153–161. https://doi.org/10.1016/j.schres.2022.01.005

Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, *70*, 1256–1269. https://doi.org/10.1037/0022-3514.70.6.1256

Tei, S., Kauppi, J.-P., Jankowski, K. F., Fujino, J., Monti, R. P., Tohka, J., Abe, N., Murai, T., Takahashi, H., & Hari, R. (2020). Brain and behavioral alterations in subjects with social anxiety dominated by empathic embarrassment. *Proceedings of the National Academy of Sciences*, *117*(8), 4385–4391. https://doi.org/10.1073/pnas.1918081117

Thompson, T., Kaminska, M., Marshall, C., & Van Zalk, N. (2019). Evaluation of the social phobia scale and social interaction anxiety scale as assessments of performance and interaction anxiety. *Psychiatry Research*, *273*, 725–731. https://doi.org/10.1016/j.psychres.2019.01.103

Tracy, J. L., & Robins, R. W. (2004). TARGET ARTICLE: "Putting the Self Into Self-Conscious Emotions: A Theoretical Model." *Psychological Inquiry*, *15*(2), 103–125. https://doi.org/10.1207/s15327965pli1502_01

Wan, L. C., & Wyer, R. S. (2020). The Role of Incidental Embarrassment in Social Interaction Behavior. *Social Cognition*, *38*(5), 422–446. https://doi.org/10.1521/soco.2020.38.5.422

Wang, H., Zhao, Q., Mu, W., Rodriguez, M., Qian, M., & Berger, T. (2020). The Effect of Shame on Patients With Social Anxiety Disorder in Internet-Based Cognitive Behavioral Therapy: Randomized Controlled Trial. *JMIR Mental Health*, *7*(7), e15797. https://doi.org/10.2196/15797

Weeks, J. W., Lee, C.-Y., Reilly, A. R., Howell, A. N., France, C., Kowalsky, J. M., & Bush, A. (2012). "The Sound of Fear": Assessing vocal fundamental frequency as a physiological indicator of social anxiety disorder. *Journal of Anxiety Disorders*, *26*(8), 811–822. https://doi.org/10.1016/j.janxdis.2012.07.005

Wong, Q. J. J., Chen, J., Gregory, B., Baillie, A. J., Nagata, T., Furukawa, T. A., Kaiya, H., Peters, L., & Rapee, R. M. (2019). Measurement equivalence of the Social Interaction Anxiety Scale (SIAS) and Social Phobia Scale (SPS) across individuals with social anxiety disorder from Japanese and Australian sociocultural contexts. *Journal of Affective Disorders*, *243*, 165–174. https://doi.org/10.1016/j.jad.2018.09.028

World Health Organization (Ed.). (2024). *ICD-11: International classification of diseases (11th revision)*. https://icd.who.int/

Appendix A

Article I – Embarrassment

By the time this dissertation was submitted for publication (21.01.25), this Article was still in the review process. Therefore, the last version (21.01.2025) that was re-submitted to *Nature Scientific Reports* is attached here.

Multidisciplinary Characterization of Embarrassment through Behavioral and Acoustic Modeling

Dajana Šipka[1]
Bogdan Vlasenko[2]
Maria Stein[1, 3]
Thomas Dierks[3]
Mathew Magimai-Doss[2]
Yosuke Morishima[3]

[1] Department of Clinical Psychology and Psychotherapy, University of Bern, Switzerland
[2] Idiap Research Institute, Martigny, Switzerland
[3] University Hospital of Psychiatry and Psychotherapy, University of Bern, Switzerland

E-Mails:
dajana.sipka@unibe.ch
bogdan.vlasenko@idiap.ch
maria.stein@unibe.ch
thomas.dierks@unibe.ch
mathew@idiap.ch
yosuke.morishima@unibe.ch

*Corresponding author at:
Dajana Šipka
University of Bern,
Institute for Psychology
Clinical Psychology and Psychotherapy Department,
Fabrikstrasse 8
3012 Bern
Switzerland
Tel.: +41 31 684 34 98
E-Mail: dajana.sipka@unibe.ch (Šipka, D.)

Abstract

Introduction: Embarrassment is a social emotion that shares many characteristics with social anxiety (SA). Most people experience embarrassment in their daily lives, but it is quite unattended in research. We characterized embarrassment in an interdisciplinary approach, introducing a behavioral paradigm and applying machine learning approaches, including acoustic analyses. **Methods:** 33 participants had to write about an embarrassing experience and had then, without knowing it prior, to read it out loud to the conductor. Embarrassment was then examined with two different approaches: Firstly, from a subjective view, with self-report measures from the participants. Secondly, from an objective, machine-learning approach, where trained models tested the robustness of our embarrassment data set (i.e., prediction accuracy), and then described embarrassment in a dimensional (i.e., dimension: valence, activation, dominance; VAD) and categorical (i.e., comparing embarrassment to other emotional states) way. **Results:** The subjective rating of embarrassment was increased after participants read their stories out loud, and participants with higher SA scores experienced higher embarrassment than participants with lower SA scores. The state of embarrassment was predicted at 86.4% at the best of the unweighted average recall rate. While the simple VAD dimensional analyses did not differentiate between the state of embarrassment and the references, the complex emotional category analyses characterized embarrassment as closer to boredom, a neutral state, and less of sadness. **Conclusion:** Combining an effective behavioral paradigm and advanced acoustic modeling, we characterized the emotional state of embarrassment, and the identified characteristics could be used as a biomarker to assess SA.

1. Introduction

Embarrassment is a social emotion that most of us experience through daily life, and it occurs when a desired social image of oneself is threatened (Miller & Tangney, 1994; Tangney et al., 1996). It can be seen as a self-conscious and, simultaneously, other-conscious emotion: While one is more self-aware in an embarrassing situation, one is also concerned about other people's judgment (Tangney et al., 1996). Embarrassment is experienced mostly in public with other people. However, it can still occur privately when the audience is imagined and is more likely to occur around strangers than loved ones. Embarrassment is usually accompanied by typical physiological changes, such as blushing (Hofmann et al., 2006; Keltner, 2005; Miller, 2012), or changes in the voice and non-verbal behavior, such as avoiding eye contact or lowering one's head. Tangney et al. (1996) categorized embarrassment as a "negatively valenced emotion" (pp. 1264), but there are also positive consequences that can arise for the person affected as well as the audience: According to Miller (2012), people often react helpfully in embarrassing situations, and showing embarrassment in the form of blushing, for example, elicits a favorable impression of the affected person. Stocks et al. (2011) additionally differed between personal and empathic embarrassment. Whilst the first is experienced for oneself, the latter is experienced for another person while, for example, observing an embarrassing task. This study, however, focused solely on personal embarrassment.

There are several emotions that are similar to embarrassment, particularly shame, since for a long time, embarrassment was considered to be part of shame. However, embarrassment is viewed as an emotion on its own today and there are distinct differences between these two emotions: Shame is related to more moral transgressions, can occur when being alone, is a more intense emotion and lasts longer than embarrassment. Additionally, embarrassment tends to occur around less familiar people in comparison to shame (see e.g., Miller & Tangney, 1994; Rozen & Aderka, 2023; Tangney et al., 1996).Core aspects of embarrassment, such as the fear of negative evaluation, fear of being rejected by others, and heightened self-consciousness, are also important aspects of social anxiety (SA) as well as social anxiety disorder (SAD). According to Rozen and Aderka (2023) embarrassment, SA, and SAD were consistently associated with each other in physiological measures, neural activities, and self-reports of emotions for clinical and non-clinical samples. Socially anxious people are generally more easily embarrassed and respond with more intense embarrassment than less socially anxious people (Leary & Hoyle, 2013). Furthermore, Rozen & Aderka (2023) showed in their review that embarrassment has found to be associated with SA in clinical as well as non-clinical samples over different studies. In both cohorts, people with either higher SA scores or SAD rated social blunders as more embarrassing in comparison to people with lower SA scores respectively without SAD. As with embarrassment, SA occurs mostly in public; if in private, an imagined audience or an imagined reaction from others is necessary. While embarrassment appears sudden, brief, and due to an actual misstep, SA appears gradually and over a longer period of time and can occur without having done anything wrong (Miller, 2009).

Embarrassment is generally neglected in research on basic emotions (Simon-Thomas et al., 2009). Especially when papers report on vocal cues in different emotions, they often examine either the classical basic emotions according to Ekman & Friesen (1971) or other emotions than embarrassment (e.g., Devillers & Vidrascu, 2007; Juslin et al., 2018; Patel et al., 2011; Sauter et al., 2010). Due to the consistent association of embarrassment and SA and the fact that embarrassment is still a neglected emotion, it is important to look more into the emotion itself, its relationship to other emotions, and how it is associated with SA from a basic research and clinical point of view.

Therefore, this paper's main goal was to exploratorily examine embarrassment and capture the emotion from different points of view. On the one hand, embarrassment can be compared categorically to other emotions; it can be described as how it relates to and shares information with them. There are different data sets in different languages, consisting of emotional speech, where the data has already been labelled and tested accordingly (see e.g., Burkhardt et al., 2005). On the other hand, embarrassment itself can be described dimensionally in more depth. Grimm et al. (2007) proposed a three-dimensional emotion space consisting of the axis valence, activation, and dominance (VAD), which was also used in this study to describe embarrassment dimensionally.

Previous studies investigating fundamental emotional research used either classical subjective clinical psychology approaches, relying mainly on the participants' self-reports and ratings of a few individual experts, or more objective measures such as neuroimaging methods and psychophysiological measures (Bastin et al., 2016). If subjective and objective measures are combined in embarrassment studies, they often focus on somatic or neuronal features (Hofmann et al., 2006; Müller-Pinzler et al., 2012) but seldom on voice parameters. The same goes for physiological indicators of SA or embarrassment: Research most often investigated body, hand, and head movements or gaze activity (Keltner et al., 2019). According to Weeks et al. (2012), there are several advantageous characteristics of voice parameters as physiological indicators of, for example, SAD, such as being less biased to subjects' responses and more objective than through self-questionnaires. So, even though participants do not say the same, one can objectively compare the paralinguistic information of their answers (Burkhardt et al., 2005).

Human speech can be divided into verbal (linguistic) and non-verbal (paralinguistic) sounds. While it is obvious that verbal sounds play an important role in communication, non-verbal aspects, such as paralinguistics, carry a lot of additional information in conversations, such as the emotional and mental state of the person speaking (Kadali & Mittal, 2020). Conversely, changes in the voice may indicate changes in a person's mental and emotional state. This study, therefore, combined subjective measures (see section 2.3.) and objective engineering and machine learning approaches (see section 2.4.) to examine embarrassment.

To describe and examine embarrassment from different points of view, this paper had the following four goals: The first goal was to induce embarrassment in participants and test whether the induction was successful and, if so, how embarrassment was related to SA. We hypothesized that

embarrassment would indeed be induced, with the participants being significantly more embarrassed during the embarrassment induction task compared to the pre- and post-induction periods. Furthermore, according to previous studies, we also assumed that participants with higher SA scores would get more embarrassed than participants with lower SA scores. Verifying that embarrassment was induced was the only goal with a hypothesis. This was the fundament for the other three following exploratory goals, where we used our acoustic data to gain further insights using automatic speech processing techniques. The second goal was to test how well our trained model could predict our sample data in pre-induction, embarrassment, and post-induction and show the robustness of our embarrassment data. The third goal was to adopt a dimensional approach and map embarrassment onto the VAD dimension. The fourth and last goal was comparing embarrassment to other emotions, thus following a categorical approach. For the third and fourth goals, publicly available emotional speech corpora (i.e., acoustic samples with emotional labels) were used to train our models.

2. Materials & Methods

2.1. Participants

Undergraduate psychology students from the University of Berne, Switzerland, were recruited for this study. The average age of the sample (N = 33) was 23.73 years (SD = 4.68), with 78.8% female participants (n = 26). The exclusion criteria were a.) any current or past neurological or psychological disorders, b.) regular medication intake, and c.) regular substance abuse. For their participation, they received 1.5 out of 12 mandatory experimental credits. All participants gave written informed consent to participate in the study. The ethics committee of the faculty of human sciences at the University of Bern approved the study protocol (2020-06-00004). The experiment was performed in accordance with the relevant guidelines and regulations according to the Declaration of Helsinki.

2.2. Procedure

See Fig. 1 for an overview of the study procedure. If the participants did not fulfil any exclusion criteria, they advanced to fill out the questionnaires Short Form Social Phobia Scale (SPS-6), Short Form Social Interaction Anxiety Scale (SIAS-6), Community Assessment of Psychic Experiences (CAPE), and demographic data online via Qualtrics not later than 12 hours before the assessment appointment.

The whole assessment (t0 – t3) was audio recorded, and the participants were informed prior to the start of the assessment. The experimenter conductor was instructed to respond neutrally respectively to not give either a positive reaction (e.g., laugh along with the participant) or a negative reaction (e.g., look dismissive) during the procedure and had all instructions written on a script to maximize standardization. This study was embedded in a larger study based on Mota et al. (2017) and represents part 2 of the procedure. Since part 1 is irrelevant to this study, it was not further explained here (see Mota et al., 2017, for the study procedure of part 1). At t0, participants were shown a neutral picture from the picture set of Mota et al. (2017) and were instructed to tell a story about the picture for at least 30 seconds. If they did not meet the 30-second mark, they were asked to talk more about the picture. At t1, they received a short explanation of the characteristics of embarrassment and how it differed from shame (see introduction for differences). The assignment then was to write about a situation where they felt embarrassed, which happened no longer than one year ago. The participants were instructed to write only a few sentences about the setting of the embarrassing situation and then to mainly focus on the following aspects of their experience in the situation: emotions (i.e., “What did you feel in this situation?”), cognitions, physiological reactions, their own behavior, and the behavior of the bystanders. Therewith, participants were more likely to be immersed in the whole experience and not only recollect mere cold facts of the situation (Tangney et al., 1996). After being informed orally, they also received all instructions in written form. Participants had a maximum of 4 minutes to come up with a situation and then a maximum of 14 minutes to write about the situation. At t2, they were asked to read the story aloud to the instructor and were instructed to talk for at least 30 seconds. If they did not meet the 30-

second mark, they were asked to talk more about the situation. At t3, they had the same task as at t0 with the same instructions. Between each time point (t0 – t3), they had to indicate their level of embarrassment on a visual analog scale (VAS; see section 2.3.3.).

At the end of the study (after t3), participants were asked to guess the study's purpose. They were then informed about the actual purpose of the study. For ethical reasons, they were asked to state their current well-being, and the instructor intervened if a participant stated particularly low well-being.

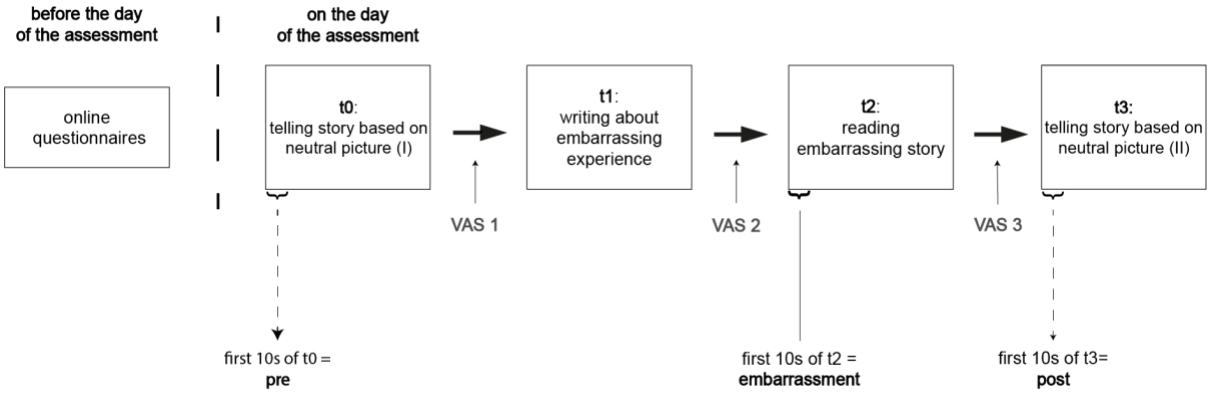


Figure 1 *Overview study procedure*

Before starting the main assessment, participants answered online questionnaires. On the day of the assessment, participants were told to tell a story based on a neutral picture before (t0) and after (t3) embarrassment induction (t2). The first ten seconds of speech in t0, t2, t3 were used as pre-embarrassment induction (pre), embarrassment induction, (emb), and post-embarrassment induction (post) for the acoustic analysis. VAS = visual analog scale.

2.3. Measures

2.3.1. Short Form Social Phobia Scale (SPS-6) and Short Form Social Interaction Anxiety Scale (SIAS-6)

The Social Phobia Scale (SPS; Mattick & Clarke, 1998) and the Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998) are two self-report questionnaires to measure different aspects of SA. The SPS measures SA in performance-related situations (e.g., fear of attracting attention while queueing), while SIAS measures SA in interactional situations (e.g., difficulty in talking to other people). For this study, the short form of the SPS (SPS-6; Peters et al., 2012) and the SIAS (SIAS-6, Peters et al., 2012) were used, which were directly translated into German for this study. The two questionnaires are presented together with 6 SPS-6 items and 6 SIAS-6 items rated on a Likert-Scale from 0, “not at all”, to 4, “extremely”, with a total sum score ranging from 0 to 48. According to Peters et al. (2012), the cut-off scores for clinically relevant SA symptoms are SPS-6 $\geq$ 2 and SIAS-6 $\geq$ 7.

2.3.2. Visual Analog Scale (VAS)

For the measurement of the strength of embarrassment participants felt before, during, and after the intervention, a paper-pencil visual analog scale (VAS) was used at each time point (i.e., after t0, t1, and t2). Participants had to assess on a 10cm long line how embarrassed they felt at the moment from 0, “not embarrassed at all”, on the left end to 10, “extremely embarrassed”, on the right end. The measured length from the zero point to the participants’ mark on each line indicated the strength of embarrassment for each time point (cf. Delgado et al., 2018). The interrater variability of two independent raters who measured the distance between all lines was $r(97) = 0.998$.

2.3.3. Recording procedure and device

Each participant's assessment was audio-recorded from t0 to t3. The recording device was always placed on the table in front of the participant with approx. 1 meter from the edge of the table opposite the participant. The recording device SONY ICD-UX570 was used. The sample rate was at 44.1 kHz with a bit depth of 16-bit. The audio data was stored in “Waveform Audio File Format” (.wav).

2.4. Statistical analysis – behavioral data

The statistical analysis of the behavioral data was performed with R Studio (Version: 2024.04.2+764; Posit team, 2024). Descriptive statistics were reported, and to examine differences and relationships, non-parametric tests were used due to a lack of normal distribution in the behavioral data (see section 3.1.). Linear mixed models (LMMs) were conducted to test the influence of gender.

2.5. Acoustic Data Analysis

The acoustic data from each participant’s first 10 seconds at the beginning of speech in t0, t2, and t3 phases was used for pre-embarrassment induction (pre), embarrassment induction (emb), and post-embarrassment induction (post), respectively. A fixed length of initial 10 seconds was chosen for the following two reasons: Firstly, from a behavioral perspective, participants were likely to experience the strongest embarrassed at the beginning of their speech. Over time, embarrassment may decrease due to habituation, resulting to greater variability across participants. By using a fixed length, the influence of habituation could be minimized. Secondly, from a machine learning perspective, fixed-length acoustic segments are more suitable for training predictive models in paralinguistic classification tasks.

2.5.1. Auxiliary emotional corpora

To train regression and classification models for the conceptual and dimensional description of embarrassment, two well-known emotional datasets in German were selected: the Berlin Emotional

Speech Database (EMO-DB) (Burkhardt et al., 2005) and the "Vera am Mittag" (VAM) dataset (Grimm et al., 2008).

The VAM corpora consist of 947 emotional German speech samples collected from 47 speakers (36 female). Speech segments were selected from 12 broadcasts of the TV talk show "Vera am Mittag" (in Engl.: "Vera at noon"). The weighted average values with evaluator weighted estimator (EWE) techniques of VAD emotional dimensions were used to train our regression models for the dimensional VAD representation of embarrassment (Grimm et al., 2007). Each speech sample in the database has EWE smoothed VAD dimensional labels in the range of [-1,+1]: valence (negative: -1 & positive: 1), activation (calm: -1 & excited: 1), and dominance (weak: -1 & strong: 1) (Kehrein, 2003).

EMO-DB covers 7 emotions (i.e., anger, joy, neutral, sadness, disgust, fear, and boredom). The corpus consists of 10 professional actors (5 female) speaking out 10 predefined phonetically balanced and emotionally neutral sentences. We utilized a subset of 493 utterances, which achieved naturalness and recognizability rates of 60% and 80%, respectively, as obtained during a perception test involving 20 subjects. This subset was used to train our emotion categories-based classification models.

2.5.2. Acoustic feature representations

Considering a comparable sparse amount of training samples for acoustic modeling, we decided to use knowledge-based handcrafted features and pre-trained data-driven feature representations. In the employed acoustic feature representations, we do not use phonetic-level information. Also, the handcrafted features are also interpretable. Thus, it can help in gaining insight into the acoustic variations related to speech production in the present study's context.

For the knowledge-based *handcrafted* feature representations (FRs), we use the (a) ComPaRE 2016 feature set provided by the openSMILE extraction tool (see Eyben et al. (2010) for more information), which has been studied for several paralinguistic speech processing tasks (Schuller et al., 2013). The ComPaRE set contains 6373 static turn-level features resulting from the computation of functionals (statistics) over low-level descriptor contours.

Considering top performance positions on challenge leaderboards and state-of-the-art neural embeddings for the Speech Emotion Recognition (SER) task from the Speech processing Universal PERformance Benchmark (SUPERB) challenge (Yang et al., 2021), we employed (b) Wav2Vec2 (WV2 EM) (Wagner et al., 2023), fine-tuned for dimensional SER on the MSP-Podcast (Lotfian & Busso, 2019) as emotion data-driven feature representation, and (c) WAVLM (large) (Chen et al., 2022) as general data-driven feature representation.

2.5.3. Machine learning experimental setup

To address our second goal (i.e., test how well our classification model could predict speaker state and distinguish between non-embarrassment and endorsement data samples), we used a speaker-independent experimental protocol, more precisely leave-one-speaker-out (LOSO) protocol. This

protocol simulates realistic conditions and maintains a good balance between training and testing subsets. In the LOSO protocol, we used speech samples from 32 speakers to train predictive models, and speech samples from the one remaining speaker were used for testing. This procedure was repeated 33 times, and predictions were aggregated into one file to estimate recognition rates. To measure recognition performance, we used unweighted average recall (UAR). During the experimental study, we used both three-class and two-class experimental settings. The three-class setting included pre-embarrassment, embarrassment, and post-embarrassment labels. In the two-class setting, we provided a selection of different combinations or grouped pre- and post-embarrassment classes into one class (i.e., non-embarrassment). Random Forest (RF) and Support Vector Machines (SVM) classifiers were used to train predictive models that take ComPaRE FR as input. The results obtained from the speaker-independent evaluation study for the second goal is presented in Section 3.2.

For the third goal (i.e., adopting a dimensional approach and mapping embarrassment onto the VAD dimension), we employed a cross-corpora approach, where first three independent Random Forest-based regressors that take as input ComPaRE FR and predict valence, arousal, and dominance values, respectively, were trained on the emotional speech samples of the VAM dataset. As mentioned earlier in Section 2.5.1, the prediction values for each of the VAD emotional dimensions are lying in the range [-1,1]. After training the regressors on the VAM database, we used the models to predict VAD labels and estimate a possible location of the embarrassment state in the circumplex model of affect introduced by Russell (1980). For this purpose, we used pre- and post-embarrassment samples as a reference and modeled possible shifts of numerical VAD labels in the context of these reference samples. The results of this experimental study are presented in Section 3.3.

For the fourth goal (i.e., comparing embarrassment to other emotions, following a categorical approach), as there are no categorical emotion labels associated with the data collected in our study, we again employed a cross-corpora approach. In this approach, a categorical emotion classifier was first built on EMO-DB corpus (see Section 2.5.1) and then the classifier outputs for our speech data were analyzed. Considering the different recording conditions and the phonetic differences between High German (in EMO-DB corpus) and Swiss German (in our data), for robust analysis, we took a multiple classifier/expert approach. As illustrated in Fig. 2, we trained three different categorical emotion classifier systems based on (a) handcrafted FR, (b) data-driven feature representation tuned for emotional analysis task (emotion data-driven FR), and (c) general data-driven FR on EMO-DB corpus. Each of the emotion classification systems were trained to predict class conditional probabilities of seven emotion classes (including neutral class). For our analysis, we passed each of our speech data samples through those three trained classifiers and took the average of the emotion class conditional probabilities estimated by the three classifiers. The resulting average emotion class conditional probabilities were then analyzed. As in the previous cases, we used pre- and post-embarrassment samples as reference samples and treated them as separate classes. The results of this experimental study are presented in Section 3.4.

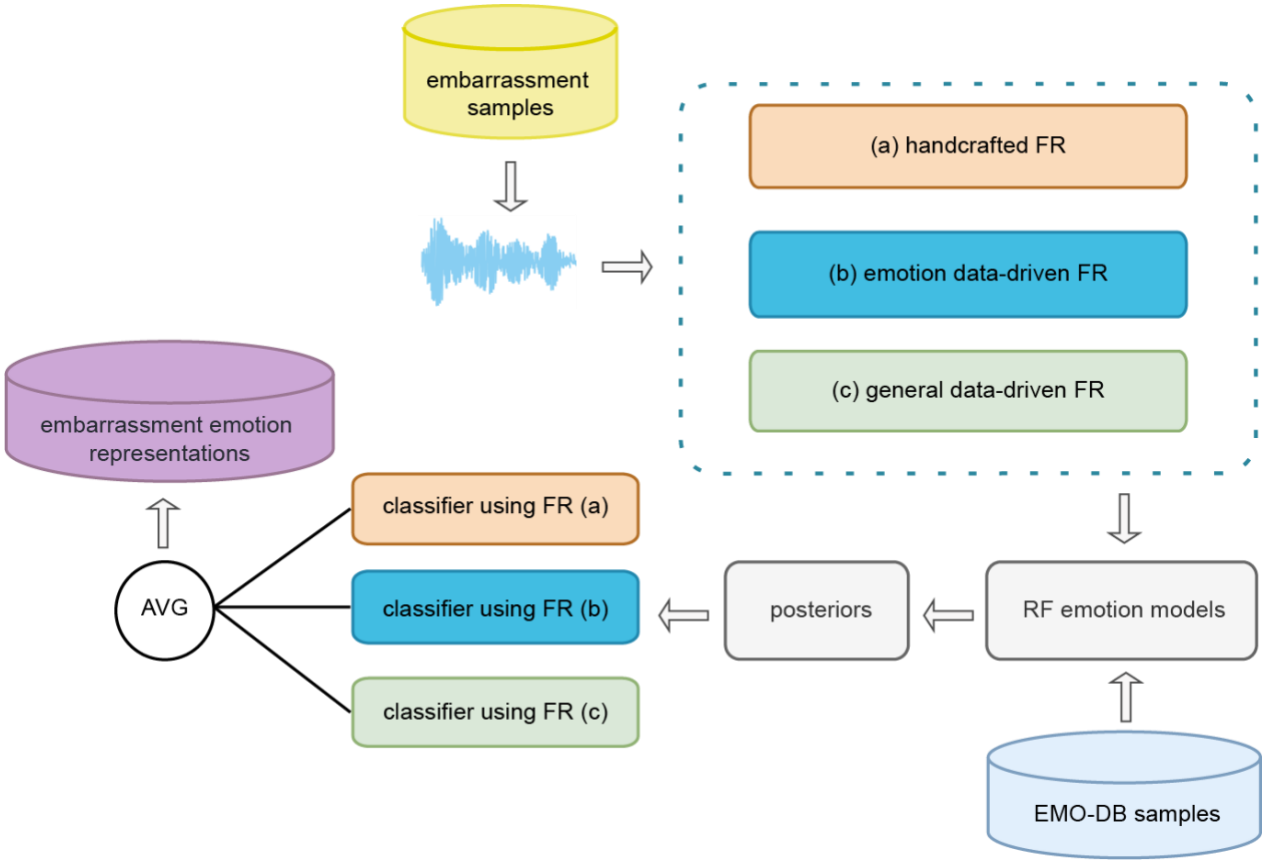


Figure 2 *Flow chart of the emotional categorical analysis with EMO-DB samples*

The presented pipeline has been used to represent pre-embarrassment, post-embarrassment, and embarrassment samples with aggregated class-conditional probabilities (referred to as posterior-based emotional representations) from categorical emotional models trained on EMO-DB. The proposed multiple classifier/expert approach is based on three expert systems trained with handcrafted FR, emotional data-driven FR, and general data-driven FR. FR = Feature Representation. RF = Random Forest. EMO-DB = a database (see Burkhardt et al. (2005) for more information). AVG = average. For (a), (b), (c), see Section 2.5.2.

2.5.4. Rate of speech analysis

To compute rate of speech, we adopted graphemes-per-second as a metric, because the grapheme-to-phoneme relation in German is shallower than the relation in English. In other words, the rate of speech in German is well represented by graphemes. We used OpenAI's Whisper model (https://huggingface.co/openai/whisper-large-v3) to transcribe the speech data. Then, we removed all punctuation marks and blank spaces in the transcription and counted the number of graphemes/characters. The grapheme count was finally divided that by 10 seconds to obtain the grapheme-per-second measure.

3. Results

2.4. SPS-6 & SIAS-6 results and embarrassment induction

Table 1

SPS-6 & SIAS-6 descriptive statistics

Questionnaires	*N*	*M*	*SD*	*Mdn*
SPS-6	33	2.52	3.01	2.00
SIAS-6	33	7.00	4.02	6.00
S (SPS–6 & SIAS–6)	33	4.76	4.18	4.00

Note. SPS-6: Short form Social Phobia Scale; SIAS: Short form Social Interaction Anxiety Scale.

Table 1 shows the results from the questionnaires SPS-6 and SIAS-6. 16 Participants were above the SIAS cut-off (i.e., sum SIAS ≥ 7), 17 participants were above the SPS cut-off (i.e., sum SPS ≥ 2), and 9 participants were above both cut-offs (i.e., sum SIAS ≥ 7 and sum SPS ≥ 2). Therefore, according to the cut-offs, 27.3% (9 / 33) showed clinically relevant SA symptoms in both questionnaires.

The first goal of the study was to verify the embarrassment induction. The Shapiro-tests for the differences of the 3 VAS scores (VAS 1 – VAS 2, VAS 1 – VAS 3, VAS 2 – VAS 3) (see Fig. 1) between t0 and t3 showed a significant result for the difference VAS 2 – VAS 3 ($p = .023$). Therefore, all further calculations were made with non-parametric tests. Multiple pairwise Wilcoxon tests showed a significant difference between the VAS 2 ($M = 2.52$, $SD = 2.09$) and VAS 3 ($M = 4.40$, $SD = 2.38$) score ($V = 551$, $p < .001$). This means that the participants got significantly more embarrassed after reading the story aloud to the conductor than after only writing the story. This can also be observed visually in Fig. 3 for 31 out of 33 participants. The other differences were not significant (VAS1 – VAS2, $p = .13$, VAS 1 – VAS 3, $p = .20$). To account for the gender disbalance ($n = 26$ females) in the significant differences between VAS 2 and VAS 3, we used LMMs. First, the time only model was calculated with time (VAS 2, VAS 3) as a predictor and ID as a random effect (i.e., accounting for the fact that all participants have different baseline values). As expected, the time only model was significant ($b = 1.88$, $SE = 0.26$, $t(32) = 7.22$, $p < .001$). Then the covariate gender was added to the model (time*gender). The interaction between time and gender was not significant ($b = 0.57$, $SE = 0.64$, $t(31) = 7.22$, $p = .89$), implying that the effect of time on the VAS values did not differ by gender.

Spearman correlations were calculated between the SIAS-6 and SPS-6 total scores, the total sum of both questionnaires and the VAS 2 – VAS 3 difference score. There was a significant negative correlation between the difference score and the total sum score of the SPS-6 and SIAS-6 ($r_s = – 0.36$, $p = .037$). This means that the higher SA is (here defined by the overall sum of both questionnaires), the more embarrassed a person became after reading the story out loud.

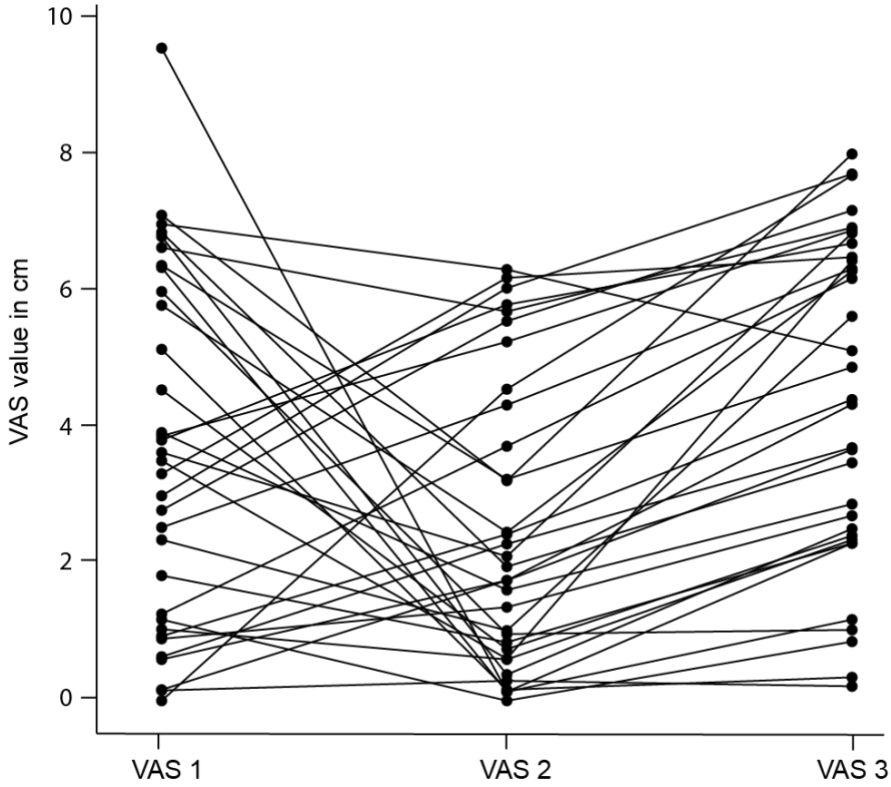


Figure 3 *VAS values connected over all time-points for each participant*
VAS-1 measured embarrassment between telling a story based on a neutral picture and writing about an embarrassing experience. VAS-2 measured embarrassment between writing about an embarrassing experience and reading the embarrassing story. VAS-3 measured embarrassment between reading the embarrassing story and telling a story based on another neutral picture. There was a significant difference between VAS-2 and VAS-3 ($V = 551$, $p < .001$), where 31 out of 33 got significantly more embarrassed after reading the embarrassing story.

2.5. Prediction of embarrassment based on acoustic features

The second goal of the study was to test the possible prediction performance of machine learning models applied to predict embarrassment. During our preliminary analysis of embarrassment prediction, we used three-class and two-class configurations for classification experimental setups. In the two-class settings, in addition to the non-embarrassment class concept (see section 2.5.3), we used a pre- vs. post-embarrassment configuration. Additionally, considering the scarcity of collected embarrassment data and the need for better interpretability, we used only knowledge-based acoustic features - the ComPaRE (see section 2.5.2.) feature set.

Table 2

Prediction performance UAR of RF and SVM for various combinations of pre-, embarrassment- and post-sets

	pre vs. emb	post vs.emb	pre vs. post	(pre+post) vs. emb	pre vs. emb vs. post
RF	**0.848**	**0.833**	0.621	**0.803**	0.636
SVM	**0.864**	**0.818**	0.591	**0.818**	0.596

Note. UAR = Unweighted Average Recall. RF = Random Forest. SVM = Support Vector Machine. emb = embarrassment.

The bold print indicates trends for a higher prediction accuracy for RF and SVM.

Table 2 shows the prediction performance (i.e., the probability of predicting the correct set) of RF and SVM. The best performance was found for predicting the pre-set vs. the embarrassment-set (pre vs. emb), where the right prediction for the respective set was made in 84.8% of the cases for RF and 86.4% for the SVM (since it was a two-class configuration, a performance by chance would be 0.5. The results here indicated a very good prediction performance). The second- and third-best performance was found for the prediction of the post vs. embarrassment (post vs. emb) and for the pre- and post-set combined vs. the embarrassment-set ((pre+post) vs. emb). The pre- vs. post-comparison was slightly above chance and could not be discriminated satisfactorily. In the case of the three-class configuration (pre vs. emb. vs. post), the main misclassification confusion was observed between the pre- and post-embarrassment classes: The pre- and post-state showed a tendency to have similar characteristics since their predictive performance was slightly above chance (see section 3.4. for quantitative proof of similarity).

Supplementary Table 2 shows the top-ranked features for each combination of prediction from Table 2. Feature ranking was conducted based on RF feature importance rates. The glossary in Supplementary Table 1 can be consulted for more information on the features. A more detailed description of acoustic features and their corresponding mathematical models can also be found in the paper of Eyben (2016). Since the best prediction performances were found for pre vs. embarrassment and (pre+post) vs. embarrassment, only their top-ranking features were examined in more detail. For selected set configurations (pre vs. embarrassment and (pre+post) vs. embarrassment), mainly vocal tract-related dynamic features are presented in the top-ranking list (i.e., the most discriminative features). Most of the top-ranked acoustic features for the (pre+post) vs. embarrassment task represent temporal dynamics of envelop of short-term spectrum which relates to modulation frequency information, indicating changes in energy in frequency bands across time due to speech articulation (movement of jaw, tongue and lips) (Hermansky, 1998).

One way to ascertain whether modulation frequency information is indeed playing a central role here is comparing speaking rate, as changes in rate of speech leads to changes in modulation spectrum (Zhang et al. 2018). More precisely, peak modulation frequency tend to correlate with

speech rate. We used the speech recognition system to transcribe the pre, post, and embarrassment speech and calculated speaking rate in terms of graphemes-per-second for each condition, respectively, each time point. As mentioned earlier, graphemes-per-second is a good indicator of speaking rate in our case, as in German the grapheme-to-phoneme relation is shallow when compared to English. Figure 4 shows the distribution of grapheme-per-second for 33 speech samples (33 speakers) for each condition. We found that the speaking rate was higher for embarrassment (mean: 11.92) than for pre-embarrassment (mean: 7.67) and post-embarrassment (mean: 8.16), suggesting that modulation frequency information indeed helps to distinguish state of embarrassment.

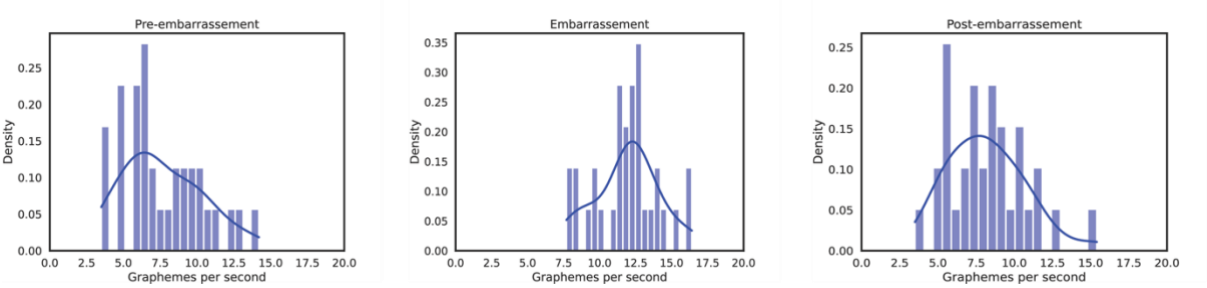


Figure 4 Histogram and Gaussian kernel density estimation for grapheme-per-second rates for pre-embarrassment (left), embarrassment (middle), and post-embarrassment (right).

2.6. Dimensional modeling of embarrassment using the dimensions valence, arousal, and dominance (VAD) from VAM

In order to find the possible location of the embarrassment state in the circumplex model of affect introduced by Russell (1980), we decided to process the embarrassment data using pre-trained VAD regressors. Fig. 5a – 5c shows a density map for the predicted valence, arousal, and dominance levels for pre vs. embarrassment and post vs. embarrassment states for all 33 participants (represented as dots). Considering the sparsity of the collected data, we used knowledge-based feature representations (i.e., ComParE) for acoustic modeling in our regression models.

Predictive models were trained on the VAM database speech samples. The regressors trained on VAD emotional dimensionalities were used to predict VAD levels for pre-, post-, and embarrassment samples. To evaluate changes in VAD levels for embarrassment phenomena, we mapped the obtained predictions in 2D plots presented in Fig. 5a – 5c.

Arousal (Fig. 5b) & dominance (Fig. 5c) seem to indicate embarrassment phenomena (more points are on the embarrassment side from the diagonal), while valence (Fig. 5a) does not show any predictive value (dots are quite evenly distributed around the diagonal) and should not be used without further linguistic post-hoc tests.

Nevertheless, none of the VAD dimensions could differentiate enough between pre- vs. embarrassment and between post- vs. embarrassment states. The obtained results showed that embarrassment phenomena could be characterized by comparably complex changes in VAD levels compared to pre vs. embarrassment and post vs. embarrassment states.

Supplementary Table 3, the mutual information (MI) between the regression task for the VAD emotional dimensions on VAM samples and the embarrassment detection task can be found. Selected acoustic features have high MI rates for both tasks (embarrassment detection and VAM-based emotion modeling). As in our previous feature ranking analysis based on RF feature importance, vocal tract-related modulation features are influential in the selected features with high shared MI. Additionally, Supplementary Table 4 shows MI between EMO-DB and embarrassment.

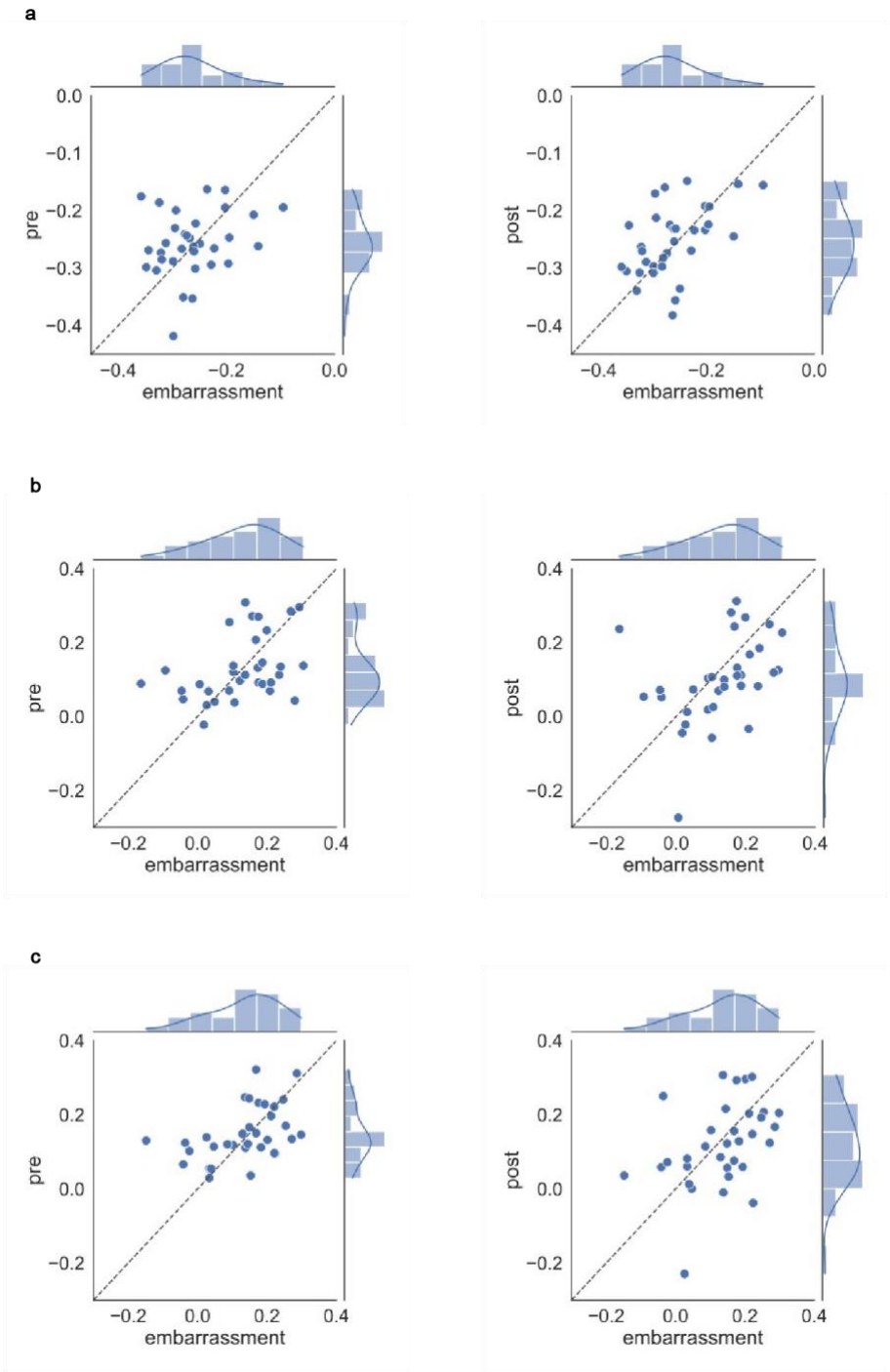


Figure 5 *Predicted levels of (a) valence, (b) arousal, and (c) dominance*

Scatter and density plot for pre vs. embarrassment and post vs. embarrassment states. The 0-point indicates a neutral state.

2.7. Categorical modeling of embarrassment using the 7 emotions from EMO-DB

For the classification task, the database EMO-DB was used to train the classifiers for acoustic emotion categories. The database contains emotional speech data from 5 female and 5 male actors speaking in German. They portrayed the following 7 emotions: happiness, sadness, disgust, fear, boredom, anger, and neutral state, to which our embarrassment data set was compared.

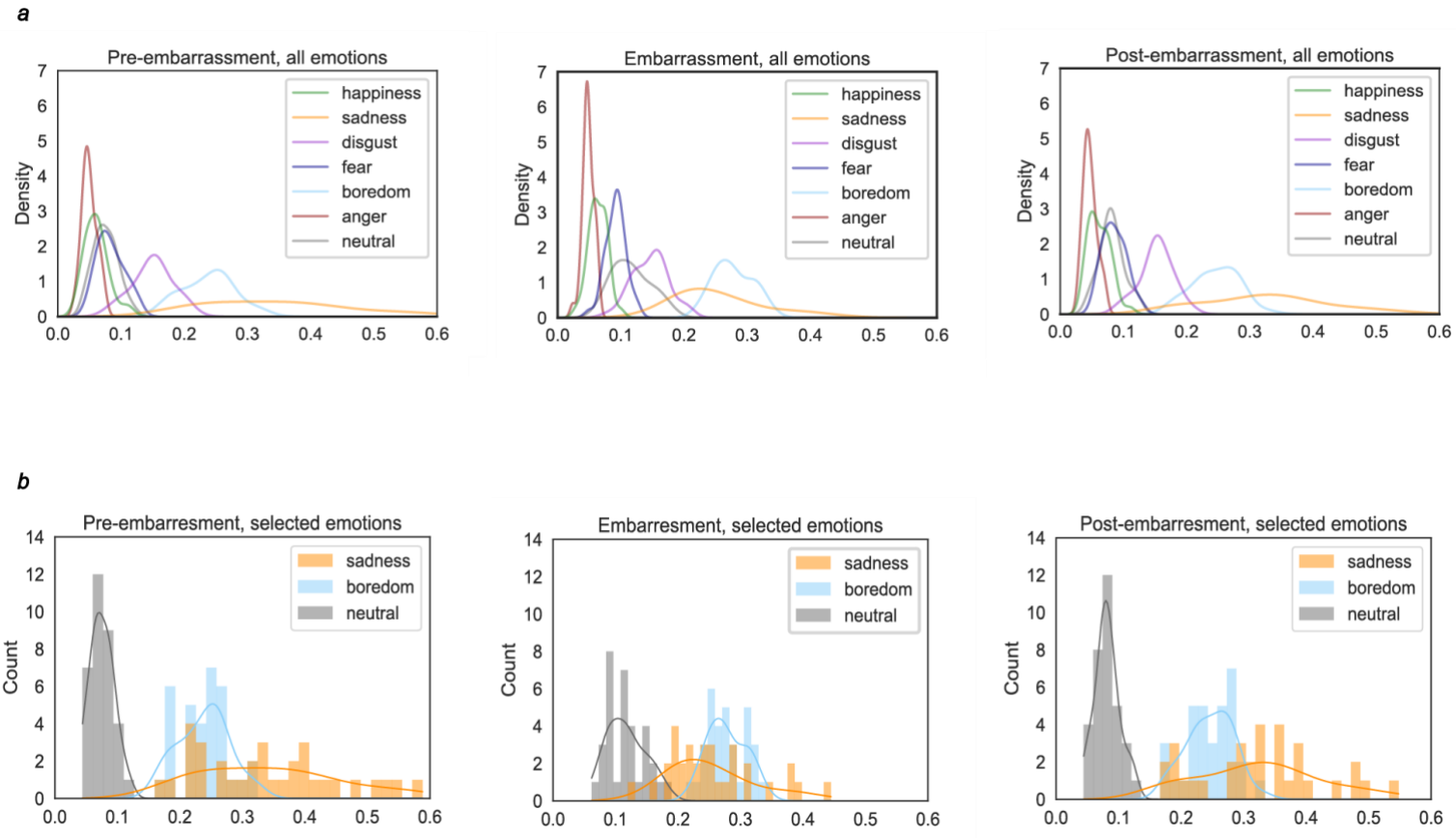


Figure 6 *The density x-axis shows the average probability for the observed emotional class over all participants (sum of all probabilities is 1). Count = number of participants.*
(a) The line plots show the distribution approximations of posterior-based seven emotional states for pre-embarrassment (left), embarrassment samples (middle), and post-embarrassment (right). (b) Histogram and line plots of selected three emotional states, showing significant changes between pre-embarrassment and embarrassment, for pre-embarrassment (left), embarrassment samples (middle), and post-embarrassment (right).

Fig. 6a shows the probability density functions (PDF) for posterior probability of all EMO-BD emotions over the 3 audio time points (pre, embarrassment, post). In other words, it shows the probability of how much each emotion is represented in our embarrassment sample for each time point. There is a skewed distribution for each time point and a clear distribution change in posterior

probabilities of the emotions boredom, sadness, and neutral state over the time points. The most indicative changes in these 3 emotions are plotted in Fig. 6b. The histogram plots show average posteriors for sadness, boredom, and neutral state.

For quantitative proof of the significance of emotional posterior modeling, we used a t-test to evaluate the average posteriors of each emotional state. The above-mentioned indicative changes were also observed in Table 3 with the mean posterior values of each emotion: The obtained p-values for sadness, boredom, and neutral state posteriors are highly significant for both pre vs. embarrassment comparisons and the post vs. embarrassment comparisons. At the same time, the p-values for pre vs. post comparisons for all emotional states are not significant. With that, it could be shown that pre- and post-states show indeed similar characteristics, which was already assumed in Table 2.

In conclusion, being embarrassed seems to shift the voice to less sadness but more boredom and neutral state-related characteristics. But this shift is only temporary and disappears after a while (at post).

Table 3

EMODB's emotional representations in mean posterior values over all subjects (N = 33) and the p-value for the t-test for different pairs of mean

emotion	mean			*p*-value		
	pre	emb.	post	pre vs. emb.	post vs. emb.	pre vs. post
happiness	0.06	0.06	0.06	.503	.652	.818
sadness	0.34	0.25	0.33	< .001	.001	.506
disgust	0.15	0.15	0.15	.699	.397	.697
fear	0.08	0.09	0.09	.079	.146	.685
boredom	0.24	0.28	0.25	< .001	< .001	.402
anger	0.05	0.05	0.05	.574	.582	.317
neutral	0.08	0.12	0.08	< .001	< .001	.200

Note. *p*-values were retrieved from t-tests.

Supplementary Table 3 and 4 shows the MI for categorical emotion classification on EMO-DB and the dimensional modeling. We used the complete emotion set presented in the EMO-DB database and binary classification mapping configurations: high/low and arousal/valence. Results in Supplementary Table 3 and 4 show that the embarrassment phenomenon could cause changes in both emotional dimensions, such as arousal and valence. Hence, the highest shared MI in the emotion classification task with the 7 emotional classes experimental setup can be observed.

4. Discussion

This study aimed to systematically examine embarrassment categorically and dimensionally with an interdisciplinary self-report and machine-learning approach.

In the first step, the induction was tested with subjective self-reported measures (i.e., SPS-6, SIAS-6, VAS), and it was shown that participants got significantly more embarrassed after reading the story aloud than only writing it and that the embarrassment induction was successful. This was particularly the case for people who reported higher SA symptoms before the experiment.

In the second step, the robustness of the embarrassment data was tested based on acoustic features. The robustness of the data set was shown with high prediction performance trends for the comparisons pre vs. embarrassment and pre- and post-set combined vs. the embarrassment-set, with SVM showing slightly better predictions than RF. The pre- vs. post-prediction performance was slightly above chance; one hypothesis is that people return to the initial state so that pre and post cannot be distinguished from each other.

In the third step, embarrassment was examined dimensionally in the VAD dimensions. It was shown that there was a shift in the dimensions of arousal and dominance, while there was none for valence. In order to obtain an applicable prediction with valence, one would have to adapt linguistic-based techniques since embarrassment is a complex emotion, and valence is dependent on different factors (as is natural language processing in general domain-dependent). A participant could, for example, try to mask their embarrassment by changing the voice so that valence would be perceived as positive, even though the content of the audio data is negative. These results suggest either that embarrassment might not be described with simply VAD dimensions since there were no differences found (i.e., distribution graphs were comparable and comparable number of points at each side of the diagonal), or that the model itself is too simple to capture a complex emotion like embarrassment.

In the fourth step, embarrassment was examined categorically by comparing it to seven other emotional states. It was shown that the state of embarrassment had characteristics of less sadness and more boredom and a neutral state in comparison to the pre- and post-state. On the other hand, the pre- and post-states did not show any significant differences, implying that the participants returned, after feeling embarrassed, to a pre-similar state. It additionally shows that our chosen method of embarrassment induction only leads to a temporary heightened embarrassment state and might, therefore, be safe to use with subclinical or clinical samples. This generally suggests that embarrassment is indeed a complex emotion since we observed a shift in multiple emotions (see Fig. 6a). There might also be cultural differences that might explain why there is a higher posterior probability of sadness at pre (one would rather expect a neutral state) or due to the emotions being acted and therefore over-exaggerated.

The behavioral data has shown that reading the story out loud induced embarrassment (in comparison to writing the embarrassing story down). The advantage of this individualized embarrassment induction compared to other induction methods, like singing or holding a presentation,

is that writing a personal story and reading it out loud makes sure that it is embarrassing for each person while singing or holding a presentation might even be pleasant for certain people. Indeed, 31 out of 33 participants reported increased embarrassment after speaking about their embarrassing experience. In addition, the machine learning models found shifts for embarrassment in acoustic features, arousal and dominance on the VAD dimensionality, and emotional characteristics of sadness, boredom and neutral state. One explanation is that the articulation changes when people get embarrassed. They might try to become more in control of their voice and to articulate more precisely, hence the change in dominance and the heightened boredom and neutral state (which could also explain the decrease in sadness). Another explanation could be that participants emotionally distanced themselves from the situation due to embarrassment and, therefore, the characteristics of the voice changes. It would have been interesting to have an additional VAS at the end of t3 to test whether the subjective feeling of embarrassment would coincide with the return of the voice to the pre-similar state.

In general, embarrassment can be predicted with high accuracy from the voice only, even though we had a small sample and fixed durations of voice samples (i.e., 10 seconds, respectively, the same amount of information for each participant), which shows that our fundamental study set-up is valid. It also shows the overall good robustness of the data set, that the classification models trained on EMO-DB can be used to understand the embarrassment phenomenon, and that the t-tests provide an additional source of information for the classification models.

The top-ranking FRs in acoustic analysis has identified two aspects in the LLDs that stand out. First, the FRs are either based on auditory spectrum or based on MFCCs. Both these FRs tend to parameterize the spectral envelop of the short-term spectrum, which carries information more related to the acoustic changes due to the change in the shape of the vocal tract system when producing speech (Childers, 1975; Hermansky, 1990; Makhoul, 1975). Second, in all the cases the LLDs were capturing the temporal dynamics of the envelop of short-term spectrum using either delta feature computation (denoted as “de” in the FR definition) or RASTA processing (denoted as “Rasta” in FR definition). Taken together, these two aspects indicate that the classification systems are focusing on modulation frequency information (Hermansky, 1998). Consistent with the changes of modulation frequency information, we further found that rate of speech was increased during embarrassment.

Another advantage of our setting is the high degree of naturalness, respectively, the ecological validity since emotion-related studies often work with acted respectively simulated emotions (Drahota et al., 2008; El Ayadi et al., 2011).

Our study shows the following limitations: Firstly, no control respectively comparable group existed. Due to the within-study design, every participant is a control for themselves for intra-individual differences, but there is no comparison for inter-individual differences for time-effects. One could, for example, add another group that writes about a neutral, instead of an embarrassing, story. Secondly, even though we were able to find robust effects, the number of participants (N = 33) in the study was rather small. Thirdly, the majority of the participants were female (n = 26). Therefore, we were not able

to control for gender differences. However, the LMMs implied that there was still a significant difference for the behavioral data when controlling for gender. Fourthly, we did not control for a linguistic bias. This was already acknowledged in the results part, where valence was not further considered. Fifthly, as no data set exists which could be used to train our models in Swiss German, they were trained in High German. Lastly, carry-over effects due to part 1 of the study cannot be ruled out. It could be, for example, that participants were more sad or bored due to part 1. Sixthly, the participants knew that they were recorded, which could have influenced the expression of embarrassment. However, it remains an open question, if there was any influence, whether it led to a more controlled or exaggerated expression of embarrassment.

By examining paralinguistic voice parameters associated with embarrassment, this paper contributes to multiple research gaps at once. Due to the especially big gap in voice analyses for embarrassment, this paper focused mostly on the fundamental description and classification of embarrassment with voice parameters rather than the association between SA and embarrassment. There was a study conducted by Simon-Thomas et al. (2009) which examined vocal bursts, however, to our knowledge, this is the first study using acoustic analyses to study embarrassment in a multidisciplinary approach.

A future paper will address this association between a healthy sample (control group) and an SAD sample. The SAD sample would be particularly interesting to investigate since we found a correlation between SA and the height of embarrassment, assuming at least more robust and bigger effects. From a clinical point of view, it would additionally be interesting to ask the participants with SAD in anxiety inducing social situations whether they notice changes in their voice and if they think that others might notice it, since one of the key processes in people with SAD in anxiety-inducing social situations is the heightened inward attention on themselves and their interoceptive information and with that the fear that others might notice their anxiety (Clark & Wells, 1995). Additionally, due to the potential negative consequences of experiencing repeated embarrassment (cf. Bas-Hoogendam et al., 2018) as well as embarrassment being part of psychiatric diagnoses (e.g., in SAD, Taijin kyofusho, agoraphobia) or consequences of psychiatric disorder (e.g., in Trichotillomania, excoriation disorder) (Diagnostic and Statistical Manual of Mental Disorders 5th edition (text rev.); American Psychiatric Association, 2022). Future intervention studies could focus on how to cope with embarrassment. For example, by providing psychoeducation on embarrassment (cf. Dijk et al., 2012) or practicing cognitive restructuring (both typical components of a cognitive behavioral therapy) for embarrassing and SA inducing situations. Furthermore, distinguishing embarrassment from other emotions remains an open question. From a behavioral perspective, we tried to ensure that participants reported embarrassment rather than shame by explaining the differences between the two emotions and asking specifically on the VAS scales about how embarrassed they felt at the moment. From an acoustic perspective, we used EMO-DB to differentiate between embarrassment and other emotions, but some confounding factors remain, such as dialect differences (EMO-DB used high German vs. our sample was in Swiss German).

We assumed that acoustic patterns for embarrassment and reference speech are highly speaker-specific (as is typical for most paralinguistic concepts); hence, the reported UAR rates for the speaker-independent evaluation protocol appeared promising. Given similarities between embarrassment and SA (cf. Hofmann, 2006), the current framework is also promising to detect SA in clinical settings. Future studies could apply more advanced speaker-adaptive techniques, incorporating precise acoustic analysis, to further improve classification performance. Lastly, in the light of multidisciplinary approaches, it would also be interesting to compare acoustic indicators with additional (neuro)physiological indicators.

In conclusion, we were able to describe the characteristic way of speaking for people in a state of embarrassment. With a multidisciplinary approach, we have established an effective behavioral paradigm to induce embarrassment and characterize participants' response to the embarrassment induction in acoustic characteristics. Those identified characteristics could be used for the assessment of SA and SAD.

References

American Psychiatric Association (Ed.). (2022). Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR).

Bas-Hoogendam, J. M., Steenbergen, H. van, Wee, N. J. A. van der, & Westenberg, P. M. (2018). Not intended, still embarrassed: Social anxiety is related to increased levels of embarrassment in response to unintentional social norm violations. European Psychiatry, 52, 15–21. https://doi.org/10.1016/j.eurpsy.2018.03.002

Bastin, C., Harrison, B. J., Davey, C. G., Moll, J., & Whittle, S. (2016). Feelings of shame, embarrassment and guilt and their neural correlates: A systematic review. *Neuroscience & Biobehavioral Reviews*, *71*, 455–471. https://doi.org/10.1016/j.neubiorev.2016.09.019

Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., & Weiss, B. (2005). A database of German emotional speech. *Interspeech 2005*, 1517–1520. https://doi.org/10.21437/Interspeech.2005-446

Chen, S., Wang, C., Chen, Z., Wu, Y., Liu, S., Chen, Z., Li, J., Kanda, N., Yoshioka, T., Xiao, X., Wu, J., Zhou, L., Ren, S., Qian, Y., Qian, Y., Wu, J., Zeng, M., Yu, X., & Wei, F. (2022). WavLM: Large-Scale Self-Supervised Pre-Training for Full Stack Speech Processing. *IEEE Journal of Selected Topics in Signal Processing*, *16*(6), 1505–1518. https://doi.org/10.1109/JSTSP.2022.3188113

Childers, D. G., Skinner, D. P., & Kemerait, R. C. (1977). The cepstrum: A guide to processing. Proceedings of the IEEE, 65(10), 1428–1443. https://doi.org/10.1109/PROC.1977.10747

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. Hope, & F. Schneider (Eds.), *Social Phobia: Diagnosis, Assessment, and Treatment (pp. 69–93)*. Guilford Press.

Delgado, D. A., Lambert, B. S., Boutris, N., McCulloch, P. C., Robbins, A. B., Moreno, M. R., & Harris, J. D. (2018). Validation of Digital Visual Analog Scale Pain Scoring With a

Traditional Paper-based Visual Analog Scale in Adults. *JAAOS: Global Research and Reviews*, *2*(3), e088. https://doi.org/10.5435/JAAOSGlobal-D-17-00088

Devillers, L., & Vidrascu, L. (2007). Real-Life Emotion Recognition in Speech. In C. Müller (Ed.), *Speaker Classification II: Selected Projects* (pp. 34–42). Springer. https://doi.org/10.1007/978-3-540-74122-0_4

Dijk, C., Buwalda, F. M., & de Jong, P. J. (2012). Dealing with Fear of Blushing: A Psychoeducational Group Intervention for Fear of Blushing. Clinical Psychology & Psychotherapy, 19(6), 481–487. https://doi.org/10.1002/cpp.764

Drahota, A., Costall, A., & Reddy, V. (2008). The vocal communication of different kinds of smile. *Speech Communication*, *50*(4), 278–287. https://doi.org/10.1016/j.specom.2007.10.001

El Ayadi, M., Kamel, M. S., & Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, *44*(3), 572–587. https://doi.org/10.1016/j.patcog.2010.09.020

Eyben, F. (2016). *Real-time Speech and Music Classification by Large Audio Feature Space Extraction*. Springer International Publishing. https://doi.org/10.1007/978-3-319-27299-3

Eyben, F., Wöllmer, M., & Schuller, B. (2010). Opensmile: The munich versatile and fast open-source audio feature extractor. *Proceedings of the 18th ACM International Conference on Multimedia*, 1459–1462. https://doi.org/10.1145/1873951.1874246

Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., André, E., Busso, C., ... & Truong, K. P. (2015). The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Transactions on Affective Computing*, 7(2), 190–202. 10.1109/TAFFC.2015.2457417

Grimm, M., Kroschel, K., Mower, E., & Narayanan, S. (2007). Primitives-based evaluation and estimation of emotions in speech. *Speech Communication*, *49*(10), 787–800. https://doi.org/10.1016/j.specom.2007.01.010

Grimm, M., Kroschel, K., & Narayanan, S. (2008). The Vera am Mittag German audio-visual emotional speech database. *2008 IEEE International Conference on Multimedia and Expo*, 865–868. https://doi.org/10.1109/ICME.2008.4607572

Hermansky, H. (1990). Perceptual linear predictive (PLP) analysis of speech. *The Journal of the Acoustical Society of America*, 87(4), 1738-1752.

Hermansky, H. (1998). Should recognizers have ears? Speech Communication, 25(1–3), 3 27. https://doi.org/10.1016/S0167-6393(98)00027-2

Hofmann, S. G., Moscovitch, D. A., & Kim, H.-J. (2006). Autonomic correlates of social anxiety and embarrassment in shy and non-shy individuals. *International Journal of Psychophysiology*, *61*(2), 134–142. https://doi.org/10.1016/j.ijpsycho.2005.09.003

Juslin, P. N. (2018). The Mirror to Our Soul? Comparisons of Spontaneous and Posed Vocal Expression of Emotion. *J Nonverbal Behav*, 40.

Kadali, D. B., & Mittal, V. K. (2020). Studies on Paralinguistic Sounds, Emotional Speech and Expressive Voices. *Workshop on Speech, Music and Mind (SMM 2020)*, 11–15. https://doi.org/10.21437/SMM.2020-3

Kehrein, R. (2003). Die Prosodie authentischer Emotionen. *Sprache · Stimme · Gehör*, *27*(02), 55–61. https://doi.org/10.1055/s-2003-40251

Keltner, D. (2005). Signs of Appeasement: Evidence for the Distinct Displays of Embarrassment, Amusement, and Shame. In P. Ekman & E. L. Rosenberg (Eds.), *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195179644.003.0007

Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional Expression: Advances in Basic Emotion Theory. *Journal of Nonverbal Behavior*, *43*(2), 133–160. https://doi.org/10.1007/s10919-019-00293-3

Leary, M. R., & Hoyle, R. H. (2013). *Handbook of Individual Differences in Social Behavior*. Guilford Publications.

Lotfian, R., & Busso, C. (2019). Building Naturalistic Emotionally Balanced Speech Corpus by Retrieving Emotional Speech from Existing Podcast Recordings. *IEEE Transactions on Affective Computing*, *10*(4), 471–483. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2017.2736999

Makhoul, J. (1975). Linear prediction: A tutorial review. Proceedings of the IEEE, 63(4), 561 580. https://doi.org/10.1109/PROC.1975.9807

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, 16.

Miller, R. S. (2009). Social anxiousness, shyness, and embarrassability. In *Handbook of individual differences in social behavior* (pp. 176–191). The Guilford Press.

Miller, R. S. (2012). The interactive origins and outcomes of embarrassment. In W. R. Crozier & P. J. de Jong (Eds.), *The Psychological Significance of the Blush* (1st ed., pp. 185–202). Cambridge University Press. https://doi.org/10.1017/CBO9781139012850.013

Miller, R. S., & Tangney, J. P. (1994). Differentiating Embarrassment and Shame. *Journal of Social and Clinical Psychology*, *13*(3), 273–287.

Mota, N. B., Copelli, M., & Ribeiro, S. (2017). Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance. *Npj Schizophrenia*, *3*(1), 18. https://doi.org/10.1038/s41537-017-0019-3

Müller-Pinzler, L., Paulus, F. M., Stemmler, G., & Krach, S. (2012). Increased autonomic activation in vicarious embarrassment. *International Journal of Psychophysiology*, *86*(1), 74–82. https://doi.org/10.1016/j.ijpsycho.2012.07.183

Patel, S., Scherer, K. R., Björkner, E., & Sundberg, J. (2011). Mapping emotions into acoustic space: The role of voice production. *Biological Psychology*, *87*(1), 93–98. https://doi.org/10.1016/j.biopsycho.2011.02.010

Peters, L., Sunderland, M., Andrews, G., Rapee, R. M., & Mattick, R. P. (2012). Development of a short form Social Interaction Anxiety (SIAS) and Social Phobia Scale (SPS) using nonparametric item response theory: The SIAS-6 and the SPS-6. *Psychological Assessment*, *24*(1), 66–76. https://doi.org/10.1037/a0024544

Posit team. (2024). *RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA.* (2023.12.1.402) [Computer software]. http://www.posit.co/.

Rozen, N., & Aderka, I. M. (2023). Emotions in social anxiety disorder: A review. *Journal of Anxiety Disorders*, *95*, 102696. https://doi.org/10.1016/j.janxdis.2023.102696

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010). Perceptual Cues in Nonverbal Vocal Expressions of Emotion. *Quarterly Journal of Experimental Psychology*, *63*(11), 2251–2272. https://doi.org/10.1080/17470211003721642

Schuller, B., Steidl, S., Batliner, A., Vinciarelli, A., Scherer, K., Ringeval, F., Chetouani, M., Weninger, F., Eyben, F., Marchi, E., Mortillaro, M., Salamin, H., Polychroniou, A., Valente, F., & Kim, S. (2013). *The INTERSPEECH 2013 computational paralinguistics challenge: Social signals, conflict, emotion, autism*. 148–152. https://doi.org/10.21437/Interspeech.2013-56

Simon-Thomas, E. R., Keltner, D. J., Sauter, D., Sinicropi-Yao, L., & Abramson, A. (2009). The voice conveys specific emotions: Evidence from vocal burst displays. *Emotion*, *9*(6), 838–846. https://doi.org/10.1037/a0017810

Stocks, E. L., Lishner, D. A., Waits, B. L., & Downum, E. M. (2011). I'm Embarrassed for You: The Effect of Valuing and Perspective Taking on Empathic Embarrassment and Empathic Concern: EMPATHIC EMBARRASSMENT. *Journal of Applied Social Psychology*, *41*(1), 1–26. https://doi.org/10.1111/j.1559-1816.2010.00699.x

Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, *70*, 1256–1269. https://doi.org/10.1037/0022-3514.70.6.1256

Vlasenko, B., Vyas, S., & Magimai-Doss, M. (2024). Comparing data-Driven and Handcrafted Features for Dimensional Emotion Recognition. *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 11841–11845. https://doi.org/10.1109/ICASSP48485.2024.10446134

Wagner, J., Triantafyllopoulos, A., Wierstorf, H., Schmitt, M., Burkhardt, F., Eyben, F., & Schuller, B. W. (2023). Dawn of the Transformer Era in Speech Emotion Recognition: Closing the Valence Gap. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(9), 10745–10759. https://doi.org/10.1109/TPAMI.2023.3263585

Weeks, J. W., Lee, C.-Y., Reilly, A. R., Howell, A. N., France, C., Kowalsky, J. M., & Bush, A. (2012). “The Sound of Fear”: Assessing vocal fundamental frequency as a physiological indicator of social anxiety disorder. *Journal of Anxiety Disorders*, *26*(8), 811–822. https://doi.org/10.1016/j.janxdis.2012.07.005

Yang, S., Chi, P.-H., Chuang, Y.-S., Lai, C.-I. J., Lakhotia, K., Lin, Y. Y., Liu, A. T., Shi, J., Chang, X., Lin, G.-T., Huang, T.-H., Tseng, W.-C., Lee, K., Liu, D.-R., Huang, Z., Dong, S., Li, S.-W., Watanabe, S., Mohamed, A., & Lee, H. (2021). SUPERB: Speech Processing Universal PERformance Benchmark. *Interspeech 2021*, 1194–1198. https://doi.org/10.21437/Interspeech.2021-1775

Zhang, Y., Zou, J., & Ding, N. (2023) Acoustic correlates of the syllabic rhythm of speech: Modulation spectrum or local features of the temporal envelope. *Neuroscience and Biobehavioral Reviews*, 147, 1-12. https://doi.org/10.1016/j.neubiorev.2023.105111

Statements

Acknowledgments

We thank Lara Dinner for helping us with the data acquisition. Furthermore, we want to thank all our brave participants, without whom this study would not have been possible.

Competing Interests:

none.

Funding:

UPD university-based fund.

Author's Contributions

The conceptualization of this project was done by D.S. and Y.M.
The data acquisition for the sample was done by D.S., Y.M., and M.S.
The calculations of the behavioral data were done by D.S. with the support of Y.M.
The calculations of the machine learning data were done by B.V. with the support of M.MD.
The manuscript was written by D.S. with the support of Y.M. and B.V.
All authors read and approved the final manuscript.

Data Availability Statement

The behavioral and machine learning data used to support the findings of this study are available upon request. Contact D.S. for the behavioral data and B.V. for the machine learning data. However, voice data is available only after the approval of the local ethics committee due to the privacy protection of the Human Research Act in Switzerland.

Ethics Declarations

The ethics committee of the faculty of human sciences at the University of Bern approved the study protocol (2020-06-00004). All participants gave written informed consent to participate in the study. We hereby confirm that the experiment was performed in accordance with the relevant guidelines and regulations according to the Declaration of Helsinki.

Supplementary Information

Supplementary Table 1

Glossary

Low-Level-Descriptors (LLDs) specification		
Identifier	**Production type**	**Description**
audSpec_Rfilt	vocal tract	Relative Spectral Transform (RASTA)-style filtered applied to Auditory Spectrum
audspecRasta_lengthL1norm	vocal tract	Relative Spectral Transform applied to Auditory Spectrum and engthL1norm is the magnitude of the L1 norm
audspec_lengthL1norm	vocal tract	Magnitude of L1 norm of Auditory Spectrum
mfcc	vocal tract	Mel-frequency cepstral coefficients
pcm_fftMag_fband1000-4000	global	Fast Fourier transform magnitude of frequency band between 1000Hz to 4000Hz
pcm_fftMag_spectralFlux	global	Evaluates the temporal variation of the logarithmically scaled rate-map across adjacent frames

Functionals	
Identifier	**Description**
iqr[a][b]	inter-quartile range [a]-[b]
Kurtosis	kurtosis (4th order moment)
peakMeanAbs	arithmetic mean of peaks (local maxima)
percentile99.0	outlier-robust maximum value of the contour, represented by the 99% percentile
posamean	arithmetic mean of positive values only
quartile[a]	a's quartile
Range	difference between maximum and minimum observed values

stddev	standard deviation of the values in the contour

Type of feature

Identifier	Description
sma	smoothed static feature representation
_sma_de_	smoothed dynamic feature representation also known as delta coefficients

Note. There are three types of LLDs based on production type: source-based, vocal tract-based, and global (i.e., all frequencies are involved).

Supplementary Table 2

Top-ranking features for different set predictions

Pre vs. embarrassment		
Nr. of feature	**feature**	**weight**
716	audSpec_Rfilt_sma_de[4]_iqr1-3	0.00415
1175	mfcc_sma_de[3]_quartile1	0.00391
2157	mfcc_sma_de[2]_posamean	0.00322
838	audSpec_Rfilt_sma_de[14]_iqr1-2	0.00281
823	audSpec_Rfilt_sma_de[13]_quartile1	0.00275
136	audSpec_Rfilt_sma[5]_leftctime	0.00275
927	audSpec_Rfilt_sma_de[21]_quartile1	0.00271
321	audSpec_Rfilt_sma[23]_leftctime	0.00261
Post vs. embarrassment		
Nr. of feature	**feature**	**weight**
11	audspecRasta_lengthL1norm_sma_de_iqr2-3	0.00433922
12	audspecRasta_lengthL1norm_sma_de_iqr1-3	0.00346824
176	audSpec_Rfilt_sma_de[4]_quartile1	0.00330195
583	audSpec_Rfilt_sma_de[20]_posamean	0.00323615
181	audSpec_Rfilt_sma_de[4]_iqr1-3	0.0031257
238	audSpec_Rfilt_sma_de[16]_iqr1-3	0.00309433
521	mfcc_sma[3]_meanRisingSlope	0.00288177
271	audSpec_Rfilt_sma_de[21]_iqr1-3	0.00280821
267	audSpec_Rfilt_sma_de[21]_quartile1	0.00276045
328	mfcc_sma_de[3]_quartile1	0.00271235
(pre+post) vs. embarrassment		
Nr. of feature	**feature**	**weight**
852	audSpec_Rfilt_sma_de[4]_iqr1-3	0.00501
69	audspecRasta_lengthL1norm_sma_de_iqr2-3	0.00395
871	audSpec_Rfilt_sma_de[5]_iqr2-3	0.00317
70	audspecRasta_lengthL1norm_sma_de_iqr1-3	0.00302
2508	audSpec_Rfilt_sma_de[21]_posamean	0.00298
1080	audSpec_Rfilt_sma_de[19]_iqr2-3	0.00292
1401	mfcc_sma_de[3]_percentile99.0	0.00280
1373	mfcc_sma_de[1]_kurtosis	0.00271

2114	mfcc_sma_de[3]_quartile3	0.00195

Note on top-ranking features ((pre+post) vs. embarrassment)

{852 audSpec_Rfilt_sma_de[4]_iqr1-3, 871 audSpec_Rfilt_sma_de[5]_iqr2-3, 2508 audSpec_Rfilt_sma_de[21]_posamean, 1080 audSpec_Rfilt_sma_de[19]_iqr2-3}
Represents supra-segmental FR composed from: LLD - Relative Spectral Transform (RASTA)-style filter applied to Auditory Spectrum, dynamic features - delta coefficients and following functionals: inter-quartile range {1-3, 2-3} and arithmetic mean of positive values only.

{69 audspecRasta_lengthL1norm_sma_de_iqr2-3, 70 audspecRasta_lengthL1norm_sma_de_iqr1-3}
Represents supra-segmental FR composed from: LLD - Relative Spectral Transform applied to Auditory Spectrum and lengthL1norm is the magnitude of the L1 norm, dynamic features - delta coefficients and following functional: inter-quartile range {1-3} and {2-3}.

{1401 mfcc_sma_de[3]_percentile99.0, 1373 mfcc_sma_de[1]_kurtosis, 2114 mfcc_sma_de[3]_quartile3}
Represents supra-segmental FR composed from: LLD - Mel-Frequency Cepstral Coefficients (MFCC), dynamic features - delta coefficients and following functionals: outlier-robust maximum value of the contour, represented by the 99% percentile, kurtosis (4th order moment) and 3rd quartile. One could observe that top-ranked MFCC FR are composed from first and third dimension MFCCs {MFCC 1 and MFCC 3} LLDs, a coarse representation of the short-term spectral envelope.

Supplementary Table 3

Shared high Mutual information between VAM and embarrassment

VAM	Feature	MI weights for embarrassment	MI weights for regression task
Valence	*pcm_fftMag_fband1000-4000_sma_range*	0.106	0.117
Valence	*mfcc_sma_de[3]_stddev*	**0.198**	0.090
Valence	*audspec_lengthL1norm_sma_peakMeanAbs*	0.135	0.140
Arousal	*mfcc_sma_de[14]_percentile99.0*	**0.215**	0.153
Arousal	*mfcc_sma_de[14]_stddev*	0.164	0.186
Arousal	*audspec_lengthL1norm_sma_peakMeanAbs*	0.135	**0.442**
Arousal	*mfcc_sma_de[14]_posamean*	0.158	0.176
Dominance	*pcm_fftMag_spectralFlux_sma_de_iqr1-2*	0.120	**0.327**
Dominance	*mfcc_sma_de[14]_quartile3*	0.126	0.137
Dominance	*mfcc_sma_de[14]_percentile99.0*	**0.215**	0.139
Dominance	mfcc_sma_de[14]_stddev	0.164	0.195
Dominance	audspec_lengthL1norm_sma_peakMeanAbs	0.135	**0.428**

Note. MI = mutual information.

The bold print show trends for a higher and relevant shared mutual information of the features of embarrassment and the regression task.

Supplementary Table 4

Share high Mutual information between EMO-DB and embarrassment

EMO-DB	Feature	MI weights for embarrassment	MI weights for classification task
7 emotions	audSpec_Rfilt_sma_de[21]_iqr1-3	**0.301**	0.334
7 emotions	audSpec_Rfilt_sma_de[20]_iqr1-3	0.257	0.367
7 emotions	mfcc_sma_de[2]_quartile1	0.239	0.419
7 emotions	audspecRasta_lengthL1norm_sma_de_iqr1-3	0.230	**0.452**
7 emotions	audspecRasta_lengthL1norm_sma_de_iqr2-3	0.202	0.422
Valence	audSpec_Rfilt_sma_de[4]_iqr1-2	**0.321**	0.061
Valence	audSpec_Rfilt_sma_de[4]_iqr1-3	0.292	**0.075**
Valence	mfcc_sma_de[1]_iqr1-3	0.279	0.052
Valence	audSpec_Rfilt_sma_de[4]_quartile1	0.259	0.066
Arousal	audSpec_Rfilt_sma_de[21]_iqr1-3	**0.301**	0.200
Arousal	audspecRasta_lengthL1norm_sma_de_iqr1-3	0.230	0.241
Arousal	audSpec_Rfilt_sma_de[18]_iqr1-3	0.220	0.205
Arousal	audSpec_Rfilt_sma_de[19]_iqr2-3	0.209	0.233
Arousal	mfcc_sma_de[2]_iqr2-3	0.206	0.224

Note. 7 emotions = The 7 emotions from EMO-DB: happiness, sadness, disgust, fear, boredom, anger, and neutral.

MI = mutual information.

The bold lines show trends for a higher and relevant shared mutual information of the features of embarrassment and emotion classification.

Appendix B

Article II – Factor Analyses of SPS & SIAS

BMC Psychiatry

RESEARCH

Open Access

Factor structure of the Social Phobia Scale (SPS) and the Social Interaction Anxiety Scale (SIAS) in a clinical sample recruited from the community

Dajana Šipka[1*], Jeannette Brodbeck[1,2], Ava Schulz[1], Timo Stolz[1] and Thomas Berger[1]

Abstract

Background The Social Phobia Scale (SPS) and the Social Interaction Anxiety Scale (SIAS) are widely used self-report questionnaires to assess symptoms of social anxiety. While SPS measures social performance anxiety, SIAS measures social interaction anxiety. They are mostly reported simultaneously, but there have not been consistent results of the joint factor structure and therefore no consistent recommendations on how to use and evaluate the questionnaires. This study aimed (1) to evaluate the underlying joint factor structure of the SPS and SIAS and (2) to test whether SPS and SIAS are reliable scales to assess two different aspects of social anxiety.

Methods The one-factor, two-factor, and bifactor models were tested in a clinical sample recruited from the community and diagnosed with a social anxiety disorder. Exploratory and confirmatory factor analyses were conducted, bifactor-specific indices were calculated, and the content of the less fitting items was examined.

Results Confirmatory factor analyses showed that the best-fitting model was the bifactor model with a reduced set of items. The bifactor-specific indices showed that the factor structure cannot be considered unidimensional and that SPS and SIAS are reliable subscales. A closer examination of the less fitting item content and implications for future studies are discussed.

Conclusions In conclusion, SPS and SIAS can be reported together as an overall score of social anxiety and are separately reliable measures to assess different aspects of social anxiety.

Trial registration This is a secondary analysis of data from two trials registered under ISRCTN75894275 and ISRCTN10627379.

Keywords Social anxiety, Social phobia scale, Social interaction anxiety scale, Factor analysis, Bifactor model

Background

Social anxiety is a common anxiety that many people experience in their day-to-day social life. The range of feared situations can be broad, from difficulty making eye contact to giving a presentation in front of others. Over the years, there have been many categorizations of social anxiety (e.g., [1]). Still, the core fear of all these categories is the fear of negative evaluation by others and not being able to maintain a favorable impression of oneself

*Correspondence:
Dajana Šipka
dajana.sipka@unibe.ch
[1] Department of Clinical Psychology and Psychotherapy, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland
[2] School of Social Work, University of Applied Sciences and Arts Northwestern Switzerland, Olten, Switzerland

[2]. The Social Phobia Scale (SPS) and the Social Interaction Anxiety Scale (SIAS) are widely used self-report questionnaires to assess social anxiety symptoms and are often reported together [3]. They both claim to measure different aspects of social anxiety: While SPS measures social performance anxiety, SIAS measures social interaction anxiety [4].

Even though the questionnaires are widely used in research and in clinical practice, there are no consistent recommendations on how to use and assess the questionnaires. There have been two different recommendations: Gomez & Watson [5] and Thompson et al. [3] recommended, based on non-clinical samples, to use the SPS and SIAS only simultaneously for the assessment of general social anxiety (i.e., report the total sum of both questionnaires), but not separately for the assessment of social performance anxiety by SPS respectively the social interaction anxiety by SIAS (i.e., report the separate sum score of each questionnaire). By contrast, based on clinical samples, Heimberg et al. [4] and Heidenreich et al. [6] recommended using the SPS and SIAS independently of each other to assess and report social performance and social interaction anxiety separately. One possible explanation for the inconsistent recommendations may be that the two aspects of social anxiety only exist distinctly at a clinical level [3].

These recommendations were all made based on the results of factor analyses. Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are possible methods to determine the joint factor structure and draw recommendations on how to use the two questionnaires in practice. Evidence regarding the joint factor structure of the SPS and the SIAS is inconsistent. Gomez & Watson [5] found a bifactor solution with SPS and SIAS as group factors, while almost all items loaded on an additional general factor. By contrast, Safren et al. [7] found a three-factor solution, with one general factor being a higher-order factor. Carleton et al. [8] found a three-factor solution as well, without a higher-order factor but with one SIAS factor and two SPS factors. Heidenreich et al. [6] on the other hand found a two-factor solution, SPS and SIAS, but suggested due to the high correlation ($r=0.69$) "[…] that both scales represent facets of a higher-order construct." (p.583). Most studies that look at the questionnaires' factor structure, use uni- or multi-dimensional models and do not test for the bifactor model (cf. previous studies listed by Eidecker et al. [9] or Gomez & Watson [5]). This study provides a more complex and relatively new factor structure analysis than most previous studies since the current study simultaneously compared the one-factor, two-factor, and bifactor models in a clinical sample from the community. Additionally, with the bifactor-specific indices, it is possible to analyze further the resulting subscales and their reliability (see 2.5. Factor Analyses: Models).

Multiple aspects make the comparison between different factor analyses studies using these two questionnaires challenging: Some studies compare the structure of SIAS with the structure of SPS separately (e.g., [10]) or only assess one of the questionnaires (mostly SIAS, e.g., [9, 11, 12]), even though in practice the two questionnaires are often used together. Most studies do not use clinical samples (i.e., participants with a social anxiety disorder (SAD) diagnosis), but the general population or student samples. Rodebaugh et al. [11] compared the SIAS structure of clinical and undergraduate student samples. They found slight but systematic differences in the responses and suggested a separate cut-off score for the student sample. Carleton et al. [8] tested the joint factor structure of SPS and SIAS with a student and a clinical sample and found the same three-factor solution for both samples but included only 14 out of the overall 40 items in the final model. As Kupper & Denollet [10] summarized, the items may have different meanings for the clinical and the non-clinical samples.

Depending on the resulting factor model, different implications can be drawn based on the model: A one-factor model would support using an overall sum score of both questionnaires to assess general social anxiety, but it would not support using both questionnaires to assess distinct aspects of social anxiety (i.e., social performance and interaction anxiety) through two separate sum scores of each questionnaire. A two-factor model would support the distinction of the two social anxiety aspects by calculating each sum score separately, but not the joint use of the questionnaires to assess general social anxiety by calculating the overall sum score. A validly proven bifactor model with good bifactor-specific indices would firstly, support evaluating general social anxiety by calculating the overall sum score and would secondly, support using both questionnaires separately to assess social performance and interaction anxiety as distinct aspects of social anxiety by calculating each sum score separately. As Thompson et al. [3] pointed out, the question about the different aspects of social anxiety is especially interesting since the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; [13]) has introduced the performance anxiety category within the social anxiety disorder diagnosis.

Methods

Aims

The two aims of this study were 1) to assess the joint factor structure of SPS and SIAS in a clinical sample since there are only a few studies using clinical samples and 2) to test whether SPS and SIAS are reliable scales to assess

two different aspects of social anxiety beyond the assessment of social anxiety in general. Two clinical samples from two studies were used. After testing whether the two samples could be combined, the joint factor structure of SPS and SIAS was tested in a combined sample using CFA, EFA, bifactor model-specific indices and a content examination of the less fitting items in the final model. Studies usually do not thoroughly examine the less-fitting item content but rather report only the statistical characteristics (e.g., item loadings). By examining the content, one might find a common theme and might indicate not only that the items are not fitting, but also why they are not. The following models were tested: one-factor, two-factor, and bifactor model.

Measures

Social Phobia Scale and Social Interaction Anxiety Scale

The Social Phobia Scale (SPS; [14]; German Version: [15]) is a self-report questionnaire that assesses social anxiety in performance-related, routine situations and activities (e.g., eating, writing in front of others, etc.). The Social Interaction Anxiety Scale (SIAS; [14]; German Version: [15]) is a self-report questionnaire that assesses social anxiety in interactional situations (e.g., conversations with strangers, friends, etc.). Both questionnaires contain 20 items each. Response categories are from 0 = "not at all" to 4 = "extremely", with a total sum score ranging from 0 to 80 per questionnaire, respectively, from 0 to 160 for both questionnaires combined. Items 5, 9, and 11 in the SIAS are reversed scored. The internal consistency of the German versions (which were used in the samples of this study) for the SAD patient population is excellent, with $\alpha = 0.94$ for each questionnaire [15]. The internal consistency for the current combined sample in this study is lower but still excellent, with $\alpha = 0.91$. See Supplementary file 1 for the item wording.

Participants

Two samples from two different studies were used, namely from Schulz et al. ([16], sample (a)) and Stolz et al. ([17], sample (b)). Both studies recruited participants from the general population through newspaper articles, interviews on radio and TV, and online forums in Switzerland, Germany, and Austria. Participants could download the study information (where participants were informed about the whole study procedure) and the informed consent form (ICF) from the respective website and send back the signed ICF via e-mail or post. Both studies established a SAD diagnosis by using the Structured Clinical Interview for DSM-IV – Axis I disorder (SCID; [18]). Schulz et al. [16] conducted a randomized controlled trial (RCT) comparing the efficacy of an internet-based cognitive behavioral therapy (ICBT) program for SAD between a clinician-guided group ICBT with clinician-guided individual ICBT, and a waitlist control group (WL). The final sample included 149 participants with a SAD diagnosis. Stolz et al. [17] used a three-armed RCT for comparing the efficacy of an ICBT program for SAD between three groups using the program as a mobile version, as a computer version, and a WL. The final sample included 150 people with a SAD diagnosis.

The inclusion criteria for both studies were a cut-off score on the SPS of > 22 or on the SIAS of > 33, a SAD diagnosis according to the SCID, an age of at least 18 years, access to a computer with internet connection, sufficient mastery of the German language, and no psychotherapy during study participation. Exclusion criteria were active suicidal plans, a history of bipolar or psychotic disorders, and no change in psychiatric medication during the previous month (if psychiatric medication intake was present). Both studies worked with the same ICBT program for SAD [19, 20]. For this study, only the baseline data was used. All items were answered without any missing data apart from one participant who was excluded as no detailed item data were available.

Sample comparison

To test whether both samples could be combined, the samples' socio-demographic data (i.e., age, sex, education, and relationship status), as well as item responses, were compared. The program R Studio (Version 2023.6.1.524) was used to compare the two samples. Shapiro tests showed no normal distribution for age nor for any of the item responses per sample (SPS 1–SPS 20, SIAS 1–SIAS 20) ($p < 0.001$). This is because age showed a clear right-skewed distribution as 62% of the sample was between 18 to 35 years of age within a total range of 18 to 76 years. Therefore, non-parametric tests were used for all comparisons. The age difference was calculated with the Mann–Whitney U test. Since the other socio-demographic data were dummy coded, Chi-square tests were used. For the item response comparisons, multiple Mann–Whitney U tests were used.

Factor analyses: models

Three models were tested: a one-factor, a two-factor, and a bifactor model (see Fig. 1). The one-factor model assumed that one factor for all items (SPS 1–SPS 20, SIAS 1–SIAS 20) adequately represents social anxiety. The two-factor model assumed two factors (SPS and SIAS). All SPS items (SPS 1–SPS 20) were assumed to load on the SPS factor and to measure social performance anxiety and all SIAS items (SIAS 1–SIAS 20) were assumed to load on the SIAS factor and to measure social interaction anxiety. The marker indicators were SPS 1 and SIAS 1. The covariance was not set to 0. The factors were

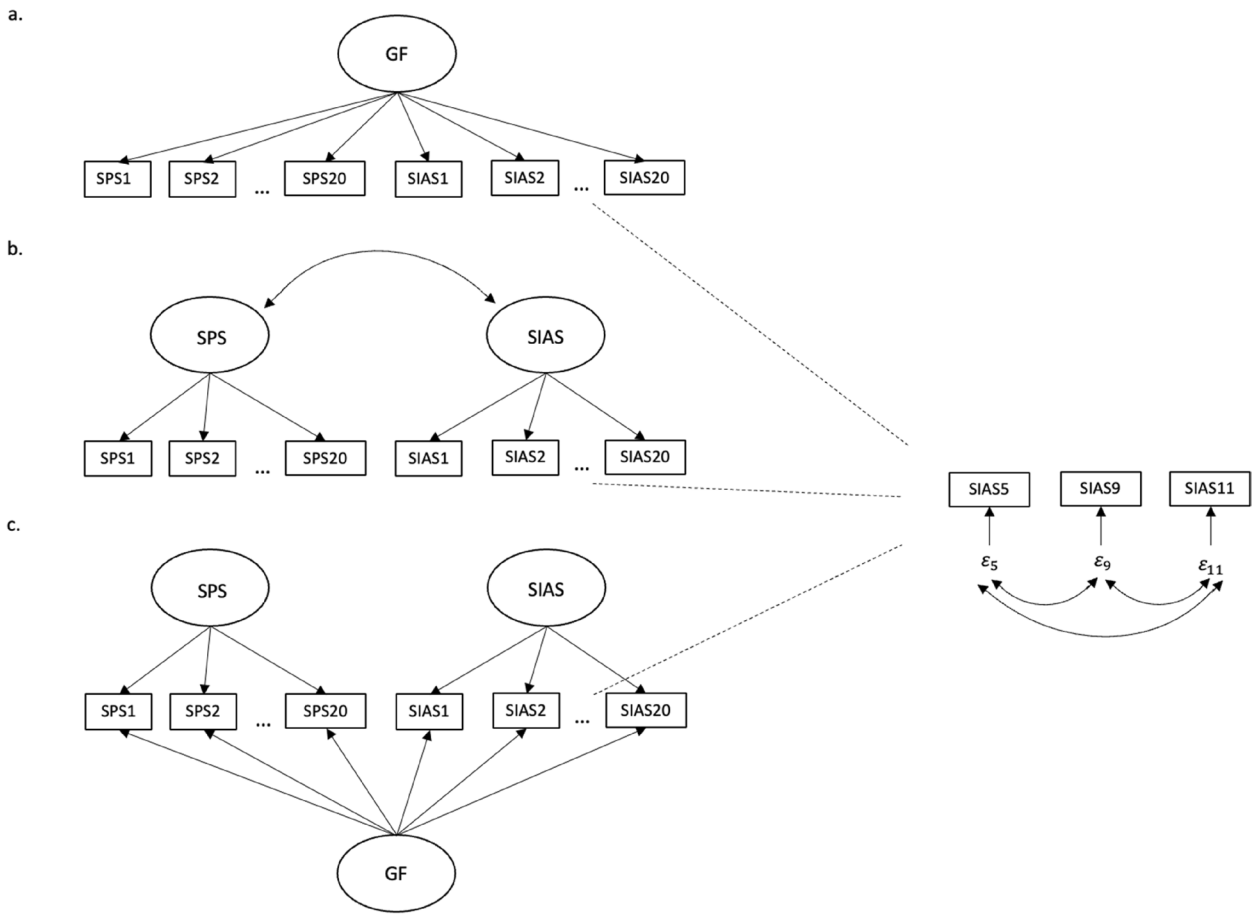


Fig. 1 One-factor (**a**), two-factor (**b**), and bifactor model (**c**). SIAS 5, SIAS 9, and SIAS 11 are reversed items; therefore, their errors were additionally correlated in each model (a, b, and c) to account for method variance

assumed to be oblique, since it was unrealistic for the two factors, which are both supposed to measure social anxiety, not to correlate at all and to be orthogonal. The bifactor model assumed three factors (GF, SPS, and SIAS) and was a combination of the first two models. All SPS items (SPS 1–SPS 20) loaded on the group factor SPS and all SIAS items (SIAS 1–SIAS 20) loaded on the group factor SIAS, while all 40 items simultaneously loaded on GF. The marker indicators were again SPS 1 and SIAS 1. As usual for the bifactor model, covariance was set to 0 for all factors. Thus, they were orthogonal [21].

The Mplus8 base program (Version 1.8.7.; Muthén & Muthén, 1998–2017) was used for the factor analyses. Model estimation was performed using the weighted least square means and variance adjusted (WLSMV) estimator, which is adequate for the categorial items and the lack of normal distribution [22]. The three reversed SIAS items SIAS 5, SIAS 9, and SIAS 11 have been repeatedly found to not belong to the other SIAS items and were assumed to rather measure extraversion [5, 9–12]. Their errors were correlated in all models as it is recommended by Brown [23] for reversed scored items in general. No cross-loadings between the items were allowed in any of the models.

Model fit was assessed through the following indices: Tucker Lewis index (TLI), comparative fit index (CFI), root mean square error of approximation (RMSEA), and standard root mean square residual (SRMR). Following Hu & Bentler [24], TLI and CFI of ≥ 0.90 were considered as acceptable fit and ≥ 0.95 as good fit. RMSEA and SRMR ≤ 0.08 were considered as acceptable fit and ≤ 0.05 as good fit. For RMSEA, Mplus calculates the 90% confidence interval (CI), where the lower CI should not be higher than 0.05 (the closer to 0, the better) and the upper not higher than 0.08 [25].

χ^2 depends on the sample, so it is recommended to use other model indices than solely χ^2 and to not reject or accept a model based on the significance of χ^2 only [25, 26]. With this sample size, χ^2 was expected to be significant (i.e., reject the proposed models) (cf. [27]). Model fit indices for a bifactor model tend to be better in comparison to a first-order factor model, especially if the bifactor

model does not reflect the true model [28]. χ^2 is in this case not a suitable comparison coefficient, therefore, additional coefficients for the evaluation of the bifactor model were used. Even if a bifactor model turns out to have the relatively best fit, it is still not proof of "enough" multidimensionality, nor of useful subscales [29]. The following coefficients were used following Rodriguez et al. [21]: omega (ω), omega hierarchical (ω_h), omega hierarchical subscale (ω_{hs}), explained common variance (ECV), percent uncontaminated correlations (PUC), factor determinacy (FD), and coefficient H (H) (for further explanation on the individual coefficients, see [21]).

Results

Comparisons of sample (a) and (b) and characteristics of the analysis sample

Sample (a) and (b) did not differ significantly in age, sex, education, or relationship status (see Table 1). The average age in sample (a) and (b) combined was 35.1 years (SD = 11.2) with 58% women. The majority of the sample had a tertiary education (70%) and was single (51%). Furthermore, none of the comparisons of the mean responses of the 40 items were significant (see Supplementary file 2, Supplementary Table 1 for all item comparisons). Thus, we combined both samples for the following analyses.

Model comparisons

Table 2 shows the goodness of fit indices and comparisons of all models. As expected, all three models showed a highly significant χ^2 with $p < 0.001$. The one-factor model (O) did not show any acceptable indices. All items loaded significantly on the factor. The two-factor model (T) showed only an acceptable RMSEA. The remaining indices were not acceptable. The factors SPS and SIAS correlated moderately with $r = 0.57$, which supports the decision to assume an oblique structure.

The bifactor model (B) showed two acceptable indices, RMSEA and SRMR, but CFI and TLI remained not acceptable. Seven items did not load on the general nor on the specific factors and one error correlation was not significant ($p > 0.05$): For group factor SPS: SPS 3 ($p = 0.407$), and SPS 18 ($p = 0.207$). For GF: SPS 1 ($p = 0.929$), SPS 7 ($p = 0.359$), SPS 10 ($p = 0.499$), SPS 19 ($p = 0.822$), and SIAS 5 ($p = 0.123$). For the error correlation: SIAS 5 × SIAS 11 ($p = 0.155$). A refined bifactor model without the non-significant elements was tested. The modification suggestion, also used by Eidecker et al. [9], to add the correlation between SIAS 12 × SIAS 17, was additionally implemented. All four indices showed acceptable values and all remaining elements were significant. Since most of the other studies initially did not include any error correlations in their models, the same three models were tested again, but without the error correlations (O1, T1, B1). All three models were not better without the error correlations in comparison to with; the same indices remained acceptable respectively not acceptable. The findings [6, 11, 12, 30] that the reversed items should be excluded from the final set of items was only partly shown since only SIAS 5 was not significant.

Table 1 Socio-demographic information of the samples (a), (b), and the full sample (both samples combined)

	Sample (a)		Sample (b)				Full sample	
	M	SD	M	SD	*U*	*p*-value	M	SD
Age	35.4	11.2	34.8	11.2	11406	.682	35.1	11.2
	Sample (a)		Sample (b)				Full sample	
	n	%	n	%	χ^2	*p*-value	n	%
Sex					2.7	.101		
women	79	53%	93	62%			172	58%
men	70	47%	56	38%			126	42%
Education					1.04	.596		
primary	5	3%	8	71%			13	4%
secondary	41	28%	36	24%			77	26%
tertiary	103	69%	105	5%			208	70%
Relationship status					7.7	.053		
single	77	52%	74	50%			151	51%
married	27	18%	36	24%			63	21%
separated, divorced, widowed	11	7%	2	1%			13	4%
in a relationship (not married)	34	23%	37	25%			71	24%
Total	149		149				298	

Sample (a) = sample from Schulz et al. [16]; Sample (b) = sample from Stolz et al. [17]; *U* = Mann–Whitney U coefficient

Table 2 Model indices and comparison of the three models

Model	RMSEA [90% CI]	SRMR	CFI	TLI	χ^2	df
One-factor (O)	0.093 [0.089, 0.097]	0.108	0.719	0.702	2648.812***	737
One-factor[1] (O1)	0.094 [0.091, 0.098]	0.110	0.711	0.695	2704.934***	740
Two-factor (T)	0.073 [0.069, 0.077]	0.089	0.830	0.820	1891.400***	736
Two-factor[1] (T1)	0.074 [0.070, 0.078]	0.091	0.825	0.815	1931.936***	739
Bifactor (B)	0.059 [0.054, 0.063]	0.067	0.894	0.882	1416.337***	697
Bifactor[1] (B1)	0.059 [0.055, 0.063]	0.067	0.893	0.881	1428.522***	700
Bifactor without non-significant elements	0.052 [0.047, 0.056]	0.067	0.917	0.908	1267.480***	704

RMSEA Root mean square error of approximation; *SRMR* Standard root mean square residual, *CFI* comparative fit index, *TLI* Tucker Lewis index

*** $p < .001$

[1] model without error correlations

To exclude any potential sample-specific biases, the models were tested separately for sample (a) and sample (b). None of the models with separate samples had more acceptable indices than the models with the combined sample. Additionally, the desired ratio χ^2/ df < 2 was only found for the bifactor model without the non-significant elements (cf. [8]). No χ^2 difference tests were reported since they cannot be conducted in a regular way for models based on the WLSMV estimator (cf. [6]).

In conclusion, only the refined bifactor model showed an acceptable, but not a good fit to the data. Moreover, there were numerous non-significant item loadings primarily in the GF. Therefore, we performed EFAs with the same data set.

Exploratory factor analysis

The requirements for an EFA as well as the estimated factor number were conducted with RStudio and Mplus. For the correlation matrix, which was used for both tests, a spearman correlation was used, since normal distribution was not given, and the spearman correlation is suited for ordinal variables. The rotation bi-geomin for orthogonal structures was used. The Bartlett test of sphericity and the Kaiser–Meyer–Olkin test indicated that the sample was adequate for an EFA (the Bartlett test of sphericity: χ^2 (780, $N = 298$) = 4651.873, $p < 0.001$; matric sample adequacy (MSA) of the Kaiser–Meyer–Olkin test: 0.88 (range: 0.73 – 0.94), which indicates excellent MSA according to Kaiser [31]).

Since the bifactor model in the CFA was the best one so far, it was assumed that the new model should be a bifactor model as well, but with potentially more than two group factors. The parallel analysis indicated 5 factors. The Kaiser-Guttman rule would have indicated 10 factors (eigenvalue > 1), while the first 3 factors seemed to explain the most variance, with an eigenvalue of 10.8, 4.4 respectively 2.3. All the other factors had an eigenvalue < 2. Moosbrugger & Hartig [32] and Brown [23] indicated that the Kaiser-Guttman rule could lead to over-factorization, which was probably the case here. Items with loadings < 0.30 (i.e., not salient loadings) were excluded [23]. When cross-loadings occurred and the difference between the two loadings was > 0.10, the item with the higher loading was chosen. When the difference was < 0.10, it was considered a cross-loading (cf. [7]).

The following models were tested: bifactor model with two, three, four, and five group factors. In every model, the following same items as in the original bifactor model without the non-significant elements were not significant for GF: SPS 1, SPS 7, SPS 19, and SIAS 5. SPS 10 was also not significant for all models except for the bifactor model with five group factors. All models had between five and eight non-significant items in the GF and between eleven and eighteen non-significant items in the group factors. With an item-driven approach, we were not able to find distinct categories for any of the models. We were not able to find a better-fitting model than the current bifactor model. Therefore, the bifactor-specific indices were calculated for this model.

Bifactor-specific indices

Table 3 shows the final refined bifactor model including the factor loadings as well as the bifactor model indices. The indices were calculated with the Bifactor Indices Calculator by Dueber [33]. Comparing ω and ω_h of GF showed that 74% (0.7 / 0.94) of the reliable variance in total scores is attributed to GF, while 24% (0.94 – 0.7) of the reliable variance in total scores is attributed to multidimensionality caused by SPS and SIAS. ω_s of SPS and SIAS indicated that SPS and SIAS are reliable subscales with 97% (0.91 / 0.94) for SPS and 98% (0.92 / 0.94) for SIAS of the reliable variance being independent of GF [29]. ω_{hs} showed that the common variance explained by the group factors, while controlling for GF, is still

Table 3 Standardized factor loadings of refined bifactor model and bifactor model indices

Item	GF	SPS	SIAS
SPS 1	–	0.44	
SPS 2	0.28	0.20	
SPS 3	0.47	–	
SPS 4	0.58	0.43	
SPS 5	0.38	0.21	
SPS 6	0.61	0.33	
SPS 7	–	0.70	
SPS 8	0.50	0.54	
SPS 9	0.33	0.51	
SPS 10	–	0.70	
SPS 11	0.22	0.57	
SPS 12	0.68	0.16	
SPS 13	0.36	0.55	
SPS 14	0.31	0.41	
SPS 15	0.63	0.30	
SPS 16	0.60	0.39	
SPS 17	0.51	0.53	
SPS 18	0.66	–	
SPS 19	–	0.56	
SPS 20	0.67	0.22	
SIAS 1	0.49		0.17
SIAS 2	0.56		0.19
SIAS 3	0.42		0.44
SIAS 4	0.32		0.39
SIAS 5	–		0.58
SIAS 6	0.42		0.27
SIAS 7	0.43		0.35
SIAS 8	0.29		0.41
SIAS 9	0.22		0.36
SIAS 10	0.40		0.68
SIAS 11	0.19		0.77
SIAS 12	0.58		0.39
SIAS 13	0.34		0.43
SIAS 14	0.40		0.41
SIAS 15	0.52		0.68
SIAS 16	0.49		0.32
SIAS 17	0.58		0.36
SIAS 18	0.42		0.45
SIAS 19	0.62		0.41
SIAS 20	0.51		0.29
omega (ω) / ω subscale (ω_s)	.94	.91	.92
omega hierarchical (ω_h) / ω_h subscale (ω_{hs})	.70	.52	.47
ECV	.51	.24	.25
PUC	.42		
FD	.94	.91	.92
H	.92	.85	.86

"–" for a factor loading means that this loading was not significant and was excluded for the final bifactor model

GF General factor, *SPS* SPS group factor, *SIAS* SIAS group factor, *ECV* Explained common variance, *PUC* Percent uncontaminated correlations, *FD* Factor determinacy; *H* Coefficient H

relatively high with 52% explained by SPS respectively 47% explained by SIAS, which is further in favor of the reliability for SPS and SIAS as separate reliable subscales to measure the two different aspects of social anxiety.

ECV and PUC were below 0.70, indicating that the common variance is not unidimensional respectively a unidimensional model may not be sufficient to represent the data. Additionally, the low PUC indicated probably biased parameters if the model was forced as unidimensional. All three FDs were above 0.9 as well as all three Hs were above 0.7, indicating a good construct reliability and therefore a stable representation of the latent variable by the indicators and a good representation of the two aspects of social anxiety [21].

In conclusion, while GF explained most of the variance (70%) and half of the common variance, the group factors explained equally the remaining common variance and were found to be reliable. There are clear indicators that the bifactor model is tendentially multidimensional.

Content analysis of final bifactor model

The final bifactor model in Table 3 shows several items which were not significant for the general factor GF and the group factor SPS. Additionally, there were some item loadings in the general as well as in the two group factors that were < 0.30. This section will first, examine the content of the none-significant items and second, the content of the items with item loadings < 0.30.

There were seven items in the group factor SPS and general factor GF, which the item loadings were not significant for. For the group factor SPS, the items SPS 3 ("I can suddenly become aware of my own voice and of others listening to me") and SPS 18 ("I get tense when I speak in front of other people") were not significant. The two items refer to aspects of the voice and being listened to respectively public speaking situations in general. For GF, items SPS 1 ("I become anxious if I have to write in front of other people"), SPS 7 ("I worry about shaking or trembling when I'm watched by other people"), SPS 10 ("I would find it difficult to drink something if in a group of people"), SPS 19 ("I worry my head will shake or nod in front of others"), SIAS 5 ("I find it easy to make friends of my own age") were not significant. SPS 1, SPS 7 and SPS 10 refer to mundane performance tasks, while SPS 19 and SIAS 5 do not match the rest with worrying about losing control and making friends. In summary, there are more items from the SPS questionnaire that do not seem to fit either performance-related social anxiety (i.e., not significant for group factor SPS), or general social anxiety (i.e., not significant for general factor). We were not able to find a common interpretation of the topics covered by said items.

There were twelve item loadings across the two group factors and GF that were below the threshold of < 0.30. Since the loadings are standardized and there were no cross-loadings, the item loadings can be interpreted as correlations between the items and the factors [23]. For the group factor SPS, the item loadings of SPS 2 ("I become self-conscious when using public toilets"), SPS 5 ("I fear I may blush when I am with others"), and SPS 12 ("I am worried people will think my behaviour odd") were below the threshold. There is no apparent similarity across these three item contents. Therefore, there is a small correlation between social performance anxiety and the fear to blush, public toilet usage, and appearing odd. For the group factor SIAS, the item loadings for SIAS 1 ("I get nervous if I have to speak with someone in authority (teacher, boss, etc.)"), SIAS 2 ("I have difficulty making eye-contact with others"), SIAS 6 ("I tense-up if I meet an acquaintance in the street"), and SIAS 20 ("I am unsure whether to greet someone I know only slightly") were below the threshold. All the items refer to situations involving people one does not know on a personal level and maintaining eye contact. Therefore, there is a small correlation between social interaction anxiety and situations involving acquaintances and maintaining eye contact. For GF, the item loadings for SPS 2 ("I become self-conscious when using public toilets"), SPS 11 ("It would make me feel self-conscious to eat in front of a stranger at a restaurant"), SIAS 8 ("I feel tense if I am alone with just one other person"), SIAS 9 ("I am at ease meeting people at parties, etc."), and SIAS 11 ("I find it easy to think of things to talk about") were below the threshold. While the last four items refer to situations involving strangers, SPS 2 again does not match the rest. Therefore, there is a small correlation between social anxiety in general and situations involving strangers and public toilet usage.

In conclusion, there were, in general, more SPS items than SIAS items that were not significant (six SPS and two SIAS items), which means that social performance anxiety is less well represented in this model than interaction anxiety. Concerning the items below threshold, more SIAS items had only a small correlation than SPS items (four SPS and seven SIAS items). All SIAS items with small correlations revolved around acquaintances, respectively strangers and situations involving talking to others, which implies that situations involving conversations and unknown people represent social anxiety in general and social interaction anxiety less well. SPS 2 was the only item that had a small correlation with both GF and group factor SPS and did not seem to represent social anxiety in general or social performance anxiety well. There was no apparent similarity between all the other SPS items (blushing, eating in front of others, and appearing odd).

Discussion

SPS and SIAS are one of the most widely used questionnaires to assess social anxiety, but to this day, there is no consensus on their joint factor structure, nor general consensual implications on how to use and evaluate the questionnaires in practice, whether they can be reliably evaluated simultaneously as well as separately. The two aims of this study were to investigate the joint factor structure of the SPS and SIAS and to provide practical implications for the use of the questionnaires based on these results. The one-factor, two-factor, and bifactor models were tested with a CFA and other possible bifactor models were tested with an EFA. Further, bifactor-specific indices as well as item contents were examined.

The best model in comparison with an acceptable fit was the bifactor model without the eight non-significant elements. The corresponding bifactor indices showed that GF, SPS, and SIAS were reliable factors, that SPS and SIAS both explained a high amount of reliable variance independent of GF and that the joint factor structure of SPS and SIAS is tendentially multidimensional. According to these results, the combined use in practice, as well as the report of the two questionnaires separately, is justified. This could not be shown for all items, since seven items in the model were not significant. Thus, the questionnaire SIAS seems to be more fitting to this model since none of the items on SIAS and only one on GF were not significant. One possible explanation is that the non-significant items do not represent the construct performance anxiety accurately. Another explanation comes from Caballo et al. [34], who found that items that referred to cognitive aspect of social anxiety, which is especially true for SPS items, were not always clear to university students and were misunderstood in their study, which could also be the case for this sample. On the other hand, there were more SIAS items than SPS items across GF and the group factors that were below the threshold. Most of the SIAS items revolved around situations involving acquaintances or strangers. There was only a small correlation between these types of situations and interaction respectively general social anxiety. SPS 2 (usage of public toilet) was the only item that was below the threshold for GF as well as for the group factor SPS. This could be an indication that the newly introduced phobia "paruresis" could be considered a separate disorder.

The deviation of the three reversed SIAS items from the other SIAS items could only partly be found in this study since only item SIAS 5 was not significant for GF. Additionally, the item loadings of SIAS 9 and SIAS 11 in

GF were below the threshold of 0.30, indicating a small correlation between the two reversed items and general social anxiety.

There are several reasons that could explain the general inconsistent factor solutions across different studies. The first explanation is according to Caballo et al. [35] the non-empirical approach of constructing the questionnaires in the first place, where not enough attention was paid to content validity. The second explanation is that the questionnaires were originally created for English-speaking users and are usually directly translated into other languages without concern for possible cultural differences [35]. Depending on the culture of the sample, the items of the questionnaires may not mean the same for every culture. Carter et al. [30] have demonstrated that the item SPS 5, the fear of blushing, for example, may not be relevant for most of the African American population. The third explanation is according to Caballo et al. [34] the often vague definition of the target group in the SPS and SIAS items. The items ask for example about "people" and "others" which could lead to a broad spectrum of interpretations and in the end not referring to the same situation. The fourth explanation may be the primary focus of the items on cognitive aspect according to Caballo et al. [34], which was mentioned before. Particularly in samples from the community, it may be that some of the items were not understood correctly. The fifth explanation concerns the problem with the number of SIAS items: Some studies use the SIAS questionnaire with 19 items, other the SIAS with 20 items. Along the way, a 20th item was added to the set of Mattick & Clarke [14]. This was also mentioned by Rodebaugh et al. [11], who pointed out that Mattick & Clarke [14] once mentioned 20 items for the SIAS on p. 462, but only reported 19 items at the end. It does not seem to be clear what exactly happened. This makes it further difficult to compare studies.

There are several limitations to this study. Firstly, CFA is a large-sample technique [26]. It may be that the sample size is not big enough to produce robust results for the resulting complex bifactor model. Secondly, no conclusions on the validity of the two questionnaires are possible. Thompson et al. [3] for example do not support the use of SPS and SIAS as separate questionnaires to assess additional to general social anxiety performance respectively interaction anxiety due to their insufficient discriminant criterion validity. This is to our knowledge the only study that tested the validity based on behavioral tests (i.e., inducing performance respectively interaction anxiety). The validity would need to be tested in future clinical studies, especially the construct validity since it assumed that SPS measures performance and SIAS interaction anxiety, but there are almost no studies that tested these assumptions with actual behavioral tests as Thompson et al. [3] did. In general, all the studies that did not recommend the separate use of SPS and SIAS were based on non-clinical samples. It may also be that the clinical population has a more clear and distinct picture of social anxiety. Thirdly, the sample is self-selective as participants are self-referred to the intervention, were prepared to download the study information from a website and send the ICF back. This suggests that participants may have had a high level of motivation to seek treatment. Also, while age, relationship status, and the balanced sex were comparable to the typical clinical SAD population (see [13]), the education level was higher. We can additionally speculate that the sample showed a higher psychosocial functioning than other clinical samples since the participants were able to work through an online therapy program with minimal guidance. These factors might have influenced the external validity of this study. Fourthly, the goodness of fit indices (i.e., RMSEA, SRMR, CFI, TRI) of the final refined bifactor model are only acceptable. This may reflect the overall inconsistent current literature on the joint factor structure as well as on the separated factor structure of SPS and SIAS. Mattick & Clarke [14] found, when developing the questionnaires, a three-factor structure for SPS and a one-factor structure for SIAS, while the literature overview in the introduction of e.g., Gomez & Watson [5] and Eidecker et al. [9] shows inconsistent results for either questionnaire over different studies.

Conclusion

The findings of this study show that SPS and SIAS are reliable questionnaires to assess specific aspects of social anxiety beyond the assessment of social anxiety in general. It still needs to be examined whether SPS measures social performance anxiety and SIAS social interaction anxiety in a clinical sample as claimed. Future studies with a clinical sample recruited from the community could try to replicate the results and test whether the bifactor model is the relatively best model and whether the same items turn out to not be significant. Social performance anxiety, measured with the SPS, does not seem to be as well represented by the model as social interaction anxiety, measured by the SIAS. It would be interesting to see, whether it is a cultural influence or due to a selection bias from the self-referred sample.

In conclusion, this study shows that the joint factor structure of SPS and SIAS is potentially multidimensional and that they are indeed reliable questionnaires to assess specific aspects of social anxiety beyond the assessment of social anxiety in general.

Abbreviations

B	Bifactor model
B1	Bifactor model without the error correlations
CI	Confidence interval
CFA	Confirmatory factor analysis
CFI	Comparative fit index
DSM-5	5^{Th} edition of the Diagnostic and Statistical Manual of Mental Disorders
ECV	Explained common variance
EFA	Exploratory factor analysis
FD	Factor determinacy
GF	General factor
H	Coefficient H
ICBT	Internet-based cognitive behavioral therapy
ICF	Informed consent form
O	One-factor model
O1	One-factor model without the error correlations
PUC	Percent uncontaminated correlations
RCT	Randomized controlled trial
RMSEA	Root mean square error of approximation
SAD	Social anxiety disorder
SCID	Structured Clinical Interview for DSM-IV – Axis I disorder
SIAS	Social interaction anxiety scale
SPS	Social phobia scale
SRMR	Standard root mean square residual
T	Two-factor model
T1	Two-factor model without the error correlations
TLI	Tucker Lewis index
U	Mann-Whitney U coefficient
WL	Waitlist control group

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12888-023-05142-8.

Additional file 1: Supplementary File 1. Item wording.

Additional file 2: Supplementary Table 1. Comparison of item responses between sample (a) & sample (b).

Acknowledgements
Not applicable.

Authors' contributions
The data acquisition for the samples (a) and (b) was done by A.S., T.S., and T.B. The calculations were done by D.S. with the support of J.B. The manuscript was written by D.S. All authors read and approved the final manuscript.

Funding
This work was supported by the Swiss National Science Foundation, through a Project Funding grant in humanities and social sciences (division I), which was given to Prof. Dr Thomas Berger, with the reference number 10001C_197475. The SNF had no influence on any of processes of this study.

Availability of data and materials
The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
All experimental protocols from the two studies that provided the here used samples ((a) & (b)) were approved by the cantonal ethics commission (ger.: Kantonale Ethikkommission Bern; KEK) and were funded by the Swiss National Science Foundation grants with the reference number: PP00P1_144824/1 for Schulz et al. (16), and the reference number: PP00P1_144824/1 for Stolz et al. (17). Both grants were given to Prof. Dr. Thomas Berger. The registration number for Schulz et al. (16) is: ISRCTN75894275 (controlled-trials.com). The registration number for Stolz et al. (17) is: ISRCTN10627379 (controlled-trials.com). A written informed consent was obtained from the participants involved in the used studies (sample (a) and (b)). The study was conducted in accordance with the Declaration of Helsinki.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 1 March 2023 Accepted: 26 August 2023
Published online: 04 September 2023

References

1. Leary MR, Kowalski RM. Social Anxiety. New York: The Guilford Press; 1995.
2. Clark DM, Wells A. A cognitive model of social phobia. In: Heimberg RG, Liebowitz RM, Hope D, Schneider F, editors. Social Phobia: Diagnosis, Assessment, and Treatment. New York: The Guilford Press; 1995. p. 69–93.
3. Thompson T, Kaminska M, Marshall C, Van Zalk N. Evaluation of the social phobia scale and social interaction anxiety scale as assessments of performance and interaction anxiety. J Psychiatr Res. 2019. https://doi.org/10.1016/j.psychres.2019.01.103.
4. Heimberg RG, Mueller GP, Holt CS, Hope DA, Liebowitz MR. Assessment of anxiety in social interaction and being observed by others: the social interaction anxiety scale and the Social Phobia Scale. Behav Ther. 1992. https://doi.org/10.1016/S0005-7894(05)80308-9.
5. Gomez R, Watson SD. Confirmatory factor analysis of the combined social phobia scale and social interaction anxiety scale: support for a bifactor model. Front Psychol. 2017. https://doi.org/10.3389/fpsyg.2017.00070.
6. Heidenreich T, Schermelleh-Engel K, Schramm E, Hofmann SG, Stangier U. The factor structure of the Social Interaction Anxiety Scale and the Social Phobia Scale. J Anxiety Disord. 2011. https://doi.org/10.1016/j.janxdis.2011.01.006.
7. Safren SA, Turk CL, Heimberg. RG. Factor structure of the Social Interaction Anxiety Scale and the Social Phobia Scale. Behav Res Ther. 1998; https://doi.org/10.1016/S0005-7967(98)00032-1.
8. Carleton RN, Collimore KC, Asmundson GJG, McCabe RE, Antony M. Refining and validating the Social Interaction Anxiety Scale and the Social Phobia Scale. J Depress Anxiety. 2009. https://doi.org/10.1002/da.20480.
9. Eidecker J, Glöckner-Rist A, Gerlach AL. Dimensional structure of the Social Interaction Anxiety Scale according to the analysis of data obtained with a German version. J Anxiety Disord. 2010. https://doi.org/10.1016/j.janxdis.2010.03.020.
10. Kupper N, Denollet J. Social anxiety in the general population: Introducing abbreviated versions of SIAS and SPS. J Affect Disord. 2012. https://doi.org/10.1016/j.jad.2011.08.014.
11. Rodebaugh TL, Woods CM, Heimberg RG, Liebowitz MR, Schneier FR. The factor structure and screening utility of the Social Interaction Anxiety Scale. Psychol Assess. 2006. https://doi.org/10.1037/1040-3590.18.2.231.
12. Rodebaugh TL, Woods CM, Heimberg RG. The Reverse of Social Anxiety Is Not Always the Opposite: The Reverse-Scored Items of the Social Interaction Anxiety Scale Do Not Belong. Behav Ther. 2007. https://doi.org/10.1016/j.beth.2006.08.001.
13. American Psychiatric Association and American Psychiatric Association, Eds., Diagnostic and statistical manual of mental disorders: DSM-5, 5th ed. Washington, D.C: American Psychiatric Association, 2013.
14. Mattick RP, Clarke JC. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behav Res Ther. 1998;36(4):455–70.
15. Stangier U, Heidenreich T, Berardi A, Golbs U, Hoyer J. Die Erfassung sozialer Phobie durch die Social Interaction Anxiety Scale (SIAS) und die Social Phobia Scale (SPS). Z für Klin Psychol Psychother. 1999. https://doi.org/10.1026//0084-5345.28.1.28.
16. Schulz A, Stolz T, Vincent A, Krieger T, Andersson G, Berger T. A sorrow shared is a sorrow halved? A three-arm randomized controlled trial comparing internet-based clinician-guided individual versus group treatment

for social anxiety disorder. Behav Res Ther. 2016. https://doi.org/10.1016/j.brat.2016.07.001.
17. Stolz T, Schulz A, Krieger T, Vincent A, Urech A, Moser C, Westermann S, Berger T. A mobile app for social anxiety disorder: A three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. J Consult Clin Psychol. 2018. https://doi.org/10.1037/ccp0000301.
18. First MB, Spitzer RL, Gibbon M, Williams JB. Structured clinical interview for DSM-IV Axis I disorders (SCID-I). Washington, DC: American Psychiatric Press; 1995.
19. Berger T, Hohl E, Caspar F. Internet-based treatment for social phobia: a randomized controlled trial. J Clin Psychol. 2009. https://doi.org/10.1002/jclp.20603.
20. Berger T, Caspar F, Richardson R, Kneubühler B, Sutter D, Andersson G. Internet-based treatment of social phobia: A randomized controlled trial comparing unguided with two types of guided self-help. Behav Res Ther. 2011. https://doi.org/10.1016/j.brat.2010.12.007.
21. Rodriguez A, Reise SP, Haviland MG. Evaluating bifactor models: Calculating and interpreting statistical indices. Psychol Methods. 2016. https://doi.org/10.1037/met0000045.
22. Beauducel A, Herzberg PY. On the Performance of Maximum Likelihood Versus Means and Variance Adjusted Weighted Least Squares Estimation in CFA. Struct Equ Modeling. 2006. https://doi.org/10.1207/s15328007sem1302_2.
23. Brown TA. Confirmatory factor analysis for applied research. 2nd ed. New York; London: The Guilford Press; 2015.
24. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling. 1999. https://doi.org/10.1080/10705519909540118.
25. Wang J, Wang X. Structural equation modeling: applications using Mplus. 2nd ed. Hoboken, NJ: Wiley; 2020.
26. Kyriazos TA. Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. Psychology. 2018. https://doi.org/10.4236/psych.2018.98126.
27. Gerbing DW, Anderson JC. Monte Carlos Evaluations of Goodness of Fit Indices for Structural Equation Models. Sociol Methods Res. 1992. https://doi.org/10.1177/0049124192021002002.
28. Dombrowski SC, McGill RJ, Canivez GL, Watkins MW, Beaujean AA. Factor Analysis and Variance Partitioning in Intelligence Test Research: Clarifying Misconceptions. J Psychoeduc Assess. 2021. https://doi.org/10.1177/0734282920961952.
29. Reise SP, Bonifay WE, Haviland MG. Scoring and Modeling Psychological Measures in the Presence of Multidimensionality. J Pers Assess. 2013. https://doi.org/10.1080/00223891.2012.725437.
30. Carter MM, Sbrocco T, Tang D, Rekrut FM, Condit C. Psychometric properties of the social phobia and social interaction anxiety scales: Evidence of construct equivalence in an African American sample. J Anxiety Disord. 2014. https://doi.org/10.1016/j.janxdis.2014.07.003.
31. Kaiser HF. An index of factorial simplicity. Psychometrika. 1974. https://doi.org/10.1007/BF02291575.
32. Moosbrugger H, Hartig J. Factor analysis in personality research: Some artefacts and their consequences for psychological assessment. Psychol Beitr. 2000;44(1):136.
33. Dueber DM. Bifactor Indices Calculator: A Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. 2017; https://doi.org/10.13023/edp.tool.01.
34. Caballo VE, Salazar IC, Irurtia MJ, Arias B, Nobre L. The assessment of social anxiety through five sef-report measures, LSAS-SR, SPAI, SPIN, SPS, and SIAS: a critical analysis of their factor structure. Behav Psychol. 2013;21(3):423–48.
35. Caballo VE, Salazar IC, Irurtia MJ, Arias B, Hofmann SG. Measuring Social Anxiety in 11 Countries: Development and Validation of the Social Anxiety Questionnaire for Adults Eur J Psychol Assess. 2010; https://doi.org/10.1027/1015-5759/a000014.

Publisher's Note

Appendix C

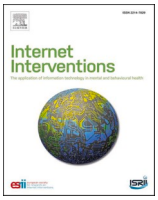
Article III – OPTIMIZE - Study Protocol

Internet Interventions 26 (2021) 100480

Contents lists available at ScienceDirect

Internet Interventions

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



Optimizing cognitive-behavioral therapy for social anxiety disorder and understanding the mechanisms of change: Study protocol for a randomized factorial trial

Rodrigo C.T. Lopes [a,*], Dajana Šipka [a], Tobias Krieger [a], Jan Philipp Klein [b], Thomas Berger [a]

[a] Department of Clinical Psychology and Psychotherapy, University of Bern, Fabrikstrasse, 8, 3012 Bern, Switzerland
[b] Department of Psychiatry, Psychosomatics and Psychotherapy, Luebeck University, Luebeck, Germany

ARTICLE INFO

Keywords:
Internet-based cognitive-behavioral therapy
Social anxiety
Factorial design
Mechanisms of change

ABSTRACT

Background: Social anxiety disorder (SAD) is characterized by a marked fear of negative evaluation in social situations and significant impairments. Even with the most effective treatments, remission rates are around 50%. An important reason for the limited effectiveness of treatments is the lack of evidence-based explanation of how treatments work and what their active ingredients might be. An approach to unpack the active ingredients and mechanisms of treatment is the factorial design.
Objectives: The study is a factorial trial aiming (1) to examine the main effects and interactions for the four main treatment components of internet-based cognitive-behavioral therapy (ICBT) for SAD (i.e., psychoeducation, cognitive restructuring, attentional training, and exposure) and (2) to examine whether and which change mechanisms mediate the relationship between treatment components and symptom reduction.
Methods: A total of 464 adults diagnosed with SAD will be randomized to one of 16 conditions containing combinations of the treatment components. The primary endpoint is SAD symptomatology at eight weeks. Secondary endpoints include symptoms of depression and anxiety, quality of life, and negative effects. Hypothesized change mechanisms are the increase of knowledge about SAD, the decrease of dysfunctional cognitions, the decrease of self-focused attention, and the decrease of avoidance and safety behaviors.
Discussion: A better understanding of the differential efficacy of treatment components and mechanisms of treatment underlying ICBT for SAD might inform clinicians and researchers to plan more potent and scalable treatments.
Trial registration: clinicaltrials.gov (NCT04879641) on June, 11th 2021. https://clinicaltrials.gov/ct2/show/NCT04879641.

1. Introduction

Social anxiety disorder (SAD) is characterized by a marked and persistent fear of negative evaluation in social situations (American Psychiatric Association, 2013) and is a prevalent and disabling disorder across the globe (Stein et al., 2017). Although effective treatments such as psychotherapy and pharmacotherapy are available, far from all individuals suffering from SAD seek and eventually find help (Dalrymple and Zimmerman, 2011). Internet interventions offer many potential benefits, such as providing broader and easier access to empirically supported treatments affordably and conveniently.

The efficacy of internet interventions such as guided self-help interventions has been demonstrated for a variety of mental disorders in many randomized controlled trials (RCTs) and meta-analyses (Andersson et al., 2014; Andersson et al., 2019b; Carlbring et al., 2018; Karyotaki et al., 2017). SAD is probably the disorder for which internet-based guided self-help treatments have the most robust empirical support (Hedman et al., 2016). In this treatment format, patients work their way through a structured self-help program, typically based on CBT manuals (Clark, 2001; Clark and Wells, 1995), and therapists (also referred to as coaches or guides) assist and support them via a secure e-mail system. Overall, the vast majority of the RCTs investigating such Internet-based cognitive-behavioral treatments (ICBT) for SAD reported substantial reductions of social anxiety symptoms (Andersson et al.,

* Corresponding author at: Universität Bern, Institut für Psychologie, Klinische Psychologie und Psychotherapie, Fabrikstrasse 8, 3012 Bern, Switzerland.
E-mail address: rodrigo.lopes@psy.unibe.ch (R.C.T. Lopes).

https://doi.org/10.1016/j.invent.2021.100480
Received 12 August 2021; Received in revised form 8 November 2021; Accepted 9 November 2021
Available online 10 November 2021

2019a; Boettcher et al., 2013) and cost-effectiveness (Hedman et al., 2011). However, and as with conventional face-to-face treatments (Loerinc et al., 2015), there is still much room for improving the efficacy of ICBT for SAD as a considerable number of patients do not recover fully after treatment. The number of participants fulfilling the criterion of clinically significant change (Jacobson and Truax, 1991) at the end of ICBT ranges between 36 and 56% across studies on ICBT for SAD (Boettcher et al., 2013).

An important reason for the limited efficacy of face-to-face and ICBT for SAD is the limited understanding of how these treatments work and what their active ingredients might be. Various reviews convincingly argued that the active ingredients of CBT need to be identified so that therapy can be made more efficacious and probably also briefer (e.g., Holmes et al., 2014; Kazdin, 2017). Identifying the active ingredients requires the use of rigorous study designs that test the presence or absence of individual therapeutic elements rather than conventional parallel-group randomized controlled trials (RCTs). Conventional RCTs are the gold standard for determining if the intervention package works by establishing the relative efficacy of one treatment intervention versus a control group (e.g., another treatment package, attention control, wait-list). However, RCTs have limitations in identifying active ingredients because they usually only compare the overall effect of an entire intervention package. Current evidence-based psychological treatments for SAD, such as CBT, are made of multiple components, including psychoeducation, cognitive restructuring, attention training, and exposure. Each of these treatment components can (a) contribute to a greater or lesser extent to the effect of the treatment package, (b) act via distinct mechanisms, and (c) act via synergistic or antagonistic interactions (Collins, 2018).

As for the specific case of SAD, little effort has been made to answer the question of the differential effects of treatment components in face-to-face settings. Evidence for differential effects of specific treatment components usually comes from underpowered clinical trials with little control over treatment integrity, and results are inconsistent (e.g., Hope et al., 1995; Mattick et al., 1989; Nortje and Posthumus, 2012). The most reliable evidence so far comes from meta-analyses, but the conclusions are still not consistent. For instance, Powers et al. (2008) found that cognitive and behavioral interventions for SAD combined were not significantly more effective than cognitive treatments alone or exposure treatments alone. Also, no significant differences were found in direct comparisons of cognitive techniques alone and exposure alone (Feske and Chambless, 1995; Gil et al., 2001; Powers et al., 2008). These findings conflict with other meta-analytic evidence showing that exposure-based interventions alone yielded the largest effect size, whether alone or combined with cognitive restructuring (Gould et al., 1997). Besides the discrepant findings, research has failed to look into the active components of treatments separately (Acarturk et al., 2009). Thus, it is unclear which intervention components work and which do not and which ones work particularly well together. In line with this, how or why well-studied interventions for SAD produce change is mostly unknown.

The factorial experiment is an efficient and economical way of studying the individual and combined effects of sets of intervention components (Collins et al., 2014; Watkins and Newbold, 2020). However, factorial designs are still rare in psychotherapy research. One of the reasons is that in traditional psychotherapy, it is challenging to clearly demarcate treatment components and avoid an unwanted drift from the treatment protocol by therapists. Standardization and the avoidance of spillover effects (e.g., avoiding therapists using techniques from other treatment components) are essential for a successful factorial experiment. With the advent of internet-delivered treatments, the possibility to successfully realize factorial designs has improved. In this new treatment format, the intervention content can be standardized, and treatment integrity (the degree to which an intervention is implemented as intended) can be controlled (Collins, 2018; Watkins et al., 2016).

Although adding a factor to a factorial experiment does not require the same relatively large increase of the number of participants as an additional treatment arm would require in an RCT (see below; Collins et al., 2014), factorial trials still need quite large sample sizes to have sufficient power. With internet interventions, it is much easier to conduct large trials than in conventional psychotherapy research, with some clinical trials with more than 1000 participants (e.g., Klein et al., 2016), which is one reason why this field has developed at a fast pace (Andersson, 2015). Due to the possibility to control the delivery of standardized treatment components and to run trials with large sample sizes, several research groups (including our own) have recently started to conduct factorial trials to identify the active ingredients of internet interventions (e.g., Berg et al., 2020b; Bur et al., 2021; Watkins et al., 2016). Factorial trials have also been recommended to understand the mechanisms of change because they "provide direct evidence about the effects and interactions of individual components within a treatment package" (Watkins and Newbold, 2020, p. 429).

1.1. Objectives

The primary objective of this trial is to investigate the active ingredients of ICBT for SAD by testing the main effects and interactions for the four main treatment components (i.e., psychoeducation, cognitive restructuring, attention training, and exposure) on primary (i.e., decrease in social anxiety symptoms) and on secondary outcomes (i.e., decrease in depressive symptoms, decrease in general anxiety, increase of quality of life, and client satisfaction). Furthermore, we also aim to investigate the effects of each treatment component on hypothesized change mechanisms and explore whether and which change mechanisms mediate the effect of the treatment components on symptom reduction. The specific secondary objectives (1) to investigate whether the specific mechanisms (i.e., knowledge gain of SAD, decrease of dysfunctional social cognitions, decrease of self-focused attention, decrease of avoidance and safety behaviors) mediate the effect of the treatment components on primary and secondary outcomes, and (2) to address additional exploratory research questions, including examining the negative effects of the treatment components and potential moderators of treatment outcome.

2. Methods

2.1. Study design

The study is a single-center, block randomized, balanced factorial trial with four treatment components (experimental factors), each evaluated at two levels (presence vs absence), resulting in 16 conditions ($2 \times 2 \times 2 \times 2$; see Table 1). Although there are 16 experimental conditions, this study should not be considered a 16-arm RCT (Collins, 2018). The purpose of the factorial experiment is not to compare the 16 conditions to each other but to estimate the main effects of the four treatment components and interactions between the components. Each estimation of the main effects and interactions is based on all the conditions and, therefore, on all participants. For example, the main effect of the cognitive restructuring component will be estimated by comparing the mean of the experimental conditions in which the cognitive restructuring component is present (5, 6, 7, 8, 13, 14, 15, 16 in Table 1) vs the mean of the experimental conditions, in which cognitive restructuring is absent (1, 2, 3, 4, 9, 10, 11, 12 in Table 1). To calculate a two-way interaction (i.e. effect of one component depending on the level of the other factor), one has to calculate the difference between the average effect of one component at the two levels of the other component (present vs absent) and then averaging over all other factors. For a detailed explanation, see Collins (2018).

2.2. Participants

A total of 464 participants with a SAD diagnosis will be included in

Table 1
Experimental conditions of the factorial design, with the presence (yes) and absence (no) of each component.

Condition	Psychoeducation	Cognitive restructuring	Attention training	Exposure
1 WL	No	No	No	No
2	No	No	No	Yes
3	No	No	Yes	No
4	No	No	Yes	Yes
5	No	Yes	No	No
6	No	Yes	No	Yes
7	No	Yes	Yes	No
8	No	Yes	Yes	Yes
9	Yes	No	No	No
10	Yes	No	No	Yes
11	Yes	No	Yes	No
12	Yes	No	Yes	Yes
13	Yes	Yes	No	No
14	Yes	Yes	No	Yes
15	Yes	Yes	Yes	No
16 full	Yes	Yes	Yes	Yes

WL = Wait-list condition. For ethical reasons, participants randomized to condition 1 will be offered treatment after eight weeks.

the study, with 29 participants each assigned to one of the 16 conditions. Participants who return the informed consent will be included in the study if they (1) are 18 years or older; (2) have access to the internet and to a smartphone, PC or tablet; (3) have sufficient knowledge of German; (4) exceed predefined cut-off scores out of two social anxiety measures (22 points on the Social Phobia Scale or 33 points on the Social Interaction Anxiety Scale; SPS & SIAS; German version: Stangier et al., 1999); (5) fulfill the diagnostic criteria of SAD according to a diagnostic telephone interview; (6) in the case of taking psychiatric medication, the treatment is stabilized over one month.

Candidates will be excluded from the study if they (1) score two or higher on the suicide item of the PHQ-9 or show active suicidal plans in the diagnostic telephone interview; (2) have other highly impairing comorbid psychiatric conditions (i.e., history of psychotic or bipolar disorder) and (3) undergo another psychological treatment at the beginning of the study.

2.3. Recruitment

Participants will be recruited using reports in newspapers, flyers, through internet forums, social media (e.g., Facebook), via our study website (https://selfhelp1.psy.unibe.ch/shyne/homepage_interessierte) and our research hub website for internet interventions (http://www.online-therapy.ch/) in German-speaking countries. The link to the study website will also be publicized using Facebook Ads and the Google Ads tool.

2.4. Treatment

The internet-based self-help program (Shyne) is based on the well-established cognitive-behavioral treatment for social anxiety disorder by Clark and Wells (1995). It has been proven efficacious in previous studies in our research hub (Berger et al., 2009, 2011; Schulz et al., 2016; Stolz et al., 2018) as well as in previous studies from other universities and countries (e.g., Boettcher et al., 2012; Chen et al., 2020; Kählke et al., 2019; Kishimoto et al., 2016; Lin et al., 2020). The Shyne program consists of the following four treatment components:

1) *Psychoeducation*: This treatment module delivers (1) detailed evidence-based information on SAD with a focus on maintaining processes (e.g., the vicious cycle of negative thoughts and emotions, cognitions, and behaviors associated with the maintenance of SAD) and (2) a brief overview of the evidence-based CBT strategies to overcome SAD (i.e., psychoeducation about the principles behind cognitive restructuring, attention training and exposure). Additionally, participants are asked to write about their anxiety-inducing situations as well as thoughts, feelings and possible avoidance behaviors associated with the described situations. Participants are encouraged to develop an individual model of their social anxiety symptoms based on the information provided.
2) *Cognitive restructuring*: In this treatment module, participants are instructed to identify and modify dysfunctional and negatively biased assumptions. It includes a thought diary to track negative beliefs in daily routine and exercises to formulate helpful and adaptive thoughts.
3) *Attention training*: In this treatment module, participants are trained to reduce self-focused and biased attention. Audio, video and text-based exercises in which participants learn to intentionally direct the attention away from themselves (i.e., less private self-consciousness) and to be less alert to potentially dangerous external social stimuli (i.e., less public self-consciousness; Duval and Wicklund, 1972; Fenigstein et al., 1975).
4) *Exposure*: In this treatment module, participants are instructed to plan and track in vivo exposures using an exposure diary. Participants are also advised to reduce safety behaviors, overt or covert acts such as avoiding eye contact or rehearsing sentences to prevent a feared outcome.

The four treatment components have the same content throughout all conditions but are slightly changed to make sense when combined in a particular treatment condition. Short, specific psychoeducation is also given as an introduction to each component (e.g., an explanation about the relationship between cognition and emotions in the cognitive restructuring component). Moreover, all participants, independent of the condition, get an introduction module at the beginning and a conclusion module at the end. The introduction module gives an overview of the program and informs the participants about how they can work with it. As a motivational strategy, participants are asked to list their personal goals with the treatment and the expected life changes after overcoming the symptoms of social anxiety. The introduction module has the same content for all conditions. In the conclusion modules, participants are provided with a summary of the steps they should follow and repeat after the program termination. They are also asked to summarize the exercises, thoughts, and behaviors that helped them cope with relapses and were generally perceived as the most helpful. We wrote sixteen different conclusion modules since the summary and recommended repetition are different for every condition.

Shyne can be accessed through a secure website from various devices such as PCs, tablets and smartphones, with each participant having a password-protected account. The program will automatically record participants' usage of Shyne, allowing for an automated measure of treatment adherence and treatment dosage.

2.5. Minimal guidance

The role of the guides is to reinforce independent program use and keep up the participants' motivation and adherence. Participants will also have the opportunity to send questions to their assigned guide throughout the program if they have difficulties with the program. As in other studies (Berger et al., 2009, 2011; Schulz et al., 2016; Stolz et al., 2018), the guides monitor the progress of the participants in the program and contact them via a secured text-based messaging system once a week to provide feedback and encourage further engagement. In case of non-adherence, the guides will remind the participants of the importance of reading the material and doing the exercises proposed by Shyne. Guides will also encourage repetition, especially in conditions with fewer components, to ensure dosage equivalence across all conditions. Minimal guidance will be provided by advanced master's students in clinical psychology and psychotherapy. Guides will be randomly assigned to participants across the conditions.

Training and weekly supervision of the master students for the diagnostic interview and guidance will be provided by the first, second and last author (two licensed psychotherapists and experienced in internet-based guided self-help treatments, and a PhD student in clinical psychology). In the supervision, the chats between participants and guides are reviewed, and it is made sure that guidance is being kept at a minimal level.

2.6. Procedures

The study procedures have been approved by the Ethics Committee of the Canton Bern (KEK Bern 2020-02952) and were registered on clinicaltrials.gov (NCT04879641). After receiving the study information and signing the informed consent, the candidates will be screened for eligibility with self-report measures. Potential participants who fulfill the inclusion criteria will be interviewed by telephone to ascertain whether they meet the criteria for SAD. After checking the inclusion criteria, the eligible participants will be randomized with equal probability to one of the 16 treatment conditions. A permuted block randomization schedule will be created using the *blockrand* package in R (Snow, 2020). The random allocation will be concealed to the investigators and done by the in-built randomization module in the REDCap software (Harris et al., 2009, 2019). Assessments of the primary and secondary outcomes and hypothesized mediators are taken at (1) baseline (pre-treatment), (2) at four weeks after randomization (mid-treatment), (3) eight weeks after randomization (post-treatment), and (4) at six months after randomization (follow-up).

As shown in Table 1, all conditions contain at least one component of ICBT for SAD, except for condition number 1, which is a wait-list control group. For ethical reasons, participants randomly assigned to the wait-list control group will receive the full treatment (i.e., condition 16) after the post-assessment. Participants randomized to one of the active treatment conditions can use the full treatment after the follow-up assessment (Fig. 1).

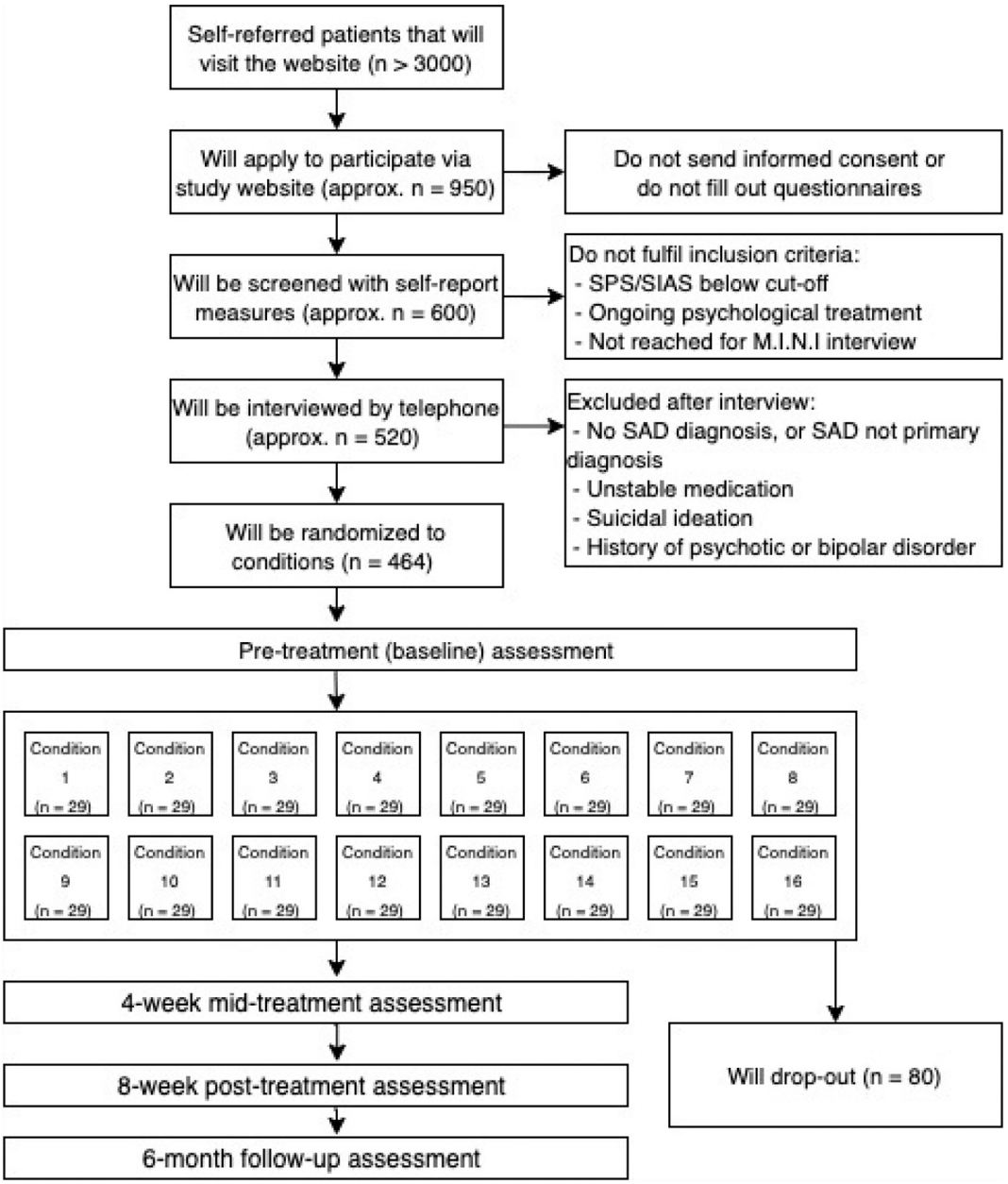


Fig. 1. Design of the study and expected participants' flow.

2.7. Instruments

Table 2 summarizes the instruments and time points of assessment that will be used.

2.7.1. Primary outcome

SAD symptoms after eight weeks are the primary outcome of the study and will be assessed with the composite score of the Social Phobia Scale and the Social Interaction Anxiety Scale (SPS & SIAS; Mattick and Clarke, 1998; German version: Stangier et al., 1999). These two self-report questionnaires complement one another and are usually administered together. The SIAS assesses fears of social interaction (e.g., "I tense up if I meet an acquaintance in the street"), while the SPS focuses on fears of being judged by others (e.g., "I become anxious if I have to write in front of others."). Both scales together consist of 40 items to be rated on a 5-point Likert scale (0 = "not at all" to 4 = "extremely", ranging from 0 to 160 points, where high scores mean more general fear of social interaction). These two companion measures have been found to be valid, reliable and useful for clinical and research purposes. The German version of SIAS has an internal consistency of Cronbach's $\alpha = 0.94$ (Stangier et al., 1999). The German version of the SPS also has an internal consistency of $\alpha = 0.94$ (Stangier et al., 1999). They are both highly correlated with the other social anxiety measures, such as the Social Phobia and Anxiety Inventory and the Liebowitz Social Anxiety Scale (Hoyer and Margraf, 2003; Stangier et al., 1999). Stangier et al. (1999) calculated a cut-off value of 22 (for SPS) and 33 (for SIAS) as a discrimination criterion between German-speaking patients with social anxiety and different comparison groups. The composite score will be the simple average of the *z*-scores of SIAS and SPS, as recommended by Song et al. (2013) for continuous variables.

2.7.2. Secondary outcomes

2.7.2.1. M.I.N.I. International Neuropsychiatric Interview for DSM-IV 6.0.0 (M.I.N.I.; Sheehan et al., 1998; German version: Ackenheil et al., 1999). M.I.N.I. is a brief structured diagnostic interview for assessing psychiatric diagnosis based on the DSM-IV. The specificity of the M.I.N.I was reported as suitable for all diagnoses (ranging from 0.72 to 0.97; 0.81 for SAD; Sheehan et al., 1997). In our study, the M.I.N.I. interview will be administered via telephone. It will assess depressive episodes, suicidality, manic and hypomanic episodes, panic disorder, agoraphobia, SAD, obsessive-compulsive disorder, post-traumatic stress disorder, substance abuse and addiction, psychosis, anorexia and bulimia nervosa and generalized anxiety disorder. The diagnosis of SAD will serve as an eligibility criterion and a secondary outcome measure (an absence of diagnosis at post-treatment and follow-up suggesting treatment success).

2.7.2.2. Patient Health Questionnaire (PHQ-9; Spitzer et al., 1999; German version: Gräfe et al., 2004). Symptoms of depression will be measured with the PHQ-9. This widely used self-report measure consists of nine questions assessing characteristic symptoms of major depression described in DSM-V distributed in nine items on a 4-point Likert scale. Higher scores indicate more severe depression. The German version of the PHQ-9 has also shown good internal consistency (Cronbach's $\alpha = 0.88$; Gräfe et al., 2004; Kroenke et al., 2010; Löwe et al., 2004).

2.7.2.3. Generalized Anxiety Disorder Scale (GAD-7; Spitzer et al., 2006; German version: Löwe et al., 2008). The GAD-7 measures seven general anxiety symptoms (i.e., feeling nervous, worrying, having trouble relaxing, restlessness, feeling annoyed or irritable, and feeling afraid that something awful might happen). Higher scores indicate more severe general anxiety symptoms. The internal consistency of the GAD-7 is good in both the original and German versions (Löwe et al., 2008; Spitzer et al., 2006).

2.7.2.4. Short-Form Health Survey-12 (SF-12; Ware et al., 1996; German version: Gandek et al., 1998). Quality of life is assessed with the SF-12. Its two subscales measure the physical and mental aspects of health-related quality of life. The SF-12 shows good psychometric properties (e.g., internal consistency of $\alpha = 0.83$) and is equivalent to the long-form, the SF-36 (Gandek et al., 1998; Ware et al., 1996).

2.7.2.5. Client Satisfaction Questionnaire (CSQ-8; Attkisson and Zwick, 1982; German version: Schmidt and Wittmann, 2002). The CSQ-8 is a self-report questionnaire that assesses the general level of satisfaction with the service received. It was developed to measure satisfaction with inpatient treatment. The original version shows good internal consistency (Cronbach's $\alpha = 0.91$; Attkisson and Zwick, 1982). In this study, we will use a version that was adapted for internet-based treatments.

2.7.2.6. Negative Effects of the Treatment (INEP; Ladwig et al., 2014). The INEP assesses any adverse effects on social, intrapersonal or work-related situations and whether they are attributed to the intervention. As in other studies, the INEP will be slightly adapted for use within internet-based interventions. The original scale was developed and validated in German and showed good internal consistency (Cronbach's $\alpha = 0.86$; Ladwig et al., 2014).

Table 2
Variables, instruments and time points of assessment.

Dimension	Instrument	Abbreviation	Authors (German version)	Timepoints
Primary outcome measure				
Social anxiety symptoms	Social Phobia Scale & Social Interaction Anxiety Scale	SPS & SIAS	Stangier et al. (1999)	Pre, Mid, Post, FU
Secondary outcome measures				
Depressive symptoms	Patient Health Questionnaire	PHQ-9	Gräfe et al. (2004)	Pre, Mid, Post, FU
General anxiety symptoms	Generalized Anxiety Disorder Scale	GAD-7	Löwe et al. (2008)	Pre, Mid, Post, FU
Quality of life	SF-12 Health Survey	SF-12	Gandek et al. (1998)	Pre, Mid, Post, FU
Client satisfaction	Client Satisfaction Questionnaire	CSQ-8	Schmidt and Wittmann (2002)	Post
Negative effects	Negative Effects of the Treatment & Symptom Deterioration	INEP	Ladwig et al. (2014)	Mid, Post, FU
Diagnoses	MINI. Neuropsychiatric Interview	MINI 6.0.0	Sheehan et al. (1998)	Pre, Post, FU
Hypothesized change mechanisms				
Knowledge of SAD	Knowledge of SAD test	KSAD	Andersson et al. (2012)	Pre, Mid, Post, FU
Dysfunctional social cognitions	Social Cognitions Questionnaire	SCQ	Stangier et al. (1997)	Pre, Mid, Post, FU
Self-focused attention	Self-Consciousness Scale	SCS	Filipp and Freudenberg (1989)	Pre, Mid, Post, FU
Fear and avoidance	Liebowitz Social Anxiety scale	LSAS-SR	Stangier and Heidenreich (2004)	Pre, Mid, Post, FU
Safety behaviors	Social Behaviors Questionnaire	SBQ	Stangier et al. (1996)	Pre, Mid, Post, FU

Notes. Pre = baseline; Mid = mid-treatment (4 weeks after baseline); Post = post-treatment (8 weeks after baseline); FU = follow-up (6 months after baseline).

2.7.2.7. Adherence. Following the suggestion of Donkin et al. (2011), a composite score to measure adherence and dosage will be created by averaging the *z*-scores of several variables: time spent in the intervention, the number of modules completed, the number of exercises completed, and the number of clicks in the intervention.

In addition, socio-demographic variables such as age, gender, country of origin, parent's country of origin, mother tongue, relationship status, educational level and employment status are assessed.

2.7.3. Assessment of hypothesized mechanisms of change

The secondary aim of this study is to better understand which mechanisms of change mediate the relationship between treatment components and symptom reduction. For this, we will also assess variables hypothesized to mediate change for every treatment component. Fig. 2 shows a conceptual model of the expected effects of the four treatment components on the hypothesized change mechanisms and primary and secondary outcomes.

2.7.3.1. Knowledge of SAD test (KSAD; Andersson et al., 2012; Berg et al., 2020a). The KSAD assesses basic knowledge around the condition of SAD and its treatment. It includes 11 questions, each with one correct answer (out of three possible choices). In addition, each response is rated in terms of how confident the participant is about the response with three response options (Guessing, Pretty Certain, Confident). A higher score indicates more knowledge. Questions cover the content of the psychoeducation component of the Shyne program (e.g., the definition of SAD, general principles of CBT, safety behaviors, avoidance, negative automatic thoughts, attentional shift, exposure). Scores of knowledge of SAD are calculated in two ways: (a) a total score based on the total number of correct answers and (b) a weighted total score in which certainty of answers was factored in. Reliability analyses showed a low Cronbach's alpha of $\alpha = 0.40$ for the raw scores, a high Cronbach's alpha of $\alpha = 0.86$ for the certainty ratings, and an alpha of $\alpha = 0.56$ for the weighted scores (Andersson et al., 2012).

2.7.3.2. Social cognitions questionnaire (SCQ; Wells et al., 1993; German version: Stangier et al., 1996b). The SCQ is a self-rating scale that assesses typical negative social cognitions of socially anxious individuals. It is composed of 22 items, ranging from 22 to 110, grouped in three subscales ("negative self", "performance anxiety", and "fear of showing bodily symptoms"). Higher scores mean more negative social cognitions. The Cronbach's α for the whole German version scale is $\alpha = 0.89$.

2.7.3.3. Self-Consciousness Scale (SCS; Fenigstein et al., 1975; German version: Filipp and Freudenberg, 1989). The SCS measures self-focused attention (or self-consciousness) in two dimensions: private self-consciousness and public self-consciousness. The German version consists of 27 items, which are rated from 1 ("very rarely") to 5 ("very often"). Higher scores indicate more self-focused attention. Both subscales have shown good internal consistency (Cronbach's $\alpha = 0.87$ and $\alpha = 0.86$, respectively; Hinz et al., 2010).

2.7.3.4. Liebowitz Social Anxiety Scale, self-report (LSAS-SR; Baker et al., 2002; German version: Stangier and Heidenreich, 2004). The LSAS-SR measures SAD symptoms. It comprises 24 items, divided into two subscales (anxiety and avoidance, 12 items each) scored on a Likert-type scale of four points and is rated in terms of frequency (never, occasionally, often and usually). In this study, and as a measure of a hypothesized mechanism of the change, only the avoidance subscale will be used. LSAS-SR shows good internal consistency ($\alpha = 0.96$ for the total scale and $\alpha = 0.92$ for the avoidance scale).

2.7.3.5. Social Behaviors Questionnaire (SBQ; Clark et al., 1995; German version: Stangier et al., 1996a). The SBQ assesses the use of safety behaviors in social situations with 27 items. The frequency of each behavior is rated on a 4-point scale (from 0 = never to 3 = always). "Avoid eye contact", "try to control shaking", "rehearse sentences in your mind" are examples of safety behavior assessed by the SBQ. The items on the SBQ are a mixture of discrete behaviors (e.g., hide your face, grip glasses tightly) and broad strategies (e.g., make an effort to come across well, try not to attract attention). Studies with adult populations revealed acceptable internal consistencies (Cronbach's $\alpha = 0.69$).

2.8. Sample size

The current study is powered for the first and primary research question, i.e., the main and interaction effects of the treatment components on the decrease of social anxiety symptoms. In the a priori power analysis, we assumed that the smallest clinically relevant difference

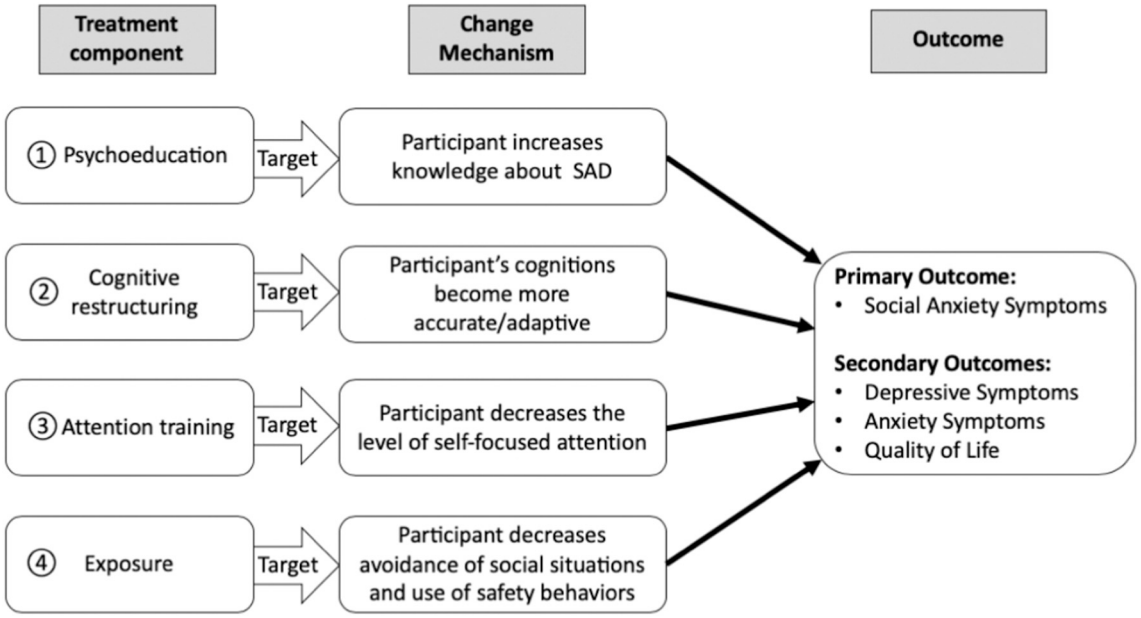


Fig. 2. Simplified conceptual model of the effects of the four treatment components on the hypothesized change mechanisms and primary and secondary outcomes.

would be a small effect size of Cohen's $d = 0.2$ for the main effect of an individual treatment component or interaction between components on pre-to-post change on social anxiety symptoms. Smaller effect sizes would be of little clinical interest and value. At an α level of 0.05, a statistical power (1-Beta) of 0.80, and a correlation between measurements with around 4- and 8-weeks interval of $r = 0.50$, based on our experience with clinical trials for SAD, we need a total of 384 participants (G*Power; Faul et al., 2007). Based on the finding that in 15 out of 17 studies, dropout rates (i.e., not providing assessment data at post-treatment) for ICBT for SAD were below 13% (Boettcher et al., 2013), we conservatively estimate a dropout rate of 20% for our study ($n = 77$ participants). Thus, we aim for a sample size of 464 participants, which results in 29 individuals per treatment condition.

As mentioned above, the logic behind how an experiment is powered differs between RCTs and factorial experiments. An RCT would compare the 16 conditions to each other, and power would be reflected in the per-condition sample size. If an RCT has a small per-condition sample size, such as 29 individuals per treatment condition, it does not have enough power to detect small effect sizes. By contrast, in factorial experiments, all participants receiving a specific component (e.g., cognitive restructuring, which is present in half of the conditions, i.e., in eight conditions) can be compared to participants who do not receive that component (also eight conditions, with 29 participants per condition). That yields a sample size of $n = 232$ per component (Collins, 2018). Since our study contains only components with two levels (absent vs present), the sample size to maintain the power to detect main effects and interactions is the same (Collins, 2018). For a detailed explanation of how factorial experiments maintain power to estimate main effects and interactions, see Collins (2018).

2.9. Statistical analyses

Reporting will follow CONSORT E-Health standards (Eysenbach and Consort-Ehealth Group, 2011). The primary outcome is the change in the composite score of SPS & SIAS from baseline to eight weeks (post-treatment). The analyses are carried out on the basis of the intention-to-treat approach (ITT; i.e., using all randomized participants). Our primary interest is in testing the main effects and interactions. For that, we will use linear mixed models repeated measures analysis of variance (ANOVA). This approach uses all available data on each subject and does not require the imputation of missing values but estimates parameters about missing values. Furthermore, mixed models account for the correlation between the repeated measurements. Main effects and interactions are calculated based on aggregates across experimental conditions. The levels of the factors will be represented numerically by −1 (absence of a component in a condition) and +1 (presence of a component in a condition), as recommended by Collins (2018). Significance testing of dichotomous data such as diagnostic status will be conducted with chi-square tests. Sensitivity analysis will be conducted to analyze the impact of dropouts on our results.

We will test mediation of the hypothesized change mechanisms (i.e., knowledge gain of SAD, decrease of dysfunctional social cognitions, decrease of self-focused attention, decrease of avoidance and safety behaviors) of the effect of the treatment components on primary and secondary outcomes (see Fig. 2 above). We will test mediation of the hypothesized change mechanisms by using an approach that allows multiple mediators in one model, as set out by Kraemer et al. (2002).

In addition, we will explore potential moderation of the treatment components by various measured variables (i.e., age, gender, country of origin, nationality, country of parents, mother tongue, relationship status, educational level, employment status, presence of comorbid disorder, use of medication, the severity of SAD). For the analysis of potential moderators, factorial ANOVA and multiple regression analysis will be used.

3. Discussion

Far from all individuals suffering from SAD seek and eventually find help, and far from all SAD patients respond fully to current evidence-based treatments. Low-threshold and cost-effective internet-based interventions can easily be distributed and flexibly used, representing a promising alternative to face-to-face therapy. With an optimized internet-based intervention, a broader population of people suffering from SAD can be reached at even lower costs and more effectiveness. The results of this trial are expected to improve current evidence-based treatments for SAD and increase the number of SAD patients fully responding to ICBT. If we know more about the active ingredients of CBT for SAD, we can probably identify better and briefer strategies that trigger change processes. Thus, understanding active ingredients and change mechanisms can optimize change and "build more potent, scalable, and efficient treatments" (Watkins et al., 2016, p. 2) of SAD.

We understand the use of the factorial trial as an appropriate approach to understand the differential effects of each component of SAD. Although other sophisticated approaches exist, for instance, component individual patient data meta-analysis (e.g., Furukawa et al., 2021), the component meta-analyses are based on the indirect comparisons between different trials. Thus, there is a higher likelihood that the observed differences can be attributed not to the various components but to the differences in the settings. Furthermore, the OPTIMIZE trial is planned to have a reasonably high sample size and sufficient power to detect even small changes.

Some potential limitations of this study should be addressed. The treatment dosage may vary across conditions and be lower in those conditions with fewer components. To prevent high variations in treatment dosage, the participants are encouraged by the program and by the guides to repeat the exercise. To address this potential limitation, we will control the overall treatment dosage of each participant by assessing adherence to the program. Also, there might be a spillover effect from the psychoeducation component once we briefly explain broad change principles used in established CBT treatment (i.e., cognitive restructuring, attention training and exposure). However, we do not offer any practical indication of implementing those techniques, and we do not provide any access to the exercises introduced in the relevant components.

Finally, the measure of knowledge gain (KSAD; Andersson et al., 2012; Berg et al., 2020a) might represent a limitation since the original authors have not found satisfactory reliability. We will replicate the reliability and test-retest analysis with our sample using all time points available to re-evaluate the ability of this scale to capture change in knowledge gain.

We aim to perform mediation analyses to test the hypothesized mechanisms of change. The mere statistical mediation is not enough to ascertain a mechanism of change (Kazdin, 2007). The field needs to show a solid theoretical foundation for a specific mediator and also strong statistical association, temporality (i.e. timeline shows that intervention leads to change in mediator which leads to change in outcome, not the other way around), specificity (i.e., to prove that one particular mediator is responsible for change), gradient (i.e., dose-response relationship between mediator and outcome), consistency (across studies with different samples) and coherence with other evidence, for instance, results coming from experimental studies (Kazdin, 2007). In our study, we will be able to assess temporality, specificity and the gradient of theoretically founded mechanisms of change. Although other studies will be needed to evaluate consistency and coherence with other experimental evidence, our results might move forward theoretical debates regarding the mechanisms involved in the maintenance of SAD and what works in treatments. We estimate that this trial's results will inform the treatment of social anxiety via internet interventions and inform face-to-face treatments. At a societal level, optimizing treatment and expanding the knowledge about mechanisms of change is essential because SAD is very common and one of the costliest psychiatric

conditions (e.g., Fehm et al., 2005). By determining the importance of each component to the overall efficacy of CBT treatment for SAD, we will be able to inform mental health policy decisions that would probably decrease its costs and increase its effectiveness.

Declaration of competing interest

All the authors have no financial or scientific competing interests to declare.

Acknowledgments

The study is entirely funded by the Swiss National Science Foundation (SNSF) through a Project Funding grant in humanities and social sciences (division I), given to Prof. Dr Thomas Berger, with the reference number 10001C_197475. The funding source has no role in the study's design, data collection, analysis, interpretation of data, writing of the manuscript, nor in the decision to submit this report.

References

Acarturk, C., Cuijpers, P., van Straten, A., de Graaf, R., 2009. Psychological treatment of social anxiety disorder: a meta-analysis. Psychol. Med. 39 (2), 241–254. https://doi.org/10.1017/S0033291708003590.

Ackenheil, M., Stotz-Ingenlath, G., Dietz-Bauer, R., Vossen, A., 1999. M.I.N.I. Mini International Neuropsychiatric Interview, German version 5.0. 0 DSM IV. Psychiatric University Clinic, Munich.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. Amer Psychiatric Pub Incorporated.

Andersson, G., 2015. The Internet and CBT: A Clinical Guide, 1st ed. Taylor & Francis Group.

Andersson, G., Carlbring, P., Furmark, T., S. O. F. I. E. Research Group, 2012. Therapist experience and knowledge acquisition in internet-delivered CBT for social anxiety disorder: a randomized controlled trial. PLoS ONE 7 (5), e37411. https://doi.org/10.1371/journal.pone.0037411.

Andersson, G., Carlbring, P., Rozental, A., 2019. Response and remission rates in internet-based cognitive behavior therapy: an individual patient data meta-analysis. Front. Psych. 10, 749. https://doi.org/10.3389/fpsyt.2019.00749.

Andersson, G., Carlbring, P., Titov, N., Lindefors, N., 2019. Internet interventions for adults with anxiety and mood disorders: a narrative umbrella review of recent meta-analyses. Can. J. Psychiatry 64 (7), 465–470. https://doi.org/10.1177/0706743719839381.

Andersson, G., Cuijpers, P., Carlbring, P., Riper, H., Hedman, E., 2014. Guided internet-based vs. Face-to-Face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 13 (3), 288–295. https://doi.org/10.1002/wps.20151.

Attkisson, C.C., Zwick, R., 1982. The client satisfaction questionnaire: psychometric properties and correlations with service utilization and psychotherapy outcome. Eval. Program Plann. 5 (3), 233–237. https://doi.org/10.1016/0149-7189(82)90074-X.

Baker, S.L., Heinrichs, N., Kim, H.-J., Hofmann, S.G., 2002. The liebowitz social anxiety scale as a self-report instrument: a preliminary psychometric analysis. Behav. Res. Ther. 40 (6), 701–715. https://doi.org/10.1016/S0005-7967(01)00060-2.

Berg, M., Andersson, G., Rozental, A., 2020. Knowledge about treatment, anxiety, and depression in association with internet-based cognitive behavioral therapy for adolescents: development and initial evaluation of a new test. SAGE Open 10 (1), 2158244019899095. https://doi.org/10.1177/2158244019899095.

Berg, M., Rozental, A., de Brun Mangs, J., Näsman, M., Strömberg, K., Viberg, L., Wallner, E., Åhman, H., Silfvernagel, K., Zetterqvist, M., Topooco, N., Capusan, A., Andersson, G., 2020. The role of learning support and chat-sessions in guided internet-based cognitive behavioral therapy for adolescents with anxiety: a factorial design study. Front. Psych. 11 https://doi.org/10.3389/fpsyt.2020.00503.

Berger, T., Caspar, F., Richardson, R., Kneubühler, B., Sutter, D., Andersson, G., 2011. Internet-based treatment of social phobia: a randomized controlled trial comparing unguided with two types of guided self-help. Behav. Res. Ther. 49 (3), 158–169. https://doi.org/10.1016/j.brat.2010.12.007.

Berger, T., Hohl, E., Caspar, F., 2009. Internet-based treatment for social phobia: a randomized controlled trial. J. Clin. Psychol. 65 (10), 1021–1035. https://doi.org/10.1002/jclp.20603.

Boettcher, J., Berger, T., Renneberg, B., 2012. Does a pre-treatment diagnostic interview affect the outcome of internet-based self-help for social anxiety disorder? A randomized controlled trial. Behav. Cogn. Psychother. 40 (05), 513–528. https://doi.org/10.1017/S1352465812000501.

Boettcher, J., Carlbring, P., Renneberg, B., Berger, T., 2013. Internet-based interventions for social anxiety disorder—an overview. Verhaltenstherapie 23 (3), 160–168. https://doi.org/10.1159/000354747.

Bur, O.T., Krieger, T., Moritz, S., Klein, J.P., Berger, T., 2021. Optimizing the context of support to improve outcomes of internet-based self-help in individuals with depressive symptoms: protocol for a randomized factorial trial. JMIR Res. Protoc. 10 (2), e21207.

Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., Hedman-Lagerlöf, E., 2018. Internet-based vs. Face-to-Face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn. Behav. Ther. 47 (1), 1–18. https://doi.org/10.1080/16506073.2017.1401115.

Chen, H., Rodriguez, M.A., Qian, M., Kishimoto, T., Lin, M., Berger, T., 2020. Predictors of treatment outcomes and adherence in internet-based cognitive behavioral therapy for social anxiety in China. Behav. Cogn. Psychother. 1–13 https://doi.org/10.1017/S1352465819000730.

Clark, D.M., 2001. A cognitive perspective on social phobia. In: Crozier, W.R., Alden, L.E. (Eds.), International Handbook of Social Anxiety: Concepts, Research and Intervention Relating to the Self and Shyness. John Wiley & Sons, Ltd, pp. 405–406.

Clark, D.M., Wells, A., 1995. A cognitive model of social phobia. In: Heimberg, R.G., Liebowitz, M.R., Hope, D., Schneider, F. (Eds.), Social Phobia: Diagnosis, Assessment, and Treatment. Guilford Press, pp. 69–93.

Clark, D.M., Wells, A., Salkoviskis, P., Hackmann, A., 1995. Social Behaviour Questionnaire [Unpublished Manuscript].

Collins, L.M., 2018. Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST). Springer International Publishing. https://doi.org/10.1007/978-3-319-72206-1.

Collins, L.M., Dziak, J.J., Kugler, K.C., Trail, J.B., 2014. Factorial experiments: efficient tools for evaluation of intervention components. Am. J. Prev. Med. 47 (4), 498–504. https://doi.org/10.1016/j.amepre.2014.06.021.

Dalrymple, K.L., Zimmerman, M., 2011. Treatment-seeking for social anxiety disorder in a general outpatient psychiatry setting. Psychiatry Res. 187 (3), 375–381. https://doi.org/10.1016/j.psychres.2011.01.004.

Donkin, L., Christensen, H., Naismith, S.L., Neal, B., Hickie, I.B., Glozier, N., 2011. A systematic review of the impact of adherence on the effectiveness of e-therapies. J. Med. Internet Res. 13 (3), e1772 https://doi.org/10.2196/jmir.1772.

Duval, S., Wicklund, R.A., 1972. In: A Theory of Objective Self Awareness. Academic Press, p. 238 pp. x.

Eysenbach, G., Consort-Ehealth Group, 2011. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J. Med. Internet Res. 13 (4), e126 https://doi.org/10.2196/jmir.1923.

Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39 (2), 175–191.

Fehm, L., Pelissolo, A., Furmark, T., Wittchen, H.-U., 2005. Size and burden of social phobia in Europe. Eur. Neuropsychopharmacol. 15 (4), 453–462. https://doi.org/10.1016/j.euroneuro.2005.04.002.

Fenigstein, A., Scheier, M.F., Buss, A.H., 1975. Public and private self-consciousness: assessment and theory. J. Consult. Clin. Psychol. 43 (4), 522–527. https://doi.org/10.1037/h0076760.

Feske, U., Chambless, D.L., 1995. Cognitive behavioral versus exposure only treatment for social phobia: a meta-analysis. Behav. Ther. 26 (4), 695–720. https://doi.org/10.1016/S0005-7894(05)80040-1.

Filipp, S.H., Freudenberg, E., 1989. Fragebogen zur Erfassung dispositionaler Selbstaufmerksamkeit. Hogrefe.

Furukawa, T.A., Suganuma, A., Ostinelli, E.G., Andersson, G., Beevers, C.G., Shumake, J., Berger, T., Boele, F.W., Buntrock, C., Carlbring, P., Choi, I., Christensen, H., Mackinnon, A., Dahne, J., Huibers, M.J.H., Ebert, D.D., Farrer, L., Forand, N.R., Strunk, D.R., Cuijpers, P., 2021. Dismantling, optimising, and personalising internet cognitive behavioural therapy for depression: a systematic review and component network meta-analysis using individual participant data. Lancet Psychiatry 8 (6), 500–511. https://doi.org/10.1016/S2215-0366(21)00077-8.

Gandek, B., Ware, J.E., Aaronson, N.K., Apolone, G., Bjorner, J.B., Brazier, J.E., Bullinger, M., Kaasa, S., Leplege, A., Prieto, L., Sullivan, M., 1998. Cross-validation of item selection and scoring for the SF-12 health survey in nine countries: results from the IQOLA project. International quality of life assessment. J. Clin. Epidemiol. 51 (11), 1171–1178. https://doi.org/10.1016/s0895-4356(98)00109-7.

Gil, P.J.M., Carrillo, F.X.M., Meca, J.S., 2001. Effectiveness of cognitive-behavioral treatment in social phobia: a meta-analytic review. Psychol. Spain 5 (1), 17–25.

Gould, R.A., Buckminster, S., Pollack, M.H., Otto, M.W., Massachusetts, L.Y., 1997. Cognitive-behavioral and pharmacological treatment for social phobia: a meta-analysis. Clin. Psychol. Sci. Pract. 4 (4), 291–306. https://doi.org/10.1111/j.1468-2850.1997.tb00123.x.

Gräfe, K., Zipfel, S., Herzog, W., Löwe, B., 2004. Screening psychischer Störungen mit dem "Gesundheitsfragebogen für patienten (PHQ-D)": ergebnisse der deutschen validierungsstudie. [Screening for psychiatric disorders with the patient health questionnaire (PHQ). Results from the german validation study.]. Diagnostica 50 (4), 171–181. https://doi.org/10.1026/0012-1924.50.4.171.

Harris, P.A., Taylor, R., Minor, B.L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., Duda, S.N., 2019. The REDCap consortium: building an international community of software platform partners. J. Biomed. Inform. 95, 103208 https://doi.org/10.1016/j.jbi.2019.103208.

Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., Conde, J.G., 2009. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J. Biomed. Inform. 42 (2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010.

Hedman, E., Andersson, E., Ljótsson, B., Andersson, G., Rück, C., Lindefors, N., 2011. Cost-effectiveness of internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: results from a randomized controlled trial. Behav. Res. Ther. 49 (11), 729–736. https://doi.org/10.1016/j.brat.2011.07.009.

Hedman, E., Botella, C., Berger, T., 2016. Internet-based cognitive behavior therapy for social anxiety disorder. In: Lindefors, N., Andersson, G. (Eds.), Guided Internet-Based

Treatments in Psychiatry. Springer International Publishing, pp. 53–78. https://doi.org/10.1007/978-3-319-06083-5_4.
Hinz, A., Blaser, G., Schmutzer, G., Bailer, H., Grulke, N., Brähler, E., Albani, C., 2010. Überprüfung und Normierung des "Fragebogen zur Erfassung dispositionaler Selbstaufmerksamkeit"(SAM) an einer repräsentativen deutschen Bevölkerungsstichprobe. Klinische Diagnostik Und Evaluation 3, 382–395.
Holmes, E.A., Craske, M.G., Graybiel, A.M., 2014. Psychological treatments: a call for mental-health science. Nat. News 511 (7509), 287. https://doi.org/10.1038/511287a.
Hope, D.A., Heimberg, R.G., Bruch, M.A., 1995. Dismantling cognitive-behavioral group therapy for social phobia. Behav. Res. Ther. 33 (6), 637–650. https://doi.org/10.1016/0005-7967(95)00013-N.
Hoyer, J., Margraf, J., 2003. Fragebögen und ratingskalen zur sozialen phobie und zu sozialen Ängsten. In: Hoyer, J., Margraf, J. (Eds.), Angstdiagnostik: Grundlagen und Testverfahren. Springer, pp. 233–319. https://doi.org/10.1007/978-3-642-18976-0_10.
Jacobson, N.S., Truax, P., 1991. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J. Consult. Clin. Psychol. 59 (1), 12–19.
Kählke, F., Berger, T., Schulz, A., Baumeister, H., Berking, M., Auerbach, R.P., Bruffaerts, R., Cuijpers, P., Kessler, R.C., Ebert, D.D., 2019. Efficacy of an unguided internet-based self-help intervention for social anxiety disorder in university students: a randomized controlled trial. Int. J. Methods Psychiatr. Res. 28 (2), e1766 https://doi.org/10.1002/mpr.1766.
Karyotaki, E., Riper, H., Twisk, J., Hoogendoorn, A., Kleiboer, A., Mira, A., Mackinnon, A., Meyer, B., Botella, C., Littlewood, E., Andersson, G., Christensen, H., Klein, J.P., Schröder, J., Bretón-López, J., Scheider, J., Griffiths, K., Farrer, L., Huibers, M.J.H., Cuijpers, P., 2017. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry 74 (4), 351–359. https://doi.org/10.1001/jamapsychiatry.2017.0044.
Kazdin, A.E., 2007. Mediators and mechanisms of change in psychotherapy research. Annu. Rev. Clin. Psychol. 3, 1–27.
Kazdin, A.E., 2017. Addressing the treatment gap: a key challenge for extending evidence-based psychosocial interventions. Behav. Res. Ther. 88, 7–18. https://doi.org/10.1016/j.brat.2016.06.004.
Kishimoto, T., Krieger, T., Berger, T., Qian, M., Chen, H., Yang, Y., 2016. Internet-based cognitive behavioral therapy for social anxiety with and without guidance compared to a wait list in China: a propensity score study. Psychother. Psychosom. 85 (5), 317–319. https://doi.org/10.1159/000446584.
Klein, J.P., Berger, T., Schröder, J., Späth, C., Meyer, B., Caspar, F., Lutz, W., Arndt, A., Greiner, W., Gräfe, V., Hautzinger, M., Fuhr, K., Rose, M., Nolte, S., Löwe, B., Andersson, G., Vettorazzi, E., Moritz, S., Hohagen, F., 2016. Effects of a psychological internet intervention in the treatment of mild to moderate depressive symptoms: results of the EVIDENT study, a randomized controlled trial. Psychother. Psychosom. 85 (4), 218–228. https://doi.org/10.1159/000445355.
Kraemer, H.C., Wilson, G.T., Fairburn, C.G., Agras, W.S., 2002. Mediators and moderators of treatment effects in randomized clinical trials. Arch. Gen. Psychiatry 59 (10), 877. https://doi.org/10.1001/archpsyc.59.10.877.
Kroenke, K., Spitzer, R.L., Williams, J.B.W., Löwe, B., 2010. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen. Hosp. Psychiatry 32 (4), 345–359. https://doi.org/10.1016/j.genhosppsych.2010.03.006.
Ladwig, I., Rief, W., Nestoriuc, Y., 2014. What are the risks and side effects of psychotherapy? - Development of an inventory for the assessment of negative effects of psychotherapy (INEP). Verhaltenstherapie 24 (4), 252–263 doi:0.1159/000367928.
Lin, L.-Y., Wang, K., Kishimoto, T., Rodriguez, M., Qian, M., Yang, Y., Zhao, Q., Berger, T., Tian, C., 2020. An internet-based intervention for individuals with social anxiety and different levels of taijin kyofusho in China. J. Cross-Cult. Psychol. 51 (5), 387–402. https://doi.org/10.1177/0022022120920720.
Loerinc, A.G., Meuret, A.E., Twohig, M.P., Rosenfield, D., Bluett, E.J., Craske, M.G., 2015. Response rates for CBT for anxiety disorders: need for standardized criteria. Clin. Psychol. Rev. 42, 72–82. https://doi.org/10.1016/j.cpr.2015.08.004.
Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., Herzberg, P.Y., 2008. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. Med. Care 46 (3), 266–274. JSTOR.
Löwe, B., Kroenke, K., Herzog, W., Gräfe, K., 2004. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the patient health questionnaire (PHQ-9). J. Affect. Disord. 81 (1), 61–66. https://doi.org/10.1016/S0165-0327(03)00198-8.
Mattick, R.P., Clarke, J.C., 1998. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behav. Res. Ther. 36 (4), 455–470. https://doi.org/10.1016/s0005-7967(97)10031-6.
Mattick, R.P., Peters, L., Clarke, J.C., 1989. Exposure and cognitive restructuring for social phobia: a controlled study. Behav. Ther. 20 (1), 3–23. https://doi.org/10.1016/S0005-7894(89)80115-7.
Nortje, C., Posthumus, T., 2012. Scores on an emotional stroop task after treatment of social anxiety disorder. Psychol. Rep. 111 (2), 461–471. https://doi.org/10.2466/02.15.20.PR0.111.5.461-471.
Powers, M.B., Sigmarsson, S.R., Emmelkamp, P.M.G., 2008. A meta-analytic review of psychological treatments for social anxiety disorder. Int. J. Cogn. Ther. 1 (2), 94–113. https://doi.org/10.1521/ijct.2008.1.2.94.
Schmidt, J., Wittmann, W.W., 2002. Fragebogen zur messung der patientenzufriedenheit. In: Brähler, E., Schumacher, J., Strauß, B. (Eds.), Diagnostische verfahren in der psychotherapie. Hogrefe, pp. 392–396.
Schulz, A., Stolz, T., Vincent, A., Krieger, T., Andersson, G., Berger, T., 2016. A sorrow shared is a sorrow halved? A three-arm randomized controlled trial comparing internet-based clinician-guided individual versus group treatment for social anxiety disorder. Behav. Res. Ther. 84, 14–26. https://doi.org/10.1016/j.brat.2016.07.001.
Sheehan, D.V., Lecrubier, Y., Harnett Sheehan, K., Janavs, J., Weiller, E., Keskiner, A., Schinka, J., Knapp, E., Sheehan, M.F., Dunbar, G., 1997. The validity of the Mini international neuropsychiatric interview (MINI) according to the SCID-P and its reliability. Eur. Psychiatry 12 (5), 232–241. https://doi.org/10.1016/S0924-9338(97)83297-X.
Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59 (Suppl. 20), 22–33. Scopus.
Snow, G., 2020. blockrand: randomization for block random clinical trials (1.5) [computer software]. https://CRAN.R-project.org/package=blockrand.
Song, M.-K., Lin, F.-C., Ward, S.E., Fine, J.P., 2013. Composite variables: when and how. Nurs. Res. 62 (1), 45–49. https://doi.org/10.1097/NNR.0b013e3182741948.
Spitzer, R.L., Kroenke, K., Williams, J.B., 1999. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. JAMA 282 (18), 1737–1744. https://doi.org/10.1001/jama.282.18.1737.
Spitzer, R.L., Kroenke, K., Williams, J.B.W., Löwe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med. 166 (10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092.
Stangier, U., Heidenreich, T., 2004. Die Liebowitz Soziale Angst-Skala (LSAS) [The Liebowitz Social Anxiety Scale (LSAS)]. In: Internationale Skalen für Psychiatrie. Collegium Internationale Psychiatriae Scalarum. Hogrefe Verlag, pp. 299–306. https://www.hogrefe.com/de/shop/internationale-skalen-fuer-psychiatrie-66480.html.
Stangier, U., Heidenreich, T., Berardi, A., Golbs, U., Hoyer, J., 1999. Die erfassung sozialer phobie durch die social interaction anxiety scale (SIAS) und die social phobia scale (SPS). Z. Klin. Psychol. Psychother. 28 (1), 28–36. https://doi.org/10.1026//0084-5345.28.1.28.
Stangier, U., Heidenreich, T., Ehlers, A., Clark, D.M., 1996. Fragebogen zu Sicherheitsverhalten in sozialen Situationen [Unpublished Manuscript].
Stangier, U., Heidenreich, T., Ehlers, A., Clark, D.M., 1996. Fragebogen zu sozialphobischen Kognitionen (SPK). Unpublished Manuscript.
Stein, D.J., Lim, C.C.W., Roest, A.M., de Jonge, P., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Benjet, C., Bromet, E.J., Bruffaerts, R., de Girolamo, G., Florescu, S., Gureje, O., Haro, J.M., Harris, M.G., He, Y., Hinkov, H., Horiguchi, I., Hu, C., WHO World Mental Health Survey Collaborators, 2017. The cross-national epidemiology of social anxiety disorder: data from the World Mental Health Survey Initiative. BMC Med. 15 (1), 143. https://doi.org/10.1186/s12916-017-0889-2.
Stolz, T., Schulz, A., Krieger, T., Vincent, A., Urech, A., Moser, C., Westermann, S., Berger, T., 2018. A mobile app for social anxiety disorder: a three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. J. Consult. Clin. Psychol. 86 (6), 493–504. https://doi.org/10.1037/ccp0000301.
Ware, J.E., Kosinski, M., Keller, S.D., 1996. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med. Care 34 (3), 220–233.
Watkins, E.R., Newbold, A., 2020. Factorial designs help to understand how psychological therapy works. Front. Psych. 11, 429. https://doi.org/10.3389/fpsyt.2020.00429.
Watkins, E.R., Newbold, A., Tester-Jones, M., Javaid, M., Cadman, J., Collins, L.M., Graham, J., Mostazir, M., 2016. Implementing multifactorial psychotherapy research in online virtual environments (IMPROVE-2): study protocol for a phase III trial of the MOST randomized component selection method for internet cognitive-behavioural therapy for depression. BMC Psychiatry 16 (1). https://doi.org/10.1186/s12888-016-1054-8.
Wells, A., Stopa, L., Clark, D.M., 1993. The Social Cognitions Questionnaire. Department of Psychiatry, University of Oxford, Warnerford Hospital, Oxford, UK. Unpublished.

Appendix D

Article IV – OPTIMIZE - Main Outcome Paper

Psychotherapy and Psychosomatics

Standard Research Article

Psychother Psychosom
DOI: 10.1159/000542425

Received: September 13, 2024
Accepted: October 31, 2024
Published online: January 10, 2025

Active Components in Internet-Based Cognitive-Behavioral Therapy for Social Anxiety Disorder: A Randomized Full Factorial Trial

Dajana Šipka[a] Rodrigo Lopes[a] Tobias Krieger[a] Jan Philipp Klein[b, c] Thomas Berger[a]

[a]Department of Clinical Psychology and Psychotherapy, University of Bern, Bern, Switzerland; [b]Department of Psychiatry, Psychosomatics and Psychotherapy, University of Lübeck, Lübeck, Germany; [c]Center for Brain, Behavior and Metabolism, University of Lübeck, Lübeck, Germany

Keywords
Anxiety disorders · Social anxiety disorder · Factorial design · Internet-based treatment · Cognitive-behavioral psychotherapy

Abstract

Introduction: Many studies have demonstrated that social anxiety disorder (SAD) can be effectively treated with psychotherapy, particularly cognitive-behavioral therapy (CBT), including internet-based CBT (ICBT). Despite evidence-based treatments, many individuals do not sufficiently benefit from them. Identifying the active components could help improve the effectiveness of SAD treatment. This study tested the effects of four treatment components (psychoeducation, cognitive restructuring, attention training, and exposure) within ICBT for SAD to investigate its active components. ***Methods:*** This randomized full factorial trial consisted of four factors (i.e., treatment components) and 16 conditions. A total of 464 adults with a diagnosed SAD were recruited from the community. The primary outcome was SAD symptoms at 8 weeks (post-assessment). Secondary outcomes included SAD diagnosis, SAD symptoms at follow-up (4 months after post), depression and anxiety symptoms, quality of life, client satisfaction, and adverse effects. ***Results:*** Conditions including psychoeducation and exposure were significantly more effective in reducing SAD symptoms at post compared to conditions without these components. Conditions including cognitive restructuring and attention training did not show superiority over conditions without them at post. However, all treatment conditions significantly reduced symptoms compared to the condition without a treatment component. At follow-up, the superiority of psychoeducation and exposure was not significant anymore due to the version without the respective components catching up. ***Conclusion:*** The findings suggest that while all treatment components of ICBT for SAD are beneficial compared to no treatment, psychoeducation and exposure include specific active components that significantly improve treatment outcomes more quickly in ICBT for SAD.

Introduction

Social anxiety disorder (SAD) is characterized by a marked and persistent fear of negative evaluation in social situations [1]. Although the prevalences of SAD vary across the globe [2], it is one of the most common mental disorders, with an estimated 12-month prevalence of 7–8% in large-scale community-based studies [3, 4]. SAD

karger@karger.com
www.karger.com/pps

Correspondence to:
Dajana Šipka, dajana.sipka@unibe.ch

tends to have a chronic course and is associated with high levels of impairment in various areas of life [5].

Numerous trials have shown that SAD can be treated effectively with pharmacotherapy and psychotherapy [6]. Among psychological treatments, cognitive-behavioral therapy (CBT) has received consistent research support in meta-analyses [6–10] and is considered one of the first-line treatments [11]. However, despite evidence-based treatments for SAD, many affected people do not seek help. Only about 20–40% consult a mental health specialist [12–14]. This low rate is partly due to common barriers like geographical location, lack of trained therapists, stigma, treatment costs, lack of awareness, not knowing where to seek help, and difficulties scheduling an appointment [15–17]. Furthermore, the fear of social situations, which is characteristic of SAD, can result in a delay in seeking professional help of up to two decades [16, 18, 19]. Internet-based cognitive-behavioral therapy (ICBT) offers several advantages, including broader and easier access to empirically supported, affordable, and convenient treatment. Over the past decade, research on ICBT has increased rapidly, demonstrating the efficacy of this novel format for various mental disorders through numerous randomized controlled trials (RCTs) and meta-analyses [20, 21]. Research on internet-based treatments for SAD has been particularly extensive in the last two decades [20, 22–24].

Despite the positive results of CBT and ICBT for SAD, a significant number of patients do not recover after undergoing empirically based treatments [6, 22, 25]. One of the reasons for limited efficacy may be the lack of understanding of how these treatments work. Various authors convincingly argued that the active components (i.e., therapy components that significantly contribute to symptom improvement) of CBT need to be identified so that therapy can be made more efficacious and possibly briefer as well [26–28]. This requires the implementation of specific and meticulous study designs that evaluate the contribution of specific therapeutic elements. While RCTs represent the gold standard for determining the efficacy of an intervention package in comparison to a control group, they are constrained in their capacity to identify active components as they typically compare the overall effects of treatment packages [28]. Factorial designs, which allow the study of the individual and combined effects of multiple factors [29], are rare in psychotherapy research due to difficulties in clearly demarcating treatment components while maintaining protocol integrity. The emergence of internet-delivered treatments allows for standardized intervention content and better control over treatment integrity, making factorial designs more feasible. Consequently, several research groups have begun using factorial trials to identify active components of ICBT [30–34].

CBT for SAD comprises multiple components, including psychoeducation, cognitive restructuring, attention training, and exposure [35]. Each of these treatment components can contribute to the efficacy of the treatment package to some extent. Most evidence for the differential efficacy of specific treatment components comes from underpowered clinical trials that lack rigorous control over treatment integrity [36–40]. The most reliable evidence concerning different components and their combinations to date comes from meta-analyses, but the conclusions are inconsistent. Powers et al. [41], for example, found that combined cognitive and behavioral treatments were not significantly more effective than cognitive treatments or exposure treatments alone and reported few differences between standard CBT and its various components. The same results were found by Podina et al. [42] where no significant difference between exposure and cognitive therapy was found. However, these findings are at odds with other meta-analytic evidence indicating that exposure-based interventions yielded the largest effect sizes, whether administered alone [43] or in conjunction with cognitive restructuring [44]. In conclusion, research has not yet sufficiently examined the active components of SAD treatment separately [7].

The current OPTIMIZE study aimed to test the main effects of the four main treatment components according to the Clark and Wells [35] model as well as Stangier et al. [45] in ICBT for SAD using a factorial design. To the best of our knowledge, this is the first factorial trial testing the effects of CBT treatment components on symptoms of SAD and other outcomes.

Methods

Study Design

The study was a randomized full factorial trial with four factors, each representing one of the components of ICBT for SAD (i.e., psychoeducation, cognitive restructuring, attention training, and exposure). Each component was either absent (−1) or present (+1), resulting in 16 (2×2×2×2) conditions. The full version (FV) of the program consists of all four main components, 14 conditions contained all possible permutations out of the four components with at least one active component, and one condition contained no treatment component (waitlist control group, WL; see our published study

protocol [46]). The study was approved by the Ethics Committee of the canton of Bern on the 26th of April 2021 (BASEC-ID: 2020-02952). Written informed consent was obtained from all participants, and no reimbursement for participation was provided. The trial is registered with ClinicalTrials.gov (NCT04879641).

Participants

See Figure 1 for the study flowchart. A total of 464 participants with a SAD diagnosis were recruited from the 11th of August 2021 until the 27th of July 2023 in Switzerland, Germany, and Austria through social media accounts (i.e., Instagram, Facebook, TikTok), Google Ads, interviews on the radio and magazines, paper fliers, SAD-related websites as well as forums, and the University of Bern website. Interested people could download the informed consent form (ICF) via our study website. Upon return of the signed ICF, participants were invited to complete online questionnaires using the REDCap software (Research Electronic Data Capture; [47, 48]). The first selection of participants was completed by evaluating the two social anxiety measures. Only participants whose scores exceeded cut-off values on the Social Phobia Scale (SPS; cut-off >22) or the Social Interaction Anxiety Scale (SIAS; cut-off >33; [49], German version: [50]) were contacted for a structured diagnostic interview (i.e., International Neuropsychiatric Interview for DSM-IV 6.0.0 (M.I.N.I.), see relevant section below for more information) conducted by phone. The interviews were conducted by the study coordinator (D.S.) and trained advanced master's students in clinical psychology and psychotherapy. The students underwent training on administering the M.I.N.I. over the course of approximately 1 month, which included multiple half-day sessions, exercises, and homework assignments. The training was led by R.L., a licensed psychotherapist, and D.S., a psychotherapist in training. Throughout the project, D.S. and R.L. provided constant supervision to ensure the quality of the students' work.

Criteria for inclusion were (a) being at least 18 years old, (b) having access to the Internet, (c) having access to a tablet, PC, or smartphone, (d) sufficient German knowledge, (e) exceeding either the SPS or SIAS cut-off, (f) receiving a SAD diagnosis from the diagnostic interview, and (g) if psychopharmaceutic intake was present, having a stable dosage since at least 1 month.

Out of 685 individuals who signed the ICF, 464 persons met all inclusion criteria and were randomized to one of the 16 conditions (see flowchart in Fig. 1). Table 1 shows the baseline characteristics of the randomized participants. Across all conditions, the mean age of the sample was 31.35 years (SD = 10.58), with the majority being female (75%), single (42%), fully employed or students (35% and 36%, respectively), and having a high-school diploma as the highest educational degree (33%).

Randomization

The eligible participants were block-randomized (block sizes: 16 and 32) to ensure an equal distribution of different characteristics across all participants. The randomization schedule was produced with a self-written R code and concealed from the study team. After the participants were included in the study, the random allocation based on the randomization schedule was carried out automatically within the REDCap software.

Interventions: The Self-Help Program Shyne

The ICBT program Shyne, which is based on the well-established CBT treatment developed by Clark and Wells [35], Stangier et al. [45], and a self-help book by Rapee [51], has been shown efficacious in various versions, settings, and countries [22, 52–58]. The four main components of Shyne are psychoeducation, cognitive restructuring, attention training, and exposure.

All participants received guidance from pre-to post-treatment from trained and supervised advanced master students in clinical psychology and psychotherapy. The master students were trained by D.S. and R.L. on how to provide guidance using material and a handout developed in previous studies conducted by our research group (e.g. [31]). The handout included numerous examples of situations (e.g., how to react if a participant is inactive for a week). The guides were also required to submit homework, where they practiced using examples from previous participants, focusing on how to provide feedback for different usage profiles and respond to messages. Throughout the process, all guides were constantly supervised by D.S. and R.L., who were, in turn, supervised by T.B. Guidance was provided via a secure text-based messaging system integrated with the Shyne program. The role of the guides was to reinforce program use and adherence to treatment as well as the attempt to hold the dose that the participants receive constantly (each participant was advised to work 50–60 min a week on the program). Each week, the participants received a message from their assigned guide, commenting on their progress, suggesting exercises, and encouraging further engagement with the program. Additionally, participants could contact their guide and ask questions at any time. Guides were instructed to respond within three working days.

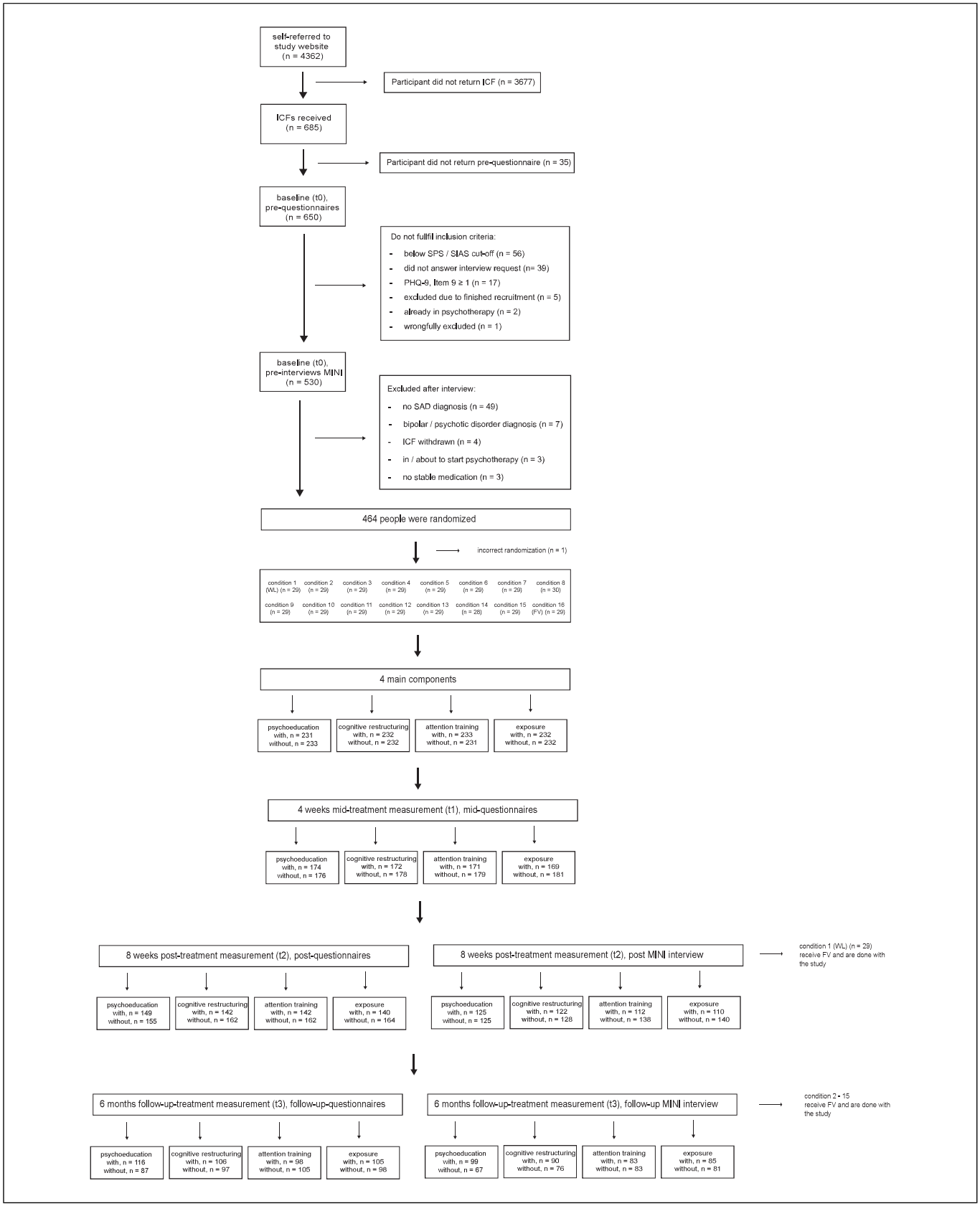


Fig. 1. Study flow. WL, waitlist control group; FV, full version.

Psychother Psychosom
DOI: 10.1159/000542425

Table 1. Characteristics from the randomized pre-sample OPTIMIZE

Characteristics	All conditions		Main component															
			psychoeducation				cognition				attention				exposure			
			yes		no		yes		no		yes		no		yes		no	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Sample size	464	100	231	50	233	50	232	50	232	50	233	50	231	50	232	50	232	50
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Age	31.35	10.58	30.41	10.36	33.07	10.79	29.34	10.90	33.07	10.23	28.00	10.69	33.07	10.48	33.24	10.53	33.07	10.64
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Sex																		
Female	347	75	177	77	170	73	178	77	169	73	169	73	178	77	172	74	175	75
Male	112	24	55	24	57	24	50	22	62	27	61	26	51	22	57	25	55	24
Non-binary	5	1	0	0	5	2	4	2	1	0	2	1	3	1	3	1	2	1
Relationship status																		
Single	195	42	99	43	96	41	103	44	92	40	97	42	98	42	100	43	95	41
In a relationship	183	39	98	42	85	36	96	41	87	38	92	39	91	39	91	39	92	40
Married	72	16	31	13	41	18	24	10	48	21	35	15	37	16	32	14	40	17
Divorced/widowed	14	3	4	2	10	4	9	4	5	2	8	3	6	3	9	4	5	2
Employment status																		
Full-time	164	35	86	37	78	33	82	35	82	35	73	31	91	39	87	38	77	33
Part-time	83	18	36	16	47	20	37	16	46	20	49	21	34	15	42	18	41	18
Student (full- or part-time)	165	36	87	38	78	33	84	36	81	35	83	36	82	35	76	33	89	38
Unemployed	37	8	15	6	22	9	22	9	15	6	20	9	17	7	21	9	16	7
Retired	10	2	6	3	4	2	6	3	4	2	5	2	5	2	3	1	7	3
Homemaker	5	1	2	1	3	1	1	0	4	2	2	1	3	1	3	1	2	1
Educational level																		
Lower than high-school diploma	73	16	106	46	64	27	37	16	36	16	28	12	45	19	36	16	37	16
High-school diploma	152	33	81	35	71	30	78	34	74	32	77	33	75	32	71	31	81	35
Bachelor's degree	101	22	50	22	51	22	58	25	43	19	52	22	49	21	56	24	45	19
Master's degree	76	16	34	15	42	18	36	16	40	17	45	19	31	13	42	18	34	15
Doctorate or higher	14	3	9	4	5	2	6	3	8	3	4	2	10	4	8	3	6	3
Other	48	10	18	8	30	13	17	7	31	13	26	11	22	10	19	8	29	13
Number of current comorbidities	555	100	259	47	296	53	291	52	264	48	261	47	294	53	281	51	274	49

Participant 71 was randomized into the false condition (condition 14 instead of 8). Therefore, the *n*'s are not divided equally.

Treatment Components

Psychoeducation. This component provided information on SAD, emphasizing the maintenance processes according to the Clark and Wells [35] model (e.g., the vicious cycle of negative thoughts, self-focused attention, anxiety, safety, and avoidance behavior) (see online suppl. material A, Fig. A1 for the model; for all online suppl. material, see https://doi.org/10.1159/000542425). Participants were asked questions about the content and their examples and experiences to provide more individualized education and boost memorization. Based on the psychoeducation, they were asked to create their personal SAD model. Additionally, participants were asked to fill out an anxiety protocol (see online suppl. material A, Fig. A2 for the anxiety protocol) in which they documented anxiety-inducing situations, along with associated thoughts, emotions, and behaviors.

Cognitive Restructuring. In this component, participants practiced cognitive restructuring for negative and dysfunctional thoughts and assumptions. They were first provided with information on the connection between cognitions, emotions, and behavior, as well as dysfunctional cognitions in people with SAD and the principle of cognitive restructuring. Then, they were asked to practice identifying and restructuring dysfunctional thoughts and assumptions in an online protocol. The content and materials were adopted from Berger et al. [52], based on the German therapy manual by Stangier et al. [45], which is itself based on the Clark and Wells [35] model, and a self-help book by Rapee [51].

Attention Training. This treatment component informed about the importance of self-focused attention in the maintenance of SAD. Participants were instructed to practice shifting away from self-focused attention and biased attention in general. Audio, video, and text-based exercises were provided, wherein participants were instructed to direct their attention away from themselves and to become less alert in social situations. For instance, a brief film showing an audience from a presenter's perspective was displayed on the computer screen, and participants were asked to summarize a brief story that was shown just before the audience appeared. They performed this task twice: first, by increasing self-focused attention, and second, by focusing on the task itself. After completing the exercise, participants were asked to report the location of their attention and the intensity of their anxiety. Finally, they received feedback on the association between self-focused attention and anxiety intensity. The concrete materials were originally developed by Berger et al. [52], based on Clark and Wells [35], Stangier et al. [45], and Rapee [51].

Exposure. Participants were informed of the importance of confronting oneself with social situations in reality and were encouraged and supported in planning and engaging in in vivo exposures. They were provided with information on safety and avoidance behavior, pre- and post-processing, and the effect of exposure. Furthermore, participants were asked to develop an anxiety hierarchy as the basis for the exposure. It was suggested to start with the easiest situation and, after mastering the situation, to gradually work through the anxiety hierarchy. Finally, an online diary for behavioral experiments was introduced, in which participants could record their planned exposures, including their expectations about anxiety levels, physical symptoms, potential safety behaviors, and other relevant variables. After completing the exposure, participants were encouraged to document their observations about the outcome (e.g., the actual intensity of anxiety experienced) and compare these with their initial expectations (see online suppl. material A, Fig. A3 for the exposure protocol).

Primary Outcomes

Social Phobia Scale & Social Interaction Anxiety Scale

The Social Phobia Scale (SPS) and Social Interaction Anxiety Scale (SIAS) ([49], German version: [50]) are self-report questionnaires that assess SAD symptoms. The questionnaires are often presented together. While SPS refers to performance-related situations (e.g., writing, holding a presentation), SIAS refers to interaction-related situations (meeting new people, maintaining eye contact). Each questionnaire comprises 20 items that can be rated from 0 ("not at all") to 4 ("extremely"). Internal consistency for both German versions was $\alpha = 0.94$ in the original validation study [50]. The internal consistencies for the baseline sample in this study were somewhat lower but still good, with $\alpha = 0.88$ for SPS, and $\alpha = 0.85$ for SIAS.

A composite score (CS) of SPS and SIAS was calculated as the primary outcome according to Song et al. [59]. The SPS and SIAS data were z-standardized based on the pre-mean and pre-standard deviation values and then summed up. Additionally, all calculations were also made for SPS and SIAS separately.

Secondary Outcomes

Generalized Anxiety Disorder Scale

The Generalized Anxiety Disorder Scale (GAD-7) ([60], German version: [61]) is a self-report questionnaire that assesses general anxiety symptoms in the last 2 weeks. The questionnaire consists of 7 items that can be rated from 0 ("not at all") to 3 ("almost every day"). The internal consistency for the German version was $\alpha = 0.89$ [61]. The internal consistency for the baseline sample in this study is lower but still good, with $\alpha = 0.81$.

Patient Health Questionnaire

The Patient Health Questionnaire (PHQ-9) ([62], German version: [63]) is a self-report questionnaire that assesses depressive symptoms according to the Diagnostic and Statistical Manual of Mental Disorders 4th edition (DSM-IV; [64]) criteria in the last 2 weeks. The questionnaire consists of 9 items that can be rated from 0 ("not at all") to 3 ("almost every day"). The internal consistency for the German version is $\alpha = 0.88$ [65]. The internal consistency for the baseline sample in this study is lower but still good, with $\alpha = 0.82$.

Short Form Health Survey

The Short Form Health Survey (SF-12) (German version: [66]) is a self-report questionnaire that assesses the impact of health on a person's everyday life and their quality of life (QoL). The results can be calculated into 2 scores that evaluate the physical and mental health impact. The internal consistencies, which Gandek et al. [66] reported as correlations, range from $r = 0.94$ to $r = 0.96$. The scores are T-standardized (M = 50, SD = 10). The R code that was used for the calculation of the subscales in this study is provided in online supplementary material B.

International Neuropsychiatric Interview for DSM-IV 6.0.0

The International Neuropsychiatric Interview for DSM-IV 6.0.0 (M.I.N.I.) ([67], German version: [68]) is a short-structured diagnostic interview based on the DSM-IV criteria. The version 6.0.0. was chosen since no German version was sufficiently validated for the DSM-5 criteria at the beginning of the study. For this study, only the SAD diagnosis is relevant; the other diagnoses (i.e., comorbidities) were only reported as the sum of all current comorbidities (see Table 1).

Client Satisfaction Questionnaire-8

The Client Satisfaction Questionnaire-8 (CSQ-8) ([69], German version: [70]) is a self-report questionnaire that assesses the patient's or participant's satisfaction with a treatment they received. The questionnaire consists of 8 items rated from 1 (implying a negative occurrence) to 4 (implying a positive occurrence). In this study, the questionnaire was slightly adapted for the internet-based treatment, and the mean was calculated as the outcome variable. Internal consistency for the German version is $\alpha = 0.92$ [70]. The internal consistency for the sample (n = 303) in this study is excellent, with $\alpha = 0.92$.

Negative Effects of the Treatment

The Negative Effects of the Treatment (INEP) [71] is a self-report questionnaire that measures the negative effects of psychotherapy. The participants rated 15 items of negative effects, first whether they occurred and, if so, whether they were due to the treatment. The questionnaire was slightly adapted for the internet-based treatment. Only items relevant to this study were examined and reported.

Time Spent on the Program

In this factorial trial, participants received a varying number of treatment components, ranging from no components at all (WL) to all four components (FV). Participants who received at least one component were asked to repeat the exercises for the entire treatment period from pre-to post-treatment. To estimate the actual dose of the intervention, we assessed the time spent on the program. Time spent on the program was estimated based on each user's click in the program, which is recorded with time stamps. Only activities that occurred between pre- and post-treatment were counted. Usage time windows under 1 min or no activity in the program for more than 24 min were not considered (since the user gets automatically logged out of Shyne after 24 min).

Power Analysis

For the power analysis, the smallest clinically relevant effect size was assumed to be Cohen's $d = 0.2$ for the main effect of an individual treatment component. Furthermore, an α error level of 0.05, a statistical power (i.e., $1 - \beta$) of 0.80, and a correlation of $r = 0.50$ between measures of the 4- and 8-week interval was chosen. Calculated with G*Power 3 [72], this resulted in a total of 384 participants needed. Considering a review on ICBT for SAD [22], a dropout rate of 20% was conservatively assumed for this study, resulting in a total of $N = 464$ participants, respectively, $n = 29$ participants per condition.

Statistical Analyses

All calculations were done with jamovi (Version: 2.4.14.0; [73]) and R Studio (Version: 2024.04.2+764; [74]). The codes of the calculations can be provided upon request.

For the baseline evaluation of the pre-sample, t tests for independent samples were used to test the between-group differences for continuous variables for each main component. The χ^2 tests were used to test the between-group differences for nominal variables for each main component.

The main analyses were carried out based on the intention-to-treat approach (ITT; i.e., using all randomized participants). We evaluated differential outcomes with linear mixed models. Time was a within-subject factor (i.e., pre, mid, post, and follow-up), and the four main components were between-subject factors (i.e., each factor is

either present [1] or absent [−1]). The fixed effects were the main component, time point, and their interaction term; the random effect was the intercept of each participant (i.e., every participant is expected to have a different pre-score). The means in the ITT sample were estimated with the restricted maximum likelihood. A separate model was estimated for each comparison/factor and outcome measure. Calculations of within- and between-group effect sizes (Cohen's *d*) were based on estimated means and the pooled standard deviations. Additionally, we exploratorily examined the effect of the number of components. For that, the conditions were first divided into categories based on the number of components they contain, ranging from 1 to 4 (e.g., the FV contains 4 components), and then the time effect within each category and the interaction effect between time and the category was calculated, using the CS.

Results

Results are presented for the ITT sample. The findings based on observed data can be found in online supplementary material C, Tables C1 and C2.

Baseline Evaluation

Independent *t* tests did not show any significant between-group differences (e.g., with vs. without psychoeducation) for age and number of current comorbidities as well as for all primary outcomes (i.e., CS, SPS, SIAS). For the secondary outcomes, there was a significant difference for GAD-7 in psychoeducation ($t(462) = 2.10$, $p = 0.036$), where the version without psychoeducation had a lower GAD-7 mean ($d = 0.20$ [0.01, 0.38]). For the other secondary outcomes (i.e., PHQ-9, GAD-7, SF-12 mental, SF-12 physical), there were no significant differences found (for all comparisons *p*s > 0.10). Furthermore, there were no significant between-group differences for sex, relationship status, employment status, or educational level, except for relationship status regarding the component cognitive restructuring ($\chi^2(4, N = 464) = 11.67$, $p = 0.020$). The group without cognitive restructuring had double the number of married participants as the group with cognitive restructuring (48 vs. 24).

Dropouts from the Study

Concerning the primary outcome, out of the 464 participants, 304 participants (66%) filled out the post-questionnaires, and 203 (44%) filled out the follow-up questionnaires. Participants who did not fill out the SPS and SIAS at post, respectively, at follow-up, were considered dropouts. Concerning the dropouts at post, no between-group differences were found for psychoeducation ($\chi^2(1, N = 464) = 0.34$, $p = 0.558$), cognitive restructuring ($\chi^2(1, N = 464) = 3.82$, $p = 0.051$), and attention training ($\chi^2(1, N = 464) = 3.82$, $p = 0.051$). There was, however, a significant between-group difference for exposure, where participants with the component exposure were more likely to dropout ($\chi^2(1, N = 464) = 5.49$, $p = 0.019$). There was a significant difference for the CS, where participants with higher SAD symptom scores were more likely to dropout ($t(462) = -1.98$, $p = 0.049$).

Concerning the dropouts at follow-up, there was a significant between-group effect for psychoeducation, where participants without psychoeducation were more likely to dropout ($\chi^2(1, N = 464) = 7.37$, $p = 0.007$). No other between-group differences were found (*p*s > 0.400). In comparison to the post-dropouts, there was no significant difference for the CS ($t(462) = -1.91$, $p = 0.057$) anymore.

Primary Outcomes

Table 2 shows estimated means, standard deviations, and Cohen's *d* effect sizes for the primary outcome (CS), SPS, and SIAS separately, divided according to the main factors (i.e., groups that have or have not received a main component). Time effects showed significant within-group effects (Fs_{CS} (2, $df(res)_{range}$ [330.37, 360.11]) ≤ 238.01, *p*s < 0.001; Fs_{SPS} (2, $df(res)_{range}$ [326.46, 358.72]) ≤ 195.43, *p*s < 0.001; Fs_{SIAS} (2, $df(res)_{range}$ [335.79, 361.40]) ≤ 199.40, *p*s < 0.001) with effect sizes ranging from medium to large effect sizes (see Table 2 for detailed effect sizes) for all conditions from pre-to post-treatment. Thus, SAD symptoms decreased in all groups, independent of the treatment component they received. However, when looking at the results for the 16 conditions separately, there was one exception: Participants in the WL did not improve significantly from pre to post ($F(2, 55.06) = 0.02$, $p = 0.98$, $d = -0.01$, Cl 95% [−0.53, 0.50]) (see online suppl. Table D1 in supplementary material D for the main results across all 16 conditions). Furthermore, in the between-group comparison of conditions 2–16 with the WL, all active conditions were significantly superior to the WL (*F*s (2, $df(res)_{range}$ [85.33, 102.71]) ≤ 36.99, *p*s < 0.001), with Cohen's *d*s ranging from $d = 0.67$, Cl 95% [0.07, 1.26] to $d = 1.56$, Cl 95% [0.87, 2.25].

The main objective of this full factorial trial was to evaluate the main effects of the four main treatment components of ICBT for SAD in post-treatment. In the following, we report the interaction effect between time (i.e., pre, mid, post) and each factor. Measured with the CS, there was a significantly higher reduction of SAD

Table 2. Primary outcomes for all components over all time points and Cohen's *d* for within-group and between-group effects

Main component	Pre-treatment		Mid-treatment		Post-treatment		Follow-up		Within-group effect size pre-post	Within-group effect size post-follow-up	Between-group effect size at post	Between-group effect size at follow-up
	M (SD)	*n*	M (SD)	*n*	M (SD)	*n*	M (SD)	*n*	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)
SPS												
Psychoed												
Yes	36.64 (12.46)	231	27.41 (13.53)	231	21.54 (14.13)	231	20.96 (15.20)	231	1.13 (0.93–1.33)	0.04 (–0.14 to 0.22)	0.39 (0.21–0.57)	0.21 (0.03–0.39)
No	36.71 (12.52)	233	30.34 (13.59)	233	27.02 (14.04)	233	24.01 (16.64)	233	0.73 (0.54–0.92)	0.20 (–0.01 to 0.38)		
Cognition												
Yes	36.87 (12.64)	232	28.28 (13.71)	232	23.15 (14.47)	232	22.04 (15.69)	232	1.01 (0.82–1.20)	0.07 (–0.11 to 0.26)	0.16 (–0.03 to 0.34)	0.05 (–0.14 to 0.23)
No	36.47 (12.64)	232	29.44 (15.56)	232	25.35 (13.86)	232	22.79 (16.15)	232	0.84 (0.65–1.03)	0.17 (–0.01 to 0.35)		
Attention												
Yes	36.14 (12.67)	233	27.67 (13.74)	233	23.29 (14.50)	233	22.00 (16.18)	233	0.94 (0.75–1.14)	0.08 (–0.10 to 0.27)	0.14 (–0.04 to 0.32)	0.05 (–0.13 to 0.24)
No	37.21 (12.61)	231	30.04 (13.53)	231	25.28 (13.58)	231	22.86 (15.65)	231	0.90 (0.71–1.09)	0.16 (–0.02 to 0.35)		
Exposure												
Yes	36.21 (12.49)	232	26.93 (13.71)	232	21.34 (14.47)	232	20.62 (15.69)	232	1.10 (0.91–1.30)	0.05 (–0.13 to 0.23)	0.40 (0.21–0.58)	0.22 (0.04–0.41)
No	37.14 (12.49)	232	30.69 (13.40)	232	26.95 (13.71)	232	24.18 (15.84)	232	0.78 (0.59–0.97)	0.19 (–0.01 to 0.37)		
SIAS												
Psychoed												
Yes	50.04 (12.46)	231	41.98 (13.37)	231	36.01 (13.98)	231	36.11 (14.89)	231	1.06 (0.87–1.25)	−0.01 (–0.19 to 0.18)	0.32 (0.14–0.50)	0.02 (–0.16 to 0.20)
No	50.24 (12.52)	233	44.19 (13.43)	233	40.44 (13.89)	233	36.46 (16.18)	233	0.74 (0.55–0.93)	0.26 (0.08–0.45)		
Cognition												
Yes	49.90 (12.49)	232	41.96 (13.56)	232	37.12 (14.32)	232	35.35 (15.38)	232	0.95 (0.76–1.14)	0.12 (–0.06 to 0.30)	0.15 (–0.03 to 0.33)	0.14 (–0.05 to 0.32)
No	50.38 (12.49)	232	44.19 (13.40)	232	39.31 (13.71)	232	37.45 (15.69)	232	0.84 (0.65–1.03)	0.13 (–0.06 to 0.31)		
Attention												
Yes	50.34 (12.52)	233	42.38 (13.59)	233	37.00 (14.20)	233	35.22 (15.72)	233	1.00 (0.80–1.19)	0.12 (–0.06 to 0.30)	0.17 (–0.01 to 0.35)	0.15 (–0.04 to 0.33)
No	49.94 (12.46)	231	43.75 (13.37)	231	39.37 (13.68)	231	37.48 (15.35)	231	0.81 (0.54–1.10)	0.13 (–0.05 to 0.31)		
Exposure												
Yes	49.64 (12.49)	232	41.56 (13.56)	232	35.76 (14.32)	232	35.17 (15.38)	232	1.03 (0.84–1.23)	0.04 (–0.14 to –0.22)	0.34 (0.15–0.52)	0.15 (–0.03 to 0.34)
No	50.64 (12.49)	232	44.52 (13.25)	232	40.47 (13.71)	232	37.55 (15.69)	232	0.78 (0.59–0.96)	0.20 (–0.02 to 0.38)		
CS												
Psychoed												
Yes	0.52 (0.76)	231	−0.10 (0.91)	231	−0.53 (0.91)	231	−0.54 (0.91)	231	1.25 (1.05–1.45)	0.01 (–0.17 to 0.19)	0.39 (0.21–0.58)	0.12 (–0.06 to 0.30)
No	0.53 (0.76)	233	0.09 (0.92)	233	−0.17 (0.92)	233	−0.42 (1.07)	233	0.83 (0.64–1.02)	0.25 (–0.07 to 0.43)		

Table 2 (continued)

Main component	Pre-treatment		Mid-treatment		Post-treatment		Follow-up		Within-group effect size pre-post	Within-group effect size post-follow-up	Between-group effect size at post	Between-group effect size at follow-up
	M (SD)	*n*	M (SD)	*n*	M (SD)	*n*	M (SD)	*n*	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)
Cognition												
Yes	0.52 (0.76)	232	−0.07 (0.91)	232	−0.43 (0.91)	232	−0.53 (1.07)	232	1.13 (0.94–1.33)	0.10 (–0.16 to 0.36)	0.18 (0.01–0.36)	0.09 (–0.09 to 0.28)
No	0.53 (0.76)	232	0.05 (0.91)	232	−0.27 (0.91)	232	−0.43 (1.07)	232	0.95 (0.76–1.15)	0.16 (–0.02 to 0.34)		
Attention												
Yes	0.51 (0.76)	233	−0.08 (0.91)	233	−0.43 (0.91)	233	−0.54 (1.06)	233	1.11 (0.92–1.31)	0.11 (–0.07 to 0.29)	0.18 (–0.01 to 0.36)	0.11 (–0.07 to 0.30)
No	0.54 (0.76)	231	0.06 (0.92)	231	−0.27 (0.92)	231	−0.42 (1.07)	231	0.97 (0.77–1.16)	0.15 (–0.03 to 0.33)		
Exposure												
Yes	0.49 (0.76)	232	−0.13 (0.91)	232	−0.54 (0.91)	232	−0.59 (1.07)	232	1.23 (1.03–1.43)	0.05 (–0.13 to 0.23)	0.41 (0.22–0.59)	0.20 (0.01–0.38)
No	0.56 (0.76)	232	0.11 (0.91)	232	−0.17 (0.91)	232	−0.38 (1.07)	232	0.87 (0.68–1.06)	0.21 (–0.03 to 0.39)		

Psychoed, psychoeducation; cognition, cognitive restructuring; attention, attention training.

symptoms for the conditions with psychoeducation in comparison to the conditions without psychoeducation ($F(2, 690.39) = 12.56$, $p < 0.001$). The between-group effect size for this difference was small to medium, $d = 0.39$, 95% CI [0.21, 0.58]. A statistically significant difference was also found for exposure, favoring conditions with exposure ($F(2, 691.50) = 9.60$, $p < 0.001$), with a small to medium between-group effect size, $d = 0.41$, 95% CI [0.22, 0.59] over conditions without exposure. There were no other significant interaction effects between time and a factor. Similar results were found for the SPS and SIAS, with significant and small to medium effect sizes in favor of conditions including psychoeducation and exposure (see Table 2).

As participants received different numbers of components, ranging from 0 (i.e., WL) to 4, we also exploratorily examined how the number of components affected the primary outcome. Figure 2 shows the CS depending on the number of components from pre to post. Time effects showed significant within-group effects for conditions with at least one component ($Fs(2, \text{df(res)}_{\text{range}}$ [177.12, 1,027.98]) $\leq$ 1,034.93, $p < 0.001$). There were significant interaction effects between time and the number of components for 0 vs. 1 ($F(2, 227.63) = 26.63$, $p < 0.001$), 0 vs. 2 ($F(2, 282.41) = 30.63$, $p < 0.001$), 0 vs. 3 ($F(2, 221.58) = 29.55$, $p < 0.001$), and 0 vs. 4 ($F(2, 92.84) = 32.32$, $p < 0.001$), meaning that it is better to give any number of components than none. Comparing the versions with at least one component, there were no significant interaction effects between time and the number of components for any of the versions (1 vs. 2 ($F(2, 414.77) = 0.70$, $p = 0.50$), 1 vs. 3 ($F(2, 350.45) = 2.69$, $p = 0.07$), 1 vs. 4 ($F(2, 217.11) = 1.60$, $p = 0.21$), 2 vs. 3 ($F(2, 412.87) = 1.24$, $p = 0.29$), 2 vs. 4 ($F(2, 276.82) = 1.03$, $p = 0.36$), 3 vs. 4 ($F(2, 214.19) = 0.19$, $p = 0.83$)). Thus, the quantity of treatment assigned does not appear to be a pivotal factor.

Secondary Outcomes

Table 3 shows the results for the secondary outcome variables PHQ-9, GAD-7, SF-12 (with the 2 scores mental and physical QoL), SAD diagnosis (i.e., the number of SAD diagnoses, conducted with the M.I.N.I., across all participants), CSQ-8, and their corresponding Cohen's *d* effect sizes. Within-group analyses showed significant time effects (pre, mid, post) for PHQ-9, GAD-7, and SF-12 mental. This means that participants improved in regard to depressive and anxiety symptoms as well as mental QoL over time, independent of which treatment component they received. Online supplementary material E, Table E1, shows the fixed effects omnibus tests. There was only one significant time effect for the PHQ-9 for the

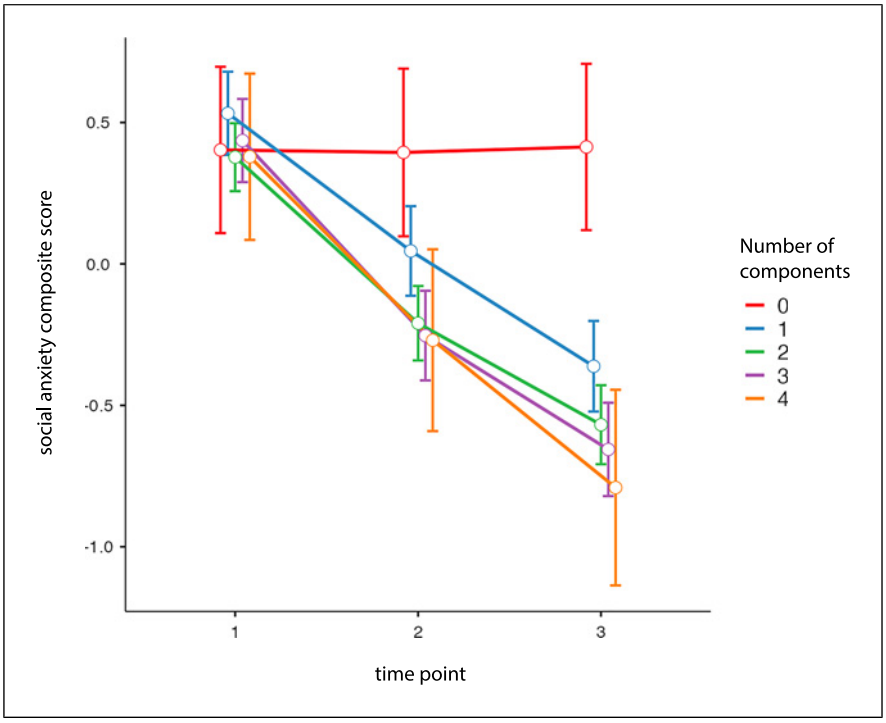


Fig. 2. Number of components over time for the primary outcome (social anxiety CS). 1 = pre; 2 = mid; 3 = post.

WL group: participants got less depressed from pre to post ($F(2, 55.18) = 4.25$, $p = 0.02$). There were no other significant time effects within the WL ($ps > 0.56$). No significant interaction effect between time (pre, mid, post) and group for PHQ-9, GAD-7, SF-12 mental or SF-12 physical was found.

The number of SAD diagnoses was the same at pre and post for the WL. For the main components, there was a significantly lower number of SAD diagnoses at post for the versions with psychoeducation in comparison to the versions without (psychoeducation: $\chi^2(1, n = 250) = 10.03$, $p = 0.002$, OR = 1.08; cognitive restructuring: $\chi^2(1, n = 250) = 0.39$, $p = 0.53$, OR = 0.85; attention training: $\chi^2(1, n = 250) = 0.22$, $p = 0.64$, OR = 0.89; exposure: $\chi^2(1, n = 250) = 1.84$, $p = 0.18$, OR = 0.71).

Concerning treatment satisfaction with the mean overall versions, participants were overall satisfied with the treatment with a mean and standard deviation ranging from 3.02 (0.52) to 3.23 (0.49). The only significant difference was found for psychoeducation, where satisfaction was higher for the conditions with psychoeducation in comparison to the conditions without psychoeducation ($t(272) = -3.40$, $p < 0.001$).

The INEP, measuring negative effects, showed that 18 participants (6%) suffered from having more experiences from the past, 4 participants (1.3%) experienced more conflicts in their romantic relationships, 4 participants (1.3%) perceived their relationship with their family and 3 participants (1%) with their friends as worse, 2 participants (0.7%) felt worse than before the treatment, and 1 participant (0.3%) had more difficulties trusting others. These results indicate that only a small percentage experienced deterioration from Shyne. There were no significant between-group differences ($ps > 0.35$).

Long-Term Effects

The estimated means at follow-up for primary and secondary outcomes are shown in Tables 2 and 3. Concerning the primary outcomes, there was a significant time (post, follow-up) effect for psychoeducation, respectively, exposure measured for the CS, in favor of the versions without psychoeducation ($F(1, 88.77) = 10.06$, $p = 0.002$), respectively, without exposure ($F(1, 100.40) = 6.48$, $p = 0.01$). Additionally, there was a significant time effect for cognitive restructuring ($F(1, 105.26) = 10.35$, $p = 0.002$) and attention training ($F(1, 95.68) = 9.05$, $p = 0.003$), in favor of the versions with the respective component. Upon further examination of the post hoc tests of the psychoeducation and exposure model with all-time points included (pre, mid, post, follow-up), the post to follow-up effect with the respective component was not significant (psychoeducation: $t(916.57) = -1.36$, $p = 0.175$; exposure: $t(917.62) = -1.74$, $p = 0.082$). This means that

Table 3. Secondary outcomes for all components over all time points and Cohen's *d* for within-group and between-group effects

Main component	Pre-treatment		Mid-treatment		Post-treatment		Follow-up		Within-group effect size pre-post	Within-group effect size post-follow-up	Between-group effect size at post	Between-group effect size at follow-up
	M (SD)	*n*	M (SD)	*n*	M (SD)	*n*	M (SD)	*n*	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)
PHQ-9												
Psychoed												
Yes	9.79 (4.56)	231	7.39 (5.02)	231	6.24 (5.17)	231	6.81 (5.62)	231	0.73 (0.54–0.92)	−0.11 (−0.29 to 0.08)	0.28 (0.10–0.47)	0.14 (−0.05 to 0.32)
No	10.29 (4.58)	233	8.37 (4.88)	233	7.71 (5.19)	233	7.61 (6.11)	233	0.53 (0.34–0.71)	0.02 (−0.16 to 0.20)		
Cognition												
Yes	10.19 (4.57)	232	8.12 (5.03)	232	7.07 (5.33)	232	7.11 (5.79)	232	0.63 (0.44–0.82)	−0.01 (−0.19 to 0.18)	−0.03 (−0.22 to 0.15)	0.04 (−0.14 to 0.22)
No	9.89 (4.57)	232	7.64 (5.03)	232	6.9 (5.03)	232	7.34 (5.94)	232	0.62 (0.44–0.81)	−0.08 (−0.26 to 0.10)		
Attention												
Yes	9.78 (4.58)	233	7.38 (5.04)	233	6.73 (5.34)	233	6.94 (5.95)	233	0.61 (0.43–0.80)	−0.04 (−0.22 to 0.15)	0.10 (−0.08 to 0.28)	0.09 (−0.09 to 0.27)
No	10.29 (4.56)	231	8.36 (4.86)	231	7.24 (5.02)	231	7.48 (5.78)	231	0.64 (0.45–0.82)	−0.04 (−0.23 to 0.14)		
Exposure												
Yes	9.97 (4.57)	232	7.52 (5.03)	232	6.94 (5.33)	232	7.14 (5.79)	232	0.61 (0.42–0.80)	−0.04 (−0.23 to 0.15)	0.02 (−0.16 to 0.20)	0.02 (−0.16 to 0.02)
No	10.11 (4.57)	232	8.22 (4.87)	232	7.04 (5.03)	232	7.27 (5.94)	232	0.64 (0.45–0.83)	−0.04 (−0.22 to 0.14)		
GAD-7												
Psychoed												
Yes	8.89 (4.10)	231	6.92 (4.56)	231	5.74 (4.71)	231	6.32 (5.02)	231	0.71 (0.53–0.90)	−0.12 (−0.30 to 0.06)	0.32 (014–0.50)	0.02 (−0.16 to 0.20)
No	9.72 (4.12)	233	7.71 (4.43)	233	7.24 (4.73)	233	6.44 (5.65)	233	0.56 (0.37–0.75)	0.15 (−0.03 to 0.34)		
Cognition												
Yes	9.52 (4.11)	232	7.49 (4.57)	232	6.61 (4.87)	232	6.64 (5.33)	232	0.65 (0.46–0.83)	−0.01 (−0.19 to 0.18)	−0.05 (−0.23 to 0.14)	−0.11 (−0.29 to 0.07)
No	9.09 (4.11)	232	7.14 (4.57)	232	6.39 (4.57)	232	6.22 (5.48)	232	0.62 (0.44–0.81)	0.03 (−0.15 to 0.22)		
Attention												
Yes	9.45 (4.12)	233	7.05 (4.58)	233	6.6 (4.88)	233	6.19 (5.50)	233	0.63 (0.45–0.82)	0.08 (−0.10 to 0.26)	−0.05 (−0.23 to 0.13)	−0.08 (−0.26 to 0.10)
No	9.16 (4.10)	231	7.57 (4.41)	231	6.41 (4.56)	231	6.66 (5.32)	231	0.63 (0.45–0.82)	−0.05 (−0.23 to 0.13)		
Exposure												
Yes	9.28 (4.11)	232	6.82 (4.57)	232	6.2 (4.87)	232	6.09 (5.33)	232	0.68 (0.50–0.87)	0.02 (−0.16 to 0.20)	0.12 (−0.06 to 0.30)	0.13 (−0.06 to 0.31)
No	9.33 (4.11)	232	7.77 (4.42)	232	6.77 (4.57)	232	6.76 (5.33)	232	0.59 (0.40–0.78)	0.002 (−0.18 to 0.18)		
SF-12, mental												
Psychoed												
Yes	34.46 (9.73)	231	40.57 (10.79)	231	42.53 (11.40)	231	41.87 (12.46)	231	0.76 (0.57–0.95)	−0.06 (−0.24 to 0.13)	0.28 (0.10–0.46)	0.07 (−0.12 to 0.25)
No	33.29 (9.77)	233	37.63 (10.84)	233	39.37 (11.30)	233	40.99 (13.89)	233	0.58 (0.39–0.76)	0.13 (−0.05 to 0.31)		

Table 3 (continued)

Main component	Pre-treatment		Mid-treatment		Post-treatment		Follow-up		Within-group effect size pre-post	Within-group effect size post-follow-up	Between-group effect size at post	Between-group effect size at follow-up
	M (SD)	*n*	M (SD)	*n*	M (SD)	*n*	M (SD)	*n*	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)	Cohen's *d* (95% CI)
Cognition												
Yes	33.78 (9.75)	232	38.47 (10.97)	232	40.71 (11.73)	232	41.82 (12.95)	232	0.64 (0.46–0.83)	0.09 (–0.09 to 0.27)	–0.04 (–0.22 to 0.14)	0.08 (–0.11 to 0.26)
No	33.98 (9.75)	232	39.7 (10.81)	232	41.15 (11.12)	232	40.81 (13.40)	232	0.69 (0.50–0.87)	–0.03 (–0.21 to 0.15)		
Attention												
Yes	34.01 (9.77)	233	39.52 (10.99)	233	40.96 (11.75)	233	42.59 (13.43)	233	0.64 (0.46–0.83)	0.13 (–0.05 to 0.31)	0.01 (–0.18 to 0.19)	0.18 (0–0.36)
No	33.75 (9.73)	231	38.69 (10.79)	231	40.89 (11.10)	231	40.19 (12.92)	231	0.68 (0.50–0.87)	–0.06 (–0.24 to 0.12)		
Exposure												
Yes	33.21 (9.75)	232	40.07 (10.97)	232	41.02 (11.73)	232	41.42 (12.95)	232	0.72 (0.54–0.91)	0.03 (–0.15 to 0.21)	0.01 (–0.17 to 0.20)	0.01 (–0.17 to 0.19)
No	34.54 (9.75)	232	38.21 (10.66)	232	40.87 (11.12)	232	41.32 (13.25)	232	0.61 (0.42–0.79)	0.04 (–0.15 to 0.22)		
SF-12, physical												
Psychoed												
Yes	52.95 (7.14)	231	53.2 (7.90)	231	52.55 (8.36)	231	53.54 (8.97)	231	–0.05 (–0.23 to 0.13)	0.11 (–0.07 to 0.30)	0.07 (–0.11 to 0.25)	0.21 (0.03–0.39)
No	52.36 (7.17)	233	52.28 (7.94)	233	51.98 (8.24)	233	51.54 (10.07)	233	–0.05 (–0.23 to 0.13)	–0.05 (–0.23 to 0.13)		
Cognition												
Yes	52.98 (7.16)	232	53.11 (7.92)	232	52.55 (8.53)	232	52.01 (9.29)	232	–0.06 (–0.24 to 0.13)	–0.06 (–0.24 to 0.12)	0.07 (–0.12 to 0.25)	–0.14 (–0.33 to –0.04)
No	52.33 (7.16)	232	52.38 (7.92)	232	51.99 (8.07)	232	53.36 (9.60)	232	–0.05 (–0.23 to 0.14)	0.15 (–0.03 to 0.34)		
Attention												
Yes	52.26 (7.17)	233	52.29 (7.94)	233	52.01 (8.55)	233	51.49 (9.62)	233	–0.03 (–0.21 to 0.15)	–0.06 (–0.24 to 0.13)	–0.06 (–0.25 to 0.12)	–0.24 (–0.42 to –0.06)
No	53.05 (7.14)	231	53.18 (7.75)	231	52.53 (8.06)	231	53.74 (9.27)	231	–0.07 (–0.25 to 0.11)	0.14 (–0.04 to 0.32)		
Exposure												
Yes	52.80 (7.16)	232	53.06 (8.07)	232	52.44 (8.53)	232	52.77 (9.44)	232	–0.05 (–0.23 to 0.14)	0.04 (–0.15 to 0.22)	0.04 (–0.14 to 0.22)	0.03 (–0.16 to 0.21)
No	52.52 (7.16)	232	52.43 (7.77)	232	51.11 (8.07)	232	52.53 (9.60)	232	–0.05 (–0.24 to 0.13)	0.05 (–0.14 to 0.23)		

Main component	Pre-treatment		Mid-treatment		Post-treatment		Follow-up	
	n	*n* (SAD = yes)	*n*	*n* (SAD = yes)	*n*	*n* (SAD = yes)	*n*	*n* (SAD = yes)
SAD diagnosis								
Psychoed								
Yes	231	231 (100%)	–	–	125	52 (42%)	99	32 (32%)
No	233	233 (100%)	–	–	125	76 (61%)	67	20 (30%)
Cognition								
Yes	232	232 (100%)	–	–	122	60 (49%)	90	25 (28%)
No	232	232 (100%)	–	–	128	68 (53%)	76	27 (36%)

Table 3 (continued)

Main component	Pre-treatment		Mid-treatment		Post-treatment		Follow-up	
	n	*n* (SAD = yes)	*n*	*n* (SAD = yes)	*n*	*n* (SAD = yes)	*n*	*n* (SAD = yes)
Attention								
Yes	233	233 (100%)	–	–	112	55 (49%)	83	21 (25%)
No	231	231 (100%)	–	–	138	73 (53%)	83	31 (37%)
Exposure								
Yes	232	232 (100%)	–	–	110	51 (46%)	85	24 (28%)
No	232	232 (100%)	–	–	140	77 (55%)	81	28 (35%)
n total per time point (%)	464 (100)	464 (100)	–	–	250 (100)	128 (51)	166 (100)	52 (31)

Psychoed, psychoeducation; cognition, cognitive restructuring; attention, attention training; PHQ-9, Patient Health Questionnaire; GAD-7, Generalized Anxiety Disorder Scale; SF-12, Short Form Health Questionnaire.

4 months after post, the initial advantages of the versions with psychoeducation and with exposure concerning the SA symptoms reduction measured with the CS disappeared (see online suppl. Fig. F1, supplementary material F for the development of psychoeducation and exposure measured with the CS over time).

Concerning the number of components (i.e., 1, 2, 3, 4), there was neither a significant time effect within each number of components (Fs(1, df(res)$_{range}$ [8.39, 84.59]) ≤ 4.30, ps ≥ 0.07) nor a significant interaction effect between any of 2 components (Fs(1, df(res)$_{range}$ [59.32, 181.99]) ≤ 3.11, ps ≥ 0.08). Concerning other secondary outcomes, there was a significant time effect for the versions with psychoeducation for the GAD-7 (F(1, 121.45) = 3.98, p = 0.048) with increased anxiety symptoms, for the version without cognitive restructuring for the SF-12 physical (F(1, 113.60) = 4.80, p = 0.03) with increased physical QoL, and for the version without attention training for the PHQ-9 (F(1, 652.42) = 6.11, p = 0.01), GAD-7 (F(1, 649.32) = 16.49, p < 0.001), SF-12 physical (F(1, 664.78) = 15.64, p < 0.001), and SF-12 mental (F(1, 656.27) = 6.93, p = 0.01) with increased depressive and anxiety symptoms, increased physical, but decreased mental QoL.

Furthermore, there was a significant interaction effect between time and GAD-7 for the component psychoeducation (F(1, 219.89) = 7.02, p = 0.01), with the conditions including psychoeducation showing more anxiety symptoms. Another significant effect between time and component for the SF-12 physical was found for the component cognitive restructuring (F(1, 238.63) = 4.07, p = 0.045) and attention training (F(1, 238.80) = 4.07, p = 0.045), with the version without cognitive restructuring, respectively, without attention training showing a higher physical QoL, and between time and attention training, was found. There were no other significant time or interaction effects on any secondary outcomes.

There were no significant differences in the number of SAD diagnoses at follow-up time point between a version with the main component in comparison to the version without (psychoeducation: χ^2(1, n = 166) = 0.05, p < 0.82, OR = 1.08; cognitive restructuring (χ^2(1, n = 166) = 1.15, p = 0.28), OR = 0.70; attention training: χ^2(1, n = 166) = 2.64, p = 0.12, OR = 0.59; exposure: χ^2(1, n = 166) = 0.77, p = 0.38, OR = 0.74).

Time Spent on the Program

Table 4 shows the time spent on the program broken down by the main components. It should be noted that the WL was not included in the calculation of the time spent on the program as this group did not use the program from pre to post. On average, participants spent around 266 min on the program. The only significant difference within each main component was found in psychoeducation, where the versions with psychoeducation showed higher use of the program than the versions without (t(426) = −4.78, p > 0.001). The effect size for this difference was small to medium, d = 0.46, 95% CI [0.27, 0.66].

Discussion

The main goal of this full factorial trial with 464 adults diagnosed with SAD was to investigate the main effects of the four treatment components (psychoeducation, cognitive restructuring, attention training, exposure) of ICBT

Table 4. Adherence in minutes for each main component

Main component	M adherence, min	SD adherence, min	*n*
Psychoeducation			
Yes	311.21	216.00	228
No	216.76	189.14	200
Cognition			
Yes	266.59	213.81	230
No	267.64	203.92	198
Attention			
Yes	281.57	214.29	228
No	250.55	202.18	200
Exposure			
Yes	269.23	199.56	228
No	264.52	220.40	200

Cognition, cognitive restructuring; attention, attention training.

for SAD on SAD symptoms at 8 weeks post-treatment. Regarding our main research question, we found a statistically significant superiority for the versions containing psychoeducation in comparison to versions without psychoeducation, and for versions containing exposure in comparison to versions without exposure. In contrast, we found no superiority regarding SAD symptom reduction for the versions that included cognitive restructuring or attention training compared to versions that did not include these two components. Considering our aim to identify the active components of the comprehensive ICBT treatment package for SAD, psychoeducation and exposure, in contrast to cognitive restructuring and attention training, seem to include specific contents that actively contribute to a better treatment outcome.

Although previous findings have indicated that psychoeducation and exposure are important and possibly sufficient components in the treatment of SAD [39, 75–77], to our knowledge, the present study is the first to find a significant superiority of interventions containing exposure or psychoeducation compared to interventions without these components. This may be due to the scarcity of research investigating the effects of individual treatment components, as well as the limited research comparing "pure" uncontaminated versions of different treatment components and their insufficient power. In face-to-face psychotherapy, it is challenging to demarcate treatment components from each other clearly and to avoid unwanted deviations from the treatment protocol by therapists. In this regard, a notable advantage of internet-based interventions is that the intervention content can be truly standardized, and treatment integrity (i.e., the degree to which an intervention is implemented as intended, can be effectively controlled). However, it is important to emphasize that such findings in studies on internet-based interventions may not necessarily generalize to face-to-face psychotherapy. Psychoeducation and exposure may be particularly well-suited for implementation in internet-based treatments and within a self-management framework. In contrast, successful cognitive restructuring may require more intensive therapist contact as given in a face-to-face setting. In the same vein, successful attention training may require environments where social anxiety is sufficiently activated [78], which may be difficult to achieve in an internet-based treatment conducted at home.

The number of treatment components did not impact treatment outcome. Interventions that included at least one treatment component were superior to the WL, and interventions comprising a single component were comparable to those comprising two or three components and the full treatment. It is important to note that participants were asked to repeat and practice the content and exercises taught for a period of 8 weeks independently of how many components they received (e.g., practicing cognitive restructuring for 8 weeks). A one-component treatment was, therefore, not necessarily a shortened treatment but a treatment that focused on just one method. The finding thus suggests that focusing on one method could be as effective as introducing several methods. However, the results may also indicate that the interventions could be shortened, aligning with the recent increased interest in internet-based brief interventions and single-session interventions [79, 80]. The present study could inform the development of brief or single-

session interventions for SAD as they are designed to deliver the active components of evidence-based treatments in a concentrated manner.

Regarding secondary outcomes, the only statistically significant interaction effect at post-treatment was found for SAD diagnosis, where the version with psychoeducation showed a significantly lower number of SAD diagnoses, again demonstrating the benefit of psychoeducation. Furthermore, we found statistically significant time effects regarding depressive, general anxiety symptoms, and mental QoL, independently of which treatment components participants received. Psychoeducation and exposure, therefore, showed no specific effect concerning most secondary outcomes. This may be because the psychoeducation and exposure treatment modules directly address SAD symptoms and not depressive or general anxiety symptoms. It may be that through the reception of any treatment, participants received hope for improvement as a general [81]. This would also be in line with the significant reduction of depressive symptoms in the WL since they knew in advance that they would receive the already tested FV after 8 weeks. Furthermore, participants were generally satisfied with the treatment independent of the version, and only a small percentage of the participants experienced a deterioration from the treatment. We found a statistically significant higher satisfaction in participants receiving psychoeducation compared to participants not receiving this treatment component. Participants who received the psychoeducation component also utilized the program for a longer duration than those who did not receive the component. This indicates that psychoeducation is not only a specifically efficacious component in reducing SAD symptoms but also meets the needs of the participants, as evidenced by the CSQ-8 and the time spent with this intervention component.

Follow-up results show that the superiority of treatments, including psychoeducation or exposure, compared to treatments in which they were absent, disappears by 4 months after treatment. The pattern of change between post-treatment and follow-up shows that this is not due to participants who have received psychoeducation and exposure deteriorating again since the effects of psychoeducation and exposure are maintained. Rather, the superiority of the treatments with psychoeducation and exposure disappears because treatments with cognitive restructuring and attention training catch up. Specifically, the treatment conditions with cognitive restructuring and attention training show a significant improvement from post-treatment to follow-up, which is not the case for the treatment conditions with psychoeducation and exposure. Despite an ongoing debate in the cognitive-behavioral literature about whether cognitive-based methods yield more lasting positive outcomes than exposure-based treatments for anxiety disorders, a longitudinal multilevel meta-analysis [42] found no evidence to support this effect, often referred to as a "sleeper effect." However, the pattern of change could also be understood in light of Fava et al. [82], who propose a sequential treatment model for mood and anxiety disorders. According to their framework, different treatment components serve distinct roles at various phases of recovery. Psychoeducation and exposure may be highly effective in the early phases of treatment by providing immediate symptom relief, but cognitive restructuring and attention training could be more beneficial in the long term by addressing deeper cognitive processes that sustain recovery. Overall, it is important to note that although the effects of the four components are similar at follow-up, psychoeducation and exposure accelerated the reduction of SAD symptoms, given their superiority in post-treatment. Accelerating the reduction of SAD symptoms is clearly an improvement in treatment, especially given the burden and chronic course of SAD symptoms.

The study has several limitations. Firstly, our participants received human guidance while using Shyne from pre to post, which could have had an effect on its own and, therefore, might have reduced potential differences between the effects of the treatment components. However, guidance aimed to ensure that participants engaged with the individual treatment components they were receiving and did not add new treatment content. Additionally, the guidance might have added a therapeutic effect (cf. [20]). Secondly, we had noteworthy dropout rates both from pre to post (34%) and from post to follow-up (56%). The rather high dropout rates are a common problem in ICBT studies [83]. The significantly higher dropout rate for exposure and more severe SAD symptoms at post could indicate that although exposure is effective, it might also be difficult for participants to expose themselves on their own and that some participants might need more support. On the other hand, the significantly higher dropout rate for psychoeducation at follow-up might stem from participants being satisfied with their treatment results and feeling no need to continue [84]. This would also be in line with the significantly higher treatment satisfaction with psychoeducation at post. In general, to ensure confidentiality, we asked participants to use anonymous email addresses for communication. However, this measure may have inadvertently contributed to participant attrition as some individuals might not have regularly monitored this email account. Thirdly, the sample was self-referred by the community, which limits the generalizability of our

findings to clinical samples. Fourthly, our sample consisted of German-speaking participants only. While the study by Thew et al. [85] and Yoshinaga et al. [86] showed that the results from ICBT, based on Clark and Wells [35], could be replicated cross-culturally with clinical samples, it would be necessary to actively outrule possible cultural biases.

Despite these limitations, this is the first randomized full factorial trial to test the effects of the main treatment components of ICBT for SAD. The findings suggest that while all treatment components of ICBT for SAD are beneficial compared to no treatment, psychoeducation and exposure include specific active components that accelerate the improvement of SAD symptoms. As a next step, potentially shorter treatments for SAD consisting of psychoeducation, exposure, and both components should be tested in confirmatory randomized controlled trials. Additionally, future research should explore other intervention components, such as imagery rescript, which has shown promise in addressing the emotional aspects of social anxiety by modifying aversive social memories [87].

Acknowledgments

We hereby express our sincere gratitude to all our participants for all their efforts and patience, without whom this study would not have been possible. We also would like to thank all our 36 master students, trainees, and employees who helped us with their exceptional work in the data acquisition process.

Statement of Ethics

The study was approved by the Ethics Committee of the canton of Bern on the 26th of April 2021 (BASEC-ID: 2020-02952). Written informed consent was obtained from all participants, and no reimbursement for participation was provided. The trial is registered with ClinicalTrials.gov (NCT04879641).

Conflict of Interest Statement

J.P.K. received funding for clinical trials (German Federal Ministry of Health, Servier), payments for presentations on psychological internet interventions (GAIA, Oberberg, Servier, Stillachhaus), consulting fees from developers and distributors of psychological internet interventions (all about me, Boehringer, Ethypharm, GAIA, sympatient), payments for workshops and books (Beltz, Elsevier, Hogrefe, and Springer) on psychotherapy for chronic depression and on psychiatric emergencies. He serves as vice chairman of the chapter "Digital Psychiatry" of the German Psychiatric Association (DGPPN).

Funding Sources

The study is funded by the Swiss National Science Foundation (SNSF) through a Project Funding Grant in Humanities and Social Sciences (Division I), given to Prof. Dr. Thomas Berger, with reference No. 10001C_197475. The funder had no role in the design, data collection, data analysis, and reporting of this study.

Author Contributions

The conceptualization of the project was done by T.B., T.K., J.P.K., R.L., and D.S. The data acquisition for the sample was done by D.S., R.L., and T.B. with the help of J.P.K. The calculations were done and the first version of the manuscript was written by D.S. with the support of T.B. All authors reviewed, edited, and approved the final manuscript.

Data Availability Statement

All data except for the demographic data are available here: https://boris-portal.unibe.ch/handle/20.500.12422/103878. Due to the demographic data containing information that could compromise the privacy of research participants, it is only available upon request from either the corresponding author (D.S.), R.L., or T.B. The codes of the calculations can be provided upon request from D.S.

References

1 American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. Amer Psychiatric Pub Incorporated; 2013; p. 991.
2 Stein DJ, Lim CCW, Roest AM, de Jonge P, Aguilar-Gaxiola S, Al-Hamzawi A, et al. The cross-national epidemiology of social anxiety disorder: data from the world mental health survey initiative. BMC Med. 2017;15(1):143. https://doi.org/10.1186/s12916-017-0889-2
3 Magee WJ, Eaton WW, Wittchen HU, McGonagle KA, Kessler RC. Agoraphobia, simple phobia, and social phobia in the National Comorbidity Survey. Arch Gen Psychiatry. 1996;53(2):159–68. https://doi.org/10.1001/archpsyc.1996.01830020077009
4 Ruscio AM, Brown TA, Chiu WT, Sareen J, Stein MB, Kessler RC. Social fears and social phobia in the USA: results from the national comorbidity survey replication. Psychol Med. 2008;38(1):15–28. https://doi.org/10.1017/S0033291707001699
5 Fehm L, Pelissolo A, Furmark T, Wittchen HU. Size and burden of social phobia in Europe. Eur Neuropsychopharmacol. 2005;15(4):453–62. https://doi.org/10.1016/j.euroneuro.2005.04.002
6 Mayo-Wilson E, Dias S, Mavranezouli I, Kew K, Clark DM, Ades AE, et al. Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. Lancet Psychiatr. 2014;1(5):368–76. https://doi.org/10.1016/S2215-0366(14)70329-3
7 Acarturk C, Cuijpers P, van Straten A, de Graaf R. Psychological treatment of social anxiety disorder: a meta-analysis. Psychol Med. 2009;39(2):241–54. https://doi.org/10.1017/S0033291708003590

8 de Ponti N, Matbouriahi M, Franco P, Harrer M, Miguel C, Papola D, et al. The efficacy of psychotherapy for social anxiety disorder, a systematic review and meta-analysis. J Anxiety Disord. 2024;104:102881. https://doi.org/10.1016/j.janxdis.2024.102881
9 Hofmann SG, Smits JAJ. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. J Clin Psychiatry. 2008;69(4):621–32. https://doi.org/10.4088/jcp.v69n0415
10 Kindred R, Bates GW, McBride NL. Long-term outcomes of cognitive behavioural therapy for social anxiety disorder: a meta-analysis of randomised controlled trials. J Anxiety Disord. 2022;92:102640. https://doi.org/10.1016/j.janxdis.2022.102640
11 National Institute for Clinical Excellence. Overview Social anxiety disorder: recognition, assessment and treatment Guidance NICE. [Internet]. [cited 2019 Sep 3]. Available from: https://www.nice.org.uk/guidance/cg159
12 Bruffaerts R, Harris MG, Kazdin AE, Vigo DV, Sampson NA, Chiu WT, et al. Perceived helpfulness of treatment for social anxiety disorder: findings from the WHO World Mental Health Surveys. Soc Psychiatry Psychiatr Epidemiol. 2022;57(10):2079–95. https://doi.org/10.1007/s00127-022-02249-3
13 Gross R, Olfson M, Gameroff MJ, Shea S, Feder A, Lantigua R, et al. Social anxiety disorder in primary care. Gen Hosp Psychiatry. 2005;27(3):161–8. https://doi.org/10.1016/j.genhosppsych.2005.01.006
14 Issakidis C, Andrews G. Service utilisation for anxiety in an Australian community sample. Soc Psychiatry Psychiatr Epidemiol. 2002;37(4):153–63. https://doi.org/10.1007/s001270200009
15 Carbonell Á, Navarro-Pérez JJ, Mestre MV. Challenges and barriers in mental healthcare systems and their impact on the family: a systematic integrative review. Health Soc Care Community. 2020;28(5):1366–79. https://doi.org/10.1111/hsc.12968
16 Haugen PT, McCrillis AM, Smid GE, Nijdam MJ. Mental health stigma and barriers to mental health care for first responders: a systematic review and meta-analysis. J Psychiatr Res. 2017;94:218–29. https://doi.org/10.1016/j.jpsychires.2017.08.001
17 Proudfoot JG. Computer-based treatment for anxiety and depression: is it feasible? Is it effective? Neurosci Biobehav Rev. 2004;28(3):353–63. https://doi.org/10.1016/j.neubiorev.2004.03.008
18 Keller MB. The lifelong course of social anxiety disorder: a clinical perspective. Acta Psychiatr Scand. 2003;108(s417):85–94. https://doi.org/10.1034/j.1600-0447.108.s417.6.x
19 Olfson M, Guardino M, Struening E, Schneier FR, Hellman F, Klein DF. Barriers to the treatment of social anxiety. Am J Psychiatry. 2000;157(4):521–7. https://doi.org/10.1176/appi.ajp.157.4.521
20 Andersson G, Berger T. Internet approaches to psychotherapy: empirical findings and future directions. In: Barkham M, Lutz W, Castonguay LG, editors. Handbook of psychotherapy and behavior change. 50th Anniversary ed. John Wiley & Sons; 2021. p. 739–62.
21 Hedman-Lagerlöf E, Carlbring P, Svärdman F, Riper H, Cuijpers P, Andersson G. Therapist-supported Internet-based cognitive behaviour therapy yields similar effects as face-to-face therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. World Psychiatr. 2023;22(2):305–14. https://doi.org/10.1002/wps.21088
22 Boettcher J, Carlbring P, Renneberg B, Berger T. Internet-based interventions for social anxiety disorder: an overview. Verhaltenstherapie. 2013;23(3):160–8. https://doi.org/10.1159/000354747
23 Guo S, Deng W, Wang H, Liu J, Liu X, Yang X, et al. The efficacy of internet-based cognitive behavioural therapy for social anxiety disorder: a systematic review and meta-analysis. Clin Psychol Psychother. 2021;28(3):656–68. https://doi.org/10.1002/cpp.2528
24 Clark DM, Wild J, Warnock-Parkes E, Stott R, Grey N, Thew G, et al. More than doubling the clinical benefit of each hour of therapist time: a randomised controlled trial of internet cognitive therapy for social anxiety disorder. Psychol Med. 2023;53(11):5022–32. https://doi.org/10.1017/S0033291722002008
25 Otto MW, Pollack MH, Gould RA, Worthington JJ, McArdle ET, Rosenbaum JF, et al. A comparison of the efficacy of clonazepam and cognitive-behavioral group therapy for the treatment of social phobia. J Anxiety Disord. 2000;14(4):345–58. https://doi.org/10.1016/s0887-6185(00)00027-x
26 Holmes EA, Craske MG, Graybiel AM. Psychological treatments: a call for mental-health science. Nat News. 2014;511(7509):287–9. https://doi.org/10.1038/511287a
27 Kazdin AE. Addressing the treatment gap: a key challenge for extending evidence-based psychosocial interventions. Behav Res Ther. 2017;88:7–18. https://doi.org/10.1016/j.brat.2016.06.004
28 Watkins ER, Newbold A. Factorial designs help to understand how psychological therapy works. Front Psychiatry. 2020;11:429. https://doi.org/10.3389/fpsyt.2020.00429
29 Collins LM, Dziak JJ, Kugler KC, Trail JB. Factorial experiments: efficient tools for evaluation of intervention components. Am J Prev Med. 2014;47(4):498–504. https://doi.org/10.1016/j.amepre.2014.06.021
30 Andersson G, Käll A, Juhlin S, Wahlström C, de Fine Licht E, Färdeman S, et al. Free choice of treatment content, support on demand and supervision in internet-delivered CBT for adults with depression: a randomized factorial design trial. Behav Res Ther. 2023;162:104265. https://doi.org/10.1016/j.brat.2023.104265
31 Bur OT, Krieger T, Moritz S, Klein JP, Berger T. Optimizing the context of support of web-based self-help in individuals with mild to moderate depressive symptoms: a randomized full factorial trial. Behav Res Ther. 2022;152:104070. https://doi.org/10.1016/j.brat.2022.104070
32 Sundström C, Peynenburg V, Chadwick C, Thiessen D, Wilhems A, Nugent M, et al. Optimizing internet-delivered cognitive behaviour therapy for alcohol misuse: a randomized factorial trial examining effects of a pre-treatment assessment interview and guidance. Addict Sci Clin Pract. 2022;17(1):37. https://doi.org/10.1186/s13722-022-00319-0
33 Watkins E, Newbold A, Tester-Jones M, Collins LM, Mostazir M. Investigation of active ingredients within internet-delivered cognitive behavioral therapy for depression: a randomized optimization trial. JAMA Psychiatry. 2023;80(9):942–51. https://doi.org/10.1001/jamapsychiatry.2023.1937
34 Watkins ER, Newbold A, Tester-Jones M, Javaid M, Cadman J, Collins LM, et al. Implementing multifactorial psychotherapy research in online virtual environments (IMPROVE-2): study protocol for a phase III trial of the MOST randomized component selection method for internet cognitive-behavioural therapy for depression. BMC Psychiatry. 2016;16(1):345. https://doi.org/10.1186/s12888-016-1054-8
35 Clark DM, Wells A. A cognitive model of social phobia. In: Heimberg RG, Liebowitz MR, Hope D, Schneider F, editors. Social Phobia: diagnosis, assessment, and treatment. New York, NY: Guilford Press; 1995. p. 69–93.
36 Hope DA, Heimberg RG, Bruch MA. Dismantling cognitive-behavioral group therapy for social phobia. Behav Res Ther. 1995;33(6):637–50. https://doi.org/10.1016/0005-7967(95)00013-n
37 Mattick RP, Peters L, Clarke JC. Exposure and cognitive restructuring for social phobia: a controlled study. Behav Ther. 1989;20(1):3–23. https://doi.org/10.1016/s0005-7894(89)80115-7
38 Nortje C, Posthumus T. Scores on an emotional stroop task after treatment of social anxiety disorder. Psychol Rep. 2012;111(2):461–71. https://doi.org/10.2466/02.15.20.PR0.111.5.461-471
39 Powers MB, Sigmarsson SR, Emmelkamp PMG. A meta-analytic review of psychological treatments for social anxiety disorder. Int J Cogn Ther. 2008;1(2):94–113. https://doi.org/10.1680/ijct.2008.1.2.94
40 Rodebaugh TL, Holaway RM, Heimberg RG. The treatment of social anxiety disorder. Clin Psychol Rev. 2004;24(7):883–908. https://doi.org/10.1016/j.cpr.2004.07.007
41 Powers MB, Sigmarsson SR, Emmelkamp PMG. A meta-analytic review of psychological treatments for social anxiety disorder. Int J Cogn Ther. 2008;1(2):94–113. https://doi.org/10.1521/ijct.2008.1.2.94

42 Podina IR, Vîslă A, Fodor LA, Flückiger C. Is there a sleeper effect of exposure-based vs. cognitive-only intervention for anxiety disorders? A longitudinal multilevel meta-analysis. Clin Psychol Rev. 2019;73:101774. https://doi.org/10.1016/j.cpr.2019.101774
43 Gil PJM, Carrillo FXM, Meca JS. Effectiveness of cognitive-behavioral treatment in social phobia: a meta-analytic review. Psychol Spain. 2001;5(1):17–25.
44 Gould RA, Buckminster S, Pollack MH, Otto MW, Yap L. Cognitive-behavioral and pharmacological treatment for social phobia: a meta-analysis. Clin Psychol Sci Pract. 1997; 4(4):291–306. https://doi.org/10.1111/j.1468-2850.1997.tb00123.x
45 Stangier U, Heidenreich T, Peitz M. Soziale Phobien: ein kognitiv-verhaltenstherapeutisches Behandlungsmanual. PVU: Beltz; 2003.
46 Lopes RCT, Šipka D, Krieger T, Klein JP, Berger T. Optimizing cognitive-behavioral therapy for social anxiety disorder and understanding the mechanisms of change: study protocol for a randomized factorial trial. Internet Interv. 2021;26:100480. https://doi.org/10.1016/j.invent.2021.100480
47 Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an International community of software platform partners. J Biomed Inform. 2019;95:103208. https://doi.org/10.1016/j.jbi.2019.103208
48 Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377–81. https://doi.org/10.1016/j.jbi.2008.08.010
49 Mattick RP, Clarke JC. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behav Res Ther. 1998;16.
50 Stangier U, Heidenreich T, Berardi A, Golbs U, Hoyer J. Die Erfassung sozialer Phobie durch Social Interaction Anxiety Scale (SIAS) und die Social Phobia Scale (SPS). Z Klin Psychol. 1999;28(1):28–36. https://doi.org/10.1026//0084-5345.28.1.28
51 Rapee RM. Overcoming shyness and social phobia: a step-by-step guide. Jason Aronson; 1998.
52 Berger T, Hohl E, Caspar F. Internet-based treatment for social phobia: a randomized controlled trial. J Clin Psychol. 2009;65(10): 1021–35. https://doi.org/10.1002/jclp.20603
53 Berger T, Caspar F, Richardson R, Kneubühler B, Sutter D, Andersson G. Internet-based treatment of social phobia: a randomized controlled trial comparing unguided with two types of guided self-help. Behav Res Ther. 2011;49(3):158–69. https://doi.org/10.1016/j.brat.2010.12.007
54 Kählke F, Berger T, Schulz A, Baumeister H, Berking M, Cuijpers P, et al. Efficacy and cost-effectiveness of an unguided, internet-based self-help intervention for social anxiety disorder in University students: protocol of a randomized controlled trial. BMC Psychiatry. 2019;19(1):197. https://doi.org/10.1186/s12888-019-2125-4
55 Kishimoto T, Krieger T, Berger T, Qian M, Chen H, Yang Y. Internet-based cognitive behavioral therapy for social anxiety with and without guidance compared to a wait list in China: a propensity score study. Psychother Psychosom. 2016;85(5):317–9. https://doi.org/10.1159/000446584
56 Schulz A, Stolz T, Vincent A, Krieger T, Andersson G, Berger T. A sorrow shared is a sorrow halved? A three-arm randomized controlled trial comparing internet-based clinician-guided individual versus group treatment for social anxiety disorder. Behav Res Ther. 2016;84:14–26. https://doi.org/10.1016/j.brat.2016.07.001
57 Skocic S, Jackson H, Hulbert C, Faber C. The maintaining factors of social anxiety: a three-group comparison of a clinical sample with highly socially anxious students and non-anxious students. Behav Cogn Psychother. 2016;44(4):385–96. https://doi.org/10.1017/S1352465815000685
58 Stolz T, Schulz A, Krieger T, Vincent A, Urech A, Moser C, et al. A mobile app for social anxiety disorder: a three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. J Consult Clin Psychol. 2018;86(6):493–504. https://doi.org/10.1037/ccp0000301
59 Song MK, Lin FC, Ward SE, Fine JP. Composite variables: when and how. Nurs Res. 2013;62(1):45–9. https://doi.org/10.1097/NNR.0b013e3182741948
60 Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006;166(10):1092–7. https://doi.org/10.1001/archinte.166.10.1092
61 Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. Med Care. 2008;46(3): 266–74. https://doi.org/10.1097/MLR.0b013e318160d093
62 Spitzer RL, Kroenke K, Williams JBW; The Patient, Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. JAMA. 1999;282(18):1737–44. https://doi.org/10.1001/jama.282.18.1737
63 Löwe B, Spitzer RL, Zipfel S, Herzog W. Gesundheitsfragebogen für Patienten (PHQ-D). Komplettversion und Kurzform. 2nd ed. Pfizer GmbH; 2002.
64 American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. 1994.
65 Gräfe K, Zipfel S, Herzog W, Löwe B. Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)": Ergebnisse der deutschen Validierungsstudie. Diagnostica. 2004;50(4): 171–81. https://doi.org/10.1026/0012-1924.50.4.171
66 Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 health survey in nine countries: results from the IQOLA project. International quality of life assessment. J Clin Epidemiol. 1998;51(11):1171–8. https://doi.org/10.1016/s0895-4356(98)00109-7
67 Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry. 1998;59(Suppl 20):22–57.
68 Ackenheil M, Stotz-Ingenlath G, Dietz-Bauer R, Vossen A. M.I.N.I Mini International Neuropsychiatric Interview, German version 5.0. 0 DSM IV. Munich Psychiatr Univ Clin; 1999.
69 Attkisson CC, Zwick R. The client satisfaction questionnaire: psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann. 1982;5(3):233–7. https://doi.org/10.1016/0149-7189(82)90074-x
70 Schmidt J, Wittmann WW. Fragebogen zur Messung der Patientenzufriedenheit. In: Brähler E, Schumacher J, Strauß B, editors. Diagnostische verfahren in der psychotherapie. Göttingen: Hogrefe; 2002. p. 392–6.
71 Ladwig I, Rief W, Nestoriuc Y. Welche Risiken und Nebenwirkungen hat Psychotherapie? Entwicklung des Inventars zur Erfassung Negativer Effekte von Psychotherapie (INEP). Verhaltenstherapie. 2014;24(4): 252–63. https://doi.org/10.1159/000367928
72 Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39(2):175–91. https://doi.org/10.3758/bf03193146
73 The jamovi project. Jamovi [Internet]. 2024. Available from: https://www.jamovi.org
74 Posit team. RStudio: integrated development environment for R. Posit software. Boston, MA: PBC; 2024. [Internet] Available from: http://www.posit.co/
75 Hofmann SG. Cognitive mediation of treatment change in social phobia. J Consult Clin Psychol. 2004;72(3):392–9. https://doi.org/10.1037/0022-006X.72.3.392
76 Kaczkurkin AN, Foa EB. Cognitive-behavioral therapy for anxiety disorders: an update on the empirical evidence. Dialogues Clin Neurosci. 2015;17(3):337–46. https://doi.org/10.31887/DCNS.2015.17.3/akaczkurkin
77 Rodrigues F, Bartolo A, Pacheco E, Pereira A, Silva CF, Oliveira C. Psycho-education for anxiety disorders in adults: a systematic review of its effectiveness. J Foren Psy. 2018; 03(02). https://doi.org/10.4172/2475-319x.1000142.

78 Heeren A, Mogoaşe C, Philippot P, McNally RJ. Attention bias modification for social anxiety: a systematic review and meta-analysis. Clin Psychol Rev. 2015;40:76–90. https://doi.org/10.1016/j.cpr.2015.06.001
79 Bisby MA, Barrett V, Staples LG, Nielssen O, Dear BF, Titov N. Things you do: a randomized controlled trial of an unguided ultra-brief intervention to reduce symptoms of depression and anxiety. J Anxiety Disord. 2024;105:102882. https://doi.org/10.1016/j.janxdis.2024.102882
80 Schleider JL, Dobias M, Sung J, Mumper E, Mullarkey MC. Acceptability and utility of an open-access, online single-session intervention platform for adolescent mental health. JMIR Ment Health. 2020;7(6):e20513. https://doi.org/10.2196/20513
81 Oehler C, Görges F, Böttger D, Hug J, Koburger N, Kohls E, et al. Efficacy of an internet-based self-management intervention for depression or dysthymia: a study protocol of an RCT using an active control condition. BMC Psychiatry. 2019;19(1):90. https://doi.org/10.1186/s12888-019-2063-1
82 Fava GA, Ruini C, Rafanelli C. Sequential treatment of mood and anxiety disorders. J Clin Psychiatry. 2005;66(11):1392–400. https://doi.org/10.4088/jcp.v66n1108
83 Taylor CB, Graham AK, Flatt RE, Waldherr K, Fitzsimmons-Craft EE. Current state of scientific evidence on Internet-based interventions for the treatment of depression, anxiety, eating disorders and substance abuse: an overview of systematic reviews and meta-analyses. Eur J Public Health. 2021;31(31 Suppl 1):i3–10. https://doi.org/10.1093/eurpub/ckz208
84 Lawler K, Earley C, Timulak L, Enrique A, Richards D. Dropout from an internet-delivered cognitive behavioral therapy intervention for adults with depression and anxiety: qualitative study. JMIR Form Res. 2021;5(11):e26221. https://doi.org/10.2196/26221
85 Thew GR, Kwok APL, Lissillour Chan MH, Powell CLYM, Wild J, Leung PWL, et al. Internet-delivered cognitive therapy for social anxiety disorder in Hong Kong: a randomized controlled trial. Internet Interv. 2022;28:100539. https://doi.org/10.1016/j.invent.2022.100539
86 Yoshinaga N, Thew GR, Hayashi Y, Matsuoka J, Tanoue H, Takanashi R, et al. Preliminary evaluation of translated and culturally adapted internet-delivered cognitive therapy for social anxiety disorder: multicenter, single-arm trial in Japan. JMIR Form Res. 2023;7(1):e45136. https://doi.org/10.2196/45136
87 Seinsche RJ, Fricke S, Neudert MK, Zimmer RI, Stark R, Hermann A. Emotional changes during imagery rescripting of aversive social memories in social anxiety disorder: a randomized controlled trial. Psychother Psychosom. 2024;93(4):264–70. https://doi.org/10.1159/000539402

Statement of Authorship

Philosophisch-humanwissenschaftliche Fakultät
Dekanat
Fabrikstrasse 8, CH-3012 Bern

UNIVERSITÄT BERN

Erklärung zur Dissertation

Hiermit bestätige ich, dass ich die Dissertation (Titel):

Multidimensional Perspective on Social Anxiety Disorder and Its Treatment Through Internet-Based Cognitive Behavioral Therapy

im Fach Klinische Psychologie und Psychotherapie

unter der Leitung von Prof. Dr. phil. Thomas Berger

ohne unerlaubte Hilfe ausgeführt und an keiner anderen Universität zur Erlangung eines akademischen Grades eingereicht habe.

Datum	Unterschrift
10.09.24	Dajana Kipka