cpe.psychopen.eu | ISSN 2625-3410 | December 2024 | Volume 6 | Issue 4 |

CLINICAL PSYCHOLOGY IN EUROPE

The Official Academic Journal of the European Association of Clinical Psychology and Psychological Treatment

PsychOpen GOLD

CLINICAL PSYCHOLOGY IN EUROPE

The Official Academic Journal of the
European Association of Clinical Psychology and Psychological Treatment

PsychOpen GOLD

CLINICAL PSYCHOLOGY IN EUROPE

December 2024 | Volume 6 | Issue 4

Contents

PsychOpen GOLD is a publishing service by
Leibniz Institute for Psychology Information (ZPID), Germany.
www.leibniz-psychology.org

PsychOpen GOLD

Clinical Psychology in Europe

PsychOpen GOLD

Editorial

Check for updates

All I Want for Christmas Is a Loo: Visualizations of Sex and Gender on Toilet Doors

Judith Rosmalen[1,2], Ilona Plug[3], Aranka Ballering[1]

[1] *Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands.* **[2]** *Department of Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands.* **[3]** *Department of Communication and Cognition, Tilburg School of Humanities and Digital Sciences, University of Tilburg, Tilburg, the Netherlands.*

Clinical Psychology in Europe, 2024, Vol. 6(4), Article e16159, https://doi.org/10.32872/cpe.16159

Published (VoR): 2024-12-20

Corresponding Author: Aranka Ballering, Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), Department of Psychiatry, University Medical Center Groningen, P.O. Box 30.001, Internal code CC72, 9700 RB, Groningen, the Netherlands. E-mail: a.v.ballering@umcg.nl

Supplementary Materials: Data, Materials [see Index of Supplementary Materials]

Preface

Ah, Christmas Eve. A time for carols, cozy fires, and, of course, culinary excellence. You've secured a reservation at one of your city's best restaurants, ready to pamper your taste buds with gastronomic delights, until it all takes an unexpected turn. You want to run from the table and not because of your eating companion's endless debates. Rather, your gut is staging a revolt, hinting at an encore performance of your prawn appetizer. Channeling your inner Olympic sprinting champion, you run to the restroom, breaking records and social decorum alike. And now, in front of the restroom doors, you face the ultimate dilemma: which door to choose? While your gut is rumbling along with Mariah Carey's song playing in the background, you can only think: "All I want for Christmas is a clue…"

We are all confronted with the importance of a nuanced understanding of both sex and gender when facing a toilet door in dire times, but this day-to-day realization does not always find its way into the realm of health research (Ballering et al., 2023). Although health-related research is becoming increasingly sex-sensitive (i.e., attentive towards the biology of intersex, male, and female bodies), sensitivity towards gender (i.e., attentiveness towards psychosocial factors related to being a woman, man, or non-binary identity) in research remains scarce (Ballering et al., 2024). This is problematic, as it ham-

pers the validity of conclusions drawn from these studies, as recent studies show that sex and gender independently affect health outcomes (Ballering et al., 2020; Mauvais-Jarvis et al., 2020).

Multiple resources are available to increase sex and gender-sensitivity in research, with the SAGER guidelines being most prominent (Heidari et al., 2016). However, we consider one key resource to be underused in raising gender-sensitivity in research: toilets. After all, we all choose on a daily basis which toilet door to open. This sensitivity to (the difference between) sex and gender in daily life, is ought to find its way to health research. Therefore, this study identified visual loo clues on toilet doors that guide the decision on which toilet to enter. We thematically analyzed, and subsequently quantified, whether these clues are related to sex or gender. We aim to sharpen the reader's understanding of sex and gender, thus facilitating increased sex and gender sensitivity in research. Also, this will aid readers so that the next time they find themselves in front of those puzzling restroom doors, they confidently choose the clue for the right loo without breaking a sweat (or a sprinting record).

Method

Data Collection

We collected pictures of international toilet doors in the period of 2020 to 2022. Initially, we visited and photographed as many toilet doors as possible whenever we were outside of our homes. We did not seek permission from an ethical committee, nor from the owners of the toilet doors. Consequently, we had to endure confused looks of others while we were taking these pictures – ah, the things we do in the name of science. We also searched online using a search string in Google containing commonly used Dutch and English variants of the terms 'toilet door(s)', 'toilet door sticker(s)', toilet icon(s)' and 'toilet entrance'. Duplicates were defined as different doors having exactly the same visual clues; these were excluded from the analyses. All raw visual data are available from OSF (Rosmalen et al., 2024S-a).

Coding of Visual Decision Clues as Related to Sex and/or Gender

Each set of toilet doors was independently analyzed by three female coders (mean age 35 [*SD* = 13.9]). The coders identified all visual elements on a given set of toilet doors that would aid the decision of which door to open in times of need. In some cases, predominantly for gender-neutral or gender-inclusive toilet doors, there was only one door to assess. Open coding of visual elements was used, followed by a thematic analysis. Consensus discussion resulted in a final definition of the visual clues per toilet door.

Visual clues were considered neutral, if these did not consistently represent either sex or gender, or as referring to either sex or gender. The three coders extensively

discussed all themes in order to reach consensus on their categorization as neutral, sex, or gender clues. Interrater agreement was assessed by means of free-marginal multi-rater kappa using the Online Kappa Calculator (http://justusrandolph.net/kappa/). The overall agreement between the three raters was 85.5%; the free-marginal kappa was 0.78 (95% CI [0.61, 0.95]).

Analysis

MS Office Excel was used to analyze the total number of clues and per set of doors.

Results

Study Sample

Appendix 1 in the Supplementary Materials shows the PRISMA statement. Our final study sample included 97 sets of toilet doors. Of these doors, six were sex-or-gender-neutral (i.e., not specifically related to women or men), unisex (i.e., specifically related to both female and male individuals) or even universal with visual clues that moved beyond sex and gender, as these depicted aged people, tattooed people, queer people, the Grim Reaper, people with various religions, aliens, fairies and mermaids.

Thematic Analysis

On the 97 doors, we identified a total of 260 visual decision clues that were categorized in 23 themes. The minimum and maximum number of clues per door were 1 and 8, with a median of 2 (IQR 1-4). The 23 themes were subdivided into neutral (n = 2), or related to sex (n = 9) or gender (n = 12). The identified visual themes and concomitant frequencies can be found in Appendix 2.

Of the 260 clues, 52.7% were gender-related and 30.0% were sex-related. The remaining 17.3% included text and icons that were categorized as neutral. Sex clues were found on 54.1% of all doors, gender clues on 72.4%.

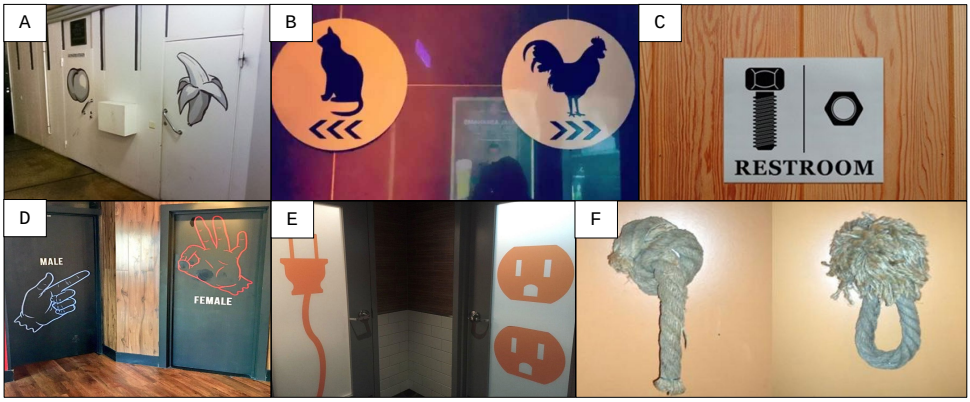
Neutrality and Naughty Nomenclature

Two themes, namely text elements and icons, were coded as neutral, since these did not consistently represent either sex or gender. Textual sex clues included variants of '*Male-Female*' or synonyms in different languages (e.g., '*Miehet-Naiset*', '*Femei-Barbati*' or '*Vrouwtjes-Mannetjes*'), or in different species (e.g., '*Bulls-Cows*'). We also categorized the text element '*Sausage-Eggs*' as a sex clue, assuming that decision clues on toilet doors are not reflecting food preferences. Gendered textual clues include names ('*Adam-Eve*'), cognitions ('*Shopping-Football*'), communication ('*Bla-Blablablabla*'), and assumed capabilities ('*Men to the left, because women are always right*')[1].

Icons predominantly included symbolic representations of the male and female reproductive organs (i.e., sex; Figure 1). Discussion was required for the doors in Figure 1B that contain a sophisticated symbol of the words used; an $N = 1$ pilot test in a young man suggested that these clues on the toilet door would take up too much time than preferred. One of the authors did not understand these doors either until the consensus discussion; as a linguist she interpreted these doors as representing gendered behavior: cocky (i.e., men-like) and catty (i.e., women-like).

Figure 1

Toilet Doors With Icons



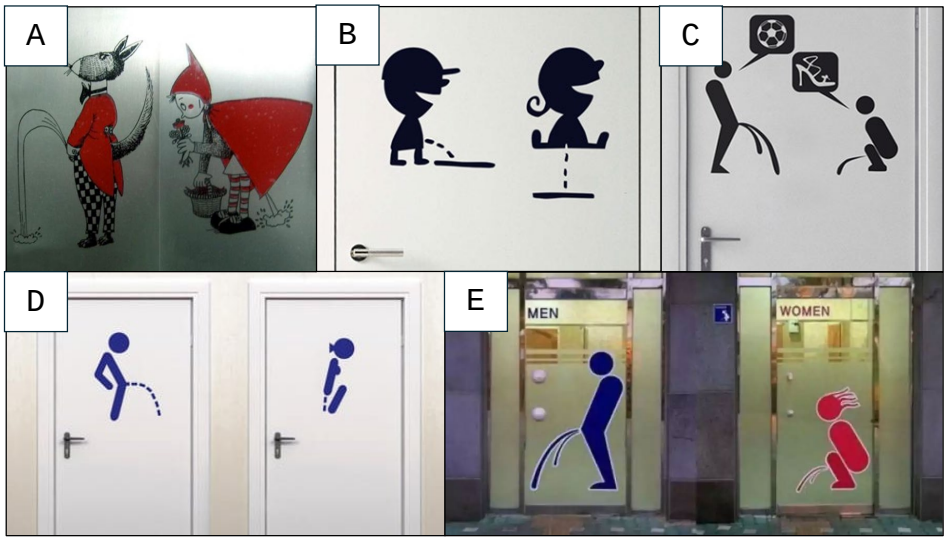
Peeing With Precision and the Angle of Male Confidence

Some visual clues required some more extensive discussion. An example hereof is the urinary stream theme (Figure 2). We initially assumed that the urinary stream was a sex clue, since it would be related to anatomical characteristics of men and women. However, it appeared more complex. We assessed the initial angle of the male urinary stream, based on the Visual Prostate Symptom Score (VPSS) (Stothers et al., 2017). Most male icons in Figure 2 depicted an angle that exceeds the highest possible score on the VPSS, namely a strong horizontal or even sky-facing stream. Potentially, this indicates male (over)confidence, which would be gender-related. A validated female equivalent of the VPSS is unfortunately not known to the authors. However, we noted a women-specific aspect that might influence a urinary stream: the women sit, bend and even jump while peeing (Figure 2). We hypothesize that this is due to the female practice of urinating without trying to touch the public toilet seat, which may be related to gender.

1) Clearly, the three coders, who identify as women, fully agree with the assumptions made on this toilet door.

Figure 2

Urinary Streams



A Case of the Jingle Balls

A second example that required somewhat more extensive discussion on the sex versus gender dimension of visual clues on toilet doors is provided in Figure 3. We initially assumed that the text element '*Balls-No balls'* in Figure 3A referred to sex (or potentially to how to decorate a Christmas tree). However, when studying doors with more direct graphical representations of balls, our interpretation appeared to be incorrect: Figure 3B shows similar balls for him and her, Figure 3C shows similar numbers of balls but qualitative differences, and Figure 3D suggests that no balls refers to men, which contrasts with Figure 3A. Since all graphical representations show that women have at least as many, or even more, balls than men, we assume that '*Balls-No balls*' in Figure 3A refers to feminine gender and masculine gender respectively, in which case we -as three women- would like to point out that the owner of the doors misplaced the male/female icons.

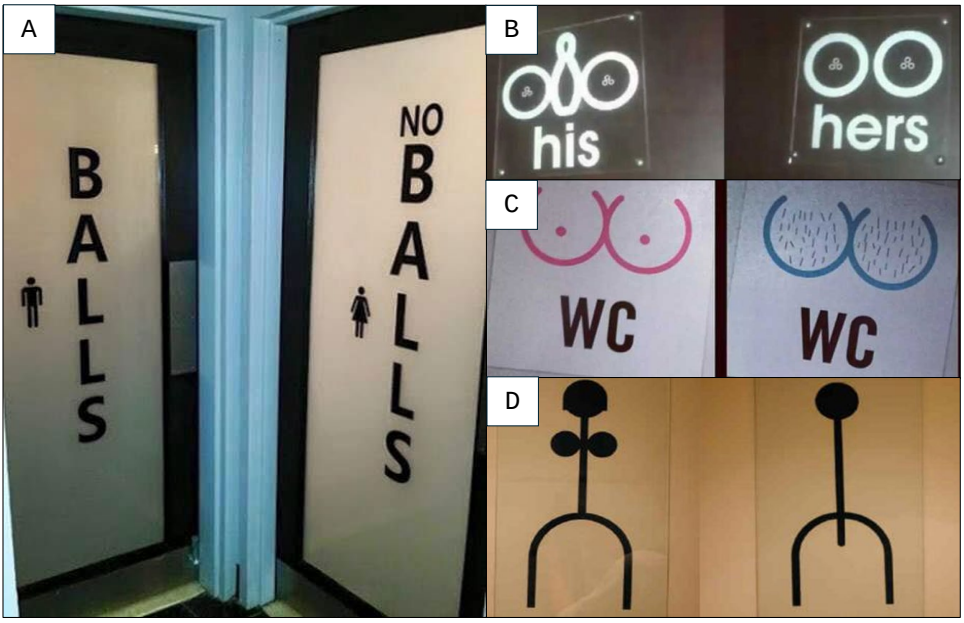
Discussion

Our study found that loo clues are primarily based on gender, as 52.7% of the clues depicted on toilet doors related to gender, while 30.0% related to sex. Twelve of the 23

identified themes related to gender and nine related to sex. Clues related to gender were found on 72.4% of the included toilet doors and sex clues on 54.1%.

Figure 3

Toilet Doors With Balls



Before interpreting our results, the following limitations need to be taken into account. First of all, data collection was performed by three women. We assume this has not influenced the offline search strategy, since we collected visual clues outside of the restrooms. One could argue that the female-oriented online search histories would have influenced the online search results, but the searches were done on computers that were also used by men. Secondly, selection bias may have occurred: it is likely that either extremely gendered or sexed toilet doors that were considered funny or interesting enough to be put online were found via our online searches. Lastly, the embodiment of gender, and what is considered feminine and masculine, depends strongly on time, place, and culture potentially affecting the interpretation of loo clues.

To the best of our knowledge, little to no previous studies have systematically assessed and quantified the clues present on toilet doors with regards to differentiating sex and gender (Iio, 2023). It has been previously argued that many toilet signs have a symbolic function (Iio, 2019). In our study, the majority of signage does not depict the

toilet itself (only 4 clues) or the activity of using the facility (7 clues related to urinary stream and 10 to urinating position, often combined). Therefore, the signage of toilets frequently involves learned behavior and culturally accepted conventions that tie closely with gendered segregation of men and women, and use of toilet facilities (Ciochetto, 2003).

Previous research argues that explicitly showing the act of urinating or defecating itself is seldom reflected in toilet signage. Potentially, this may be culturally inappropriate (Ciochetto, 2003). However, our results contrast with this view. Especially with regards to the VPSS, the results also suggest that toilets are not very inclusive for men with prostate problems. This may be attributed to the stoicism related to masculinity, which results in men not easily admitting physical complaints or bodily disturbances (Ballering et al., 2021; MacLean et al., 2010). Toilet signage involving men with prostate problems may not be attractive for men without problems.

We conclude that current toilet signage predominantly uses gender clues to indicate who should enter. Further studies could investigate the amount of time it takes to make a decision based on sex and gender clues, in relation to attitudes towards sex and gender characteristics. As a practical implication, we suggest for all to take a second look at the loo clues and try to identify the sex and/or gender clues before entering a toilet, if the urgency allows the time investment[2]. This facilitates a learning experience about the difference between sex and gender and may subsequently increase sex and gender sensitivity in research as well. But, perhaps more importantly, you can now enjoy your holiday dinner in peace, knowing that the only dilemma left to solve this Christmas is whether to indulge in another slice of pie—without a hint of restroom anxiety in sight.

2) If the urgency is relatively low, please consider snapping pictures of interesting toilet doors and send these to the corresponding author. Especially doors with Christmas trees, Santa Claus, and reindeers are very welcome.

Funding: JR, AB and IP have received funding from ZonMw (Gender and health project number 849200013), AB additionally received funding from ZonMw (project number: 849800001, 50018423 and 04520232330002).

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: The authors have no conflicts of interest to declare.

Author Contributions: JR had the original idea, conducted the statistical analysis and wrote a first draft of the article. JR, IP and AB together conducted the qualitative analysis. IP and AB helped to interpret results. AB edited the final version of the paper.

Ethics Statement: The study used routinely collected, non-identifiable data, and ethics approval was not required.

Data Availability: All raw visual data are available from OSF (see Rosmalen et al., 2024S-a).

Supplementary Materials

The Supplementary Materials contain the following items:

- All raw visual data (Rosmalen et al., 2024S-a)
- The online appendices (Rosmalen et al., 2024S-b):
  - Appendix 1: Prisma flow chart
  - Appendix 2: Frequency and categorization of the identified themes on toilet doors

Index of Supplementary Materials

Rosmalen, J., Plug, I., & Ballering, A. (2024S-a). *All I want for Christmas...* [Research data]. OSF. https://osf.io/gyqek/

Rosmalen, J., Plug, I., & Ballering, A. (2024S-b). *Supplementary materials to "All I want for Christmas is a loo: Visualizations of sex and gender on toilet doors"* [Online appendices]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.15751

References

Ballering, A. V., Bonvanie, I. J., Olde Hartman, T. C., Monden, R., & Rosmalen, J. G. M. (2020). Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine, 253*, Article 112968. https://doi.org/10.1016/j.socscimed.2020.112968

Ballering, A. V., Burke, S. M., Maeckelberghe, E. L. M., & Rosmalen, J. G. M. (2023). How to ensure inclusivity in large-scale general population cohort studies? Lessons learned with regard to including and assessing sex, gender, and sexual orientation. *Archives of Sexual Behavior, 52*(5), 2163–2172. https://doi.org/10.1007/s10508-023-02600-y

Ballering, A. V., Muijres, D., Uijen, A. A., Rosmalen, J. G. M., & Olde Hartman, T. C. (2021). Sex differences in the trajectories to diagnosis of patients presenting with common somatic symptoms in primary care: An observational cohort study. *Journal of Psychosomatic Research, 149*, Article 110589. https://doi.org/10.1016/j.jpsychores.2021.110589

Ballering, A. V., Olde Hartman, T. C., & Rosmalen, J. G. M. (2024). Gender scores in epidemiological research: Methods, advantages and implications. *The Lancet Regional Health – Europe, 43*, Article 100962. https://doi.org/10.1016/j.lanepe.2024.100962

Ciochetto, L. (2003). Toilet signage as effective communication. *Visible Language, 37*(2), 208–221.

Heidari, S., Babor, T. F., De Castro, P., Tort, S., & Curno, M. (2016). Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review, 1*, Article 2. https://doi.org/10.1186/s41073-016-0007-6

Iio, J. (2019). The semiotics of toilet signs. In M. Kurosu (Ed.), *Human-computer interaction: Design practice in contemporary societies* (pp. 285–293). https://doi.org/10.1007/978-3-030-22636-7_20

Iio, J. (2023). The restroom sign database. *Bulletin of Human Centered Design Organization, 19*(1), 57–62. https://doi.org/10.34404/hcd.19.1_57

MacLean, A., Sweeting, H., & Hunt, K. (2010). 'Rules' for boys, 'guidelines' for girls: Gender differences in symptom reporting during childhood and adolescence. *Social Science & Medicine, 70*(4), 597–604. https://doi.org/10.1016/j.socscimed.2009.10.042

Mauvais-Jarvis, F., Bairey Merz, N., Barnes, P. J., Brinton, R. D., Carrero, J.-J., DeMeo, D. L., De Vries, G. J., Epperson, C. N., Govindan, R., Klein, S. L., Lonardo, A., Maki, P. M., McCullough, L. D., Regitz-Zagrosek, V., Regensteiner, J. G., Rubin, J. B., Sandberg, K., & Suzuki, A. (2020). Sex and gender: Modifiers of health, disease, and medicine. *Lancet, 396*(10250), 565–582. https://doi.org/10.1016/S0140-6736(20)31561-0

Stothers, L., Macnab, A., Bajunirwe, F., Mutabazi, S., & Lobatt, C. (2017). Comprehension and construct validity of the visual prostate symptom score by men with obstructive lower urinary tract symptoms in rural Africa. *Canadian Urological Association Journal, 11*(11), E405–E408. https://doi.org/10.5489/cuaj.4589

Clinical Psychology in Europe

PsychOpen GOLD

Systematic Reviews and Meta-analyses

Effectiveness of Attachment-Based Family Therapy for Suicidal Adolescents and Young Adults: A Systematic Review and Meta-Analysis

Poul M. Schulte-Frankenfeld[1,2], Josefien J. F. Breedvelt[3,4], Marlies E. Brouwer[1], Nadia van der Spek[1], Guy Bosmans[5], Claudi L. Bockting[1,4,6]

[1] *Department of Psychiatry, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, The Netherlands.* **[2]** *Department of Pediatric Neurology, Charité – Universitätsmedizin Berlin, Berlin, Germany.* **[3]** *Department of Child and Adolescent Psychiatry, Institute for Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom.* **[4]** *Centre for Urban Mental Health, University of Amsterdam, Amsterdam, The Netherlands.* **[5]** *Department of Clinical Psychology, KU Leuven, Leuven, Belgium.* **[6]** *Amsterdam Public Health, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, The Netherlands.*

Clinical Psychology in Europe, 2024, Vol. 6(4), Article e13717, https://doi.org/10.32872/cpe.13717

Received: 2024-01-15 • **Accepted:** 2024-08-12 • **Published (VoR):** 2024-12-20

Handling Editor: Winfried Rief, Philipps-University of Marburg, Marburg, Germany

Corresponding Author: Claudi L. Bockting, Meibergdreef 5, 1105 AZ Amsterdam, The Netherlands. phone: +31 20 566 9111. E-mail: c.l.bockting@amsterdamumc.nl

Supplementary Materials: Code, Data, Materials, Preregistration [see Index of Supplementary Materials]

Abstract

Background: Suicide is a leading cause of death among adolescents and young adults. While only few evidence-based treatments with limited efficacy are available, family processes have recently been posed as a possible alternative target for intervention. Here, we review the evidence for Attachment-Based Family Therapy (ABFT), a guideline-listed treatment targeting intrafamilial ruptures and building protective caregiver-child relationships.

Method: PubMed, PsycINFO, Embase, and Scopus were searched for prospective trials on ABFT in youth published up until November 6th, 2023, and including measures of suicidality. Results were independently screened by two researchers following PRISMA guidelines. Risk of bias was assessed using the Cochrane RoB-2 framework. A random effects meta-analysis was conducted on suicidal ideation and depressive symptoms post-intervention scores in randomized-controlled trials (RCTs).

Results: Seven articles reporting on four RCTs (n = 287) and three open trials (n = 45) were identified. Mean age of participants was M_{pooled} = 15.2 years and the majority identified as female

(~80%). Overall, ABFT was not significantly more effective in reducing youth suicidal ideation, g_{pooled} = 0.40, 95% CI [-0.12, 0.93], nor depressive symptoms, g_{pooled} = 0.33, 95% CI [-0.18, 0.84], compared to investigated controls (Waitlist, (Enhanced) Treatment as Usual, Family-Enhanced Nondirective Supportive Therapy).

Conclusion: Evidence is strongly limited, with few available trials, small sample sizes, high sample heterogeneity, attrition rates, and risk of bias. While not generally superior to other treatments, ABFT might still be a clinically valid option in specific cases and should be further investigated. Clinicians are currently recommended to apply caution when considering ABFT as stand-alone intervention for suicidal youth and to decide on a case-by-case basis.

Keywords

suicide, suicidal ideation, adolescents, young adults, psychotherapy, family therapy

Highlights

- Current evidence on the comparative effectiveness of ABFT is strongly limited.
- Meta-analysis showed no improved effect on suicidal ideation nor depressive symptoms compared to other treatments.
- ABFT might still be a clinically valid option in specific cases and should be further investigated.
- Implications and recommendations for researchers and practitioners are discussed in the article.

Suicide is the fourth most prevalent cause of death in adolescents and young adults worldwide (World Health Organization, 2021). In 2019, more than 157.700 youths aged 15 to 29 died prematurely by intentional self-harm, accounting for about 8% of all deaths and the majority of Years of Life Lost in this cohort (Castelpietra et al., 2022; World Health Organization, 2020). Many of these cases are preceded by risk factors and early warning signs such as depressive symptoms, being part of a sexual minority group, expression of suicidal ideas, previous self-harm, or contact with primary health services, and could thus potentially be prevented through targeted treatment (Bachmann, 2018; Chiang et al., 2021; Fergusson et al., 2005; Jackman et al., 2021; Liu et al., 2020; Walby et al., 2018). Several psychological interventions that might help to attenuate the suicidal trajectory are available (e.g., Cognitive Behavioral Therapy (CBT), Dialectical Behavior Therapy (DBT)). However, efficacy of these interventions for adolescent populations was found to be low overall and has stagnated for decades (Busby et al., 2020; Fox et al., 2020; Franklin et al., 2017; Kothgassner et al., 2020; Robinson et al., 2018; Tarrier et al., 2008). Thus, developing more effective interventions to prevent adolescent suicide is warranted (Holmes et al., 2018).

Increasing attention has recently been paid to the role that early childhood interpersonal and relational factors might play in the development and trajectory of mental

health problems and suicidal thoughts and behaviors (Chu et al., 2017; Franzoi et al., 2024; Van Orden et al., 2010). Having strong and positive interpersonal connections to family members and friends was found to be protective in this regard, and the absence of such resulting in feelings of loneliness was a risk factor respectively (Gunn et al., 2018; McClelland et al., 2020). At the same time, disturbing aspects of interpersonal relations and familial environments, such as conflicts, constant negative evaluation, childhood trauma, physical abuse, and neglect, can be perceived as major stressors by individuals at risk and contribute to the suicidal trajectory (Carballo et al., 2020; King & Merchant, 2008). Although accumulating evidence suggests that interpersonal relations, specifically in the context of adolescents' family environments, could be strong targets for preventive interventions, there are only few family-focused interventions available (Frey et al., 2022; Sullivan et al., 2023).

One such family-focused intervention for suicidal youths is Attachment-Based Family Therapy, which targets interpersonal ruptures between youth at risk of suicide and their primary caregivers (G. S. Diamond, 2022; G. S. Diamond et al., 2014). Developed upon the assumption that the quality of familial relations can trigger, exacerbate, and buffer against suicide trajectories, this 16-week treatment protocol addresses interpersonal traumata and dysfunctional interaction patterns in a therapist-guided systematic process. In contrast to currently prevalent treatments mainly targeting patients' thought patterns and behavior (e.g., Cognitive Behavioral Therapy, Dialectical Behavior Therapy), the five critical treatment tasks in ABFT focus on identifying factors that damaged intrafamilial trust, motivating patients and caregivers to rediscover their innate desire for mutual closeness, and building mature, regulated, and empathic interaction patterns. Adolescents' autonomy and developmental responsibility are encouraged, and caregivers are supported in developing an empathic, empowering, and unconditionally accepting stance toward their youths. By the end of treatment, attachment security is expected to be improved, building the foundation for future positive development and alleviation of symptoms. A full description of underlying theory and mechanisms can be found in the original manual and a more recent review (G. S. Diamond et al., 2014, 2016).

Attachment-Based Family Therapy is currently listed as a 'promising' intervention for youth depression and suicidal ideation at the California Evidence-based Clearing House for Child Welfare, with ratings based on three randomized-controlled trials and four moderation studies published up until 2018 (G. S. Diamond et al., 2002, 2010, 2012; Shpigel et al., 2012; Israel & Diamond, 2013; Feder & Diamond, 2016; Ibrahim et al., 2018; CEBC, 2020). It is further mentioned as a possible intervention for youth suffering from moderate to severe depression in the NICE guidelines (National Institute for Health and Care Excellence, 2019). In recent years, increasing interest in interpersonal treatments for youth suicide bore new evidence, and new results ought to be considered in clinical guidelines and practice. Importantly, more recently published reviews on ABFT remain mostly narrative and do not provide a systematic approach to its evaluation, or do

not examine the efficacy of ABFT for youth suicidality specifically (G. S. Diamond et al., 2016, 2021; Sullivan et al., 2023; van Aswegen et al., 2023). One recently published meta-analysis on the effect of family-based treatments in youth finds a significant overall treatment benefit of family-therapy over comparator therapies for suicidal ideation, but not depressive symptoms (Waraan et al., 2023). However, this review pools together different types of family-based therapies (e.g. Family-focused Cognitive Behavioral Therapy, Systems Integrative Family Therapy), thus not allowing to assess the individual efficacy of ABFT. Therefore, this systematic review and meta-analysis aims to summarise and assess the current evidence for the effectiveness of Attachment-Based Family Therapy in treating suicidal adolescents and young adults.

Method

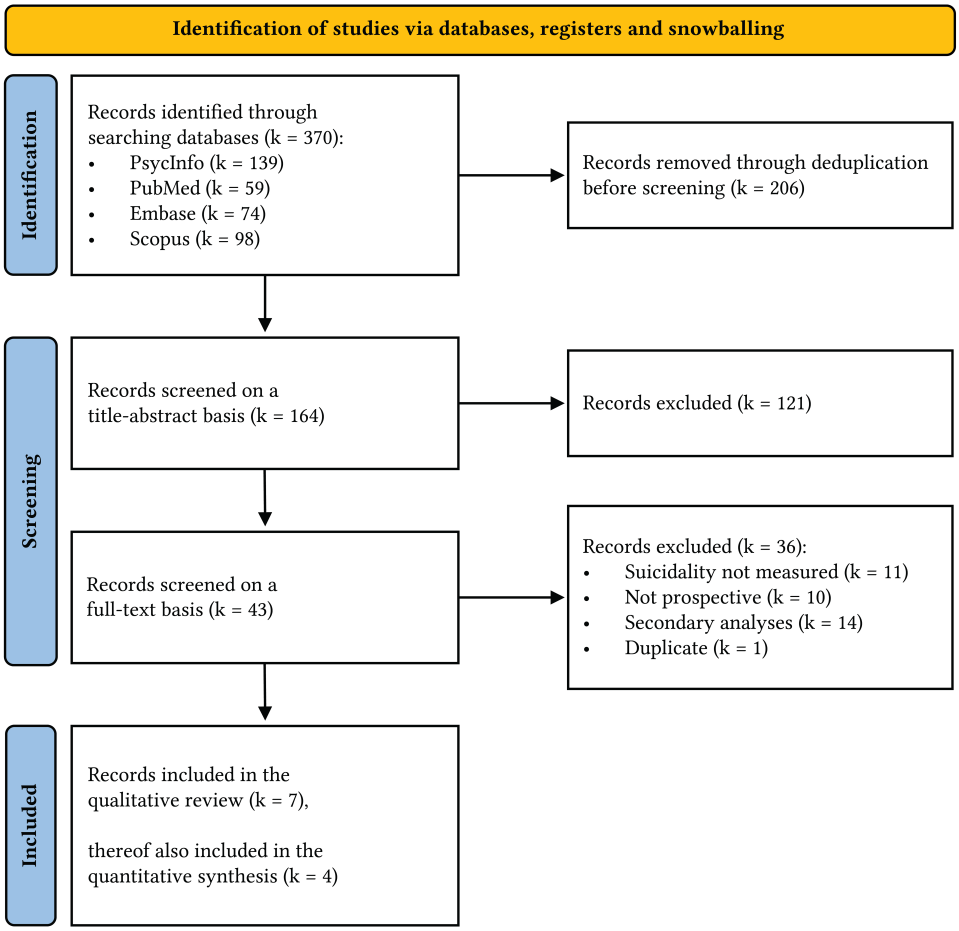
Search Strategy and Selection Criteria

This review follows the PRISMA guidelines, was initially performed in August 2021 and updated on November 6th, 2023, and preregistered at PROSPERO (CRD42021271731, see Schulte-Frankenfeld et al., 2024S-a). PubMed (Ovid MEDLINE® All 1946 to November 03, 2023), PsycINFO (APA PsycInfo 1806 – October Week 4 2023), and Embase (Embase Classic + Embase, 1946 – 2023 November 03) were searched for trials indexed for the search term "attachment-based family therapy" up until November 06th, 2023, using the Ovid interface (version 05.09.00.005). Scopus (November 06, 2023) was searched in titles, abstracts, and keywords. No further filters were applied. Snowballing was performed to identify non-indexed studies by examining published reviews, reference lists of included trials, and contacting expert authors. Search results were deduplicated using ProQuest Refworks.

Two reviewers independently screened all records based on titles and abstracts using Rayyan (Ouzzani et al., 2016). Studies were eligible if they fulfilled the following criteria: (1) sample of adolescents and young adults aged ten years and older; (2) ABFT treatment; (3) suicidality measured at post-treatment through diagnostic methods (e.g., clinical interviews, SIQ-JR, BSS, SIDAS); (4) longitudinal or prospective design; (5) written in English, Dutch, or German. Review papers, expert opinions, and case reports were excluded. Potentially eligible records were further assessed on a full-text basis. Risk of bias of eligible records was assessed using the Cochrane RoB-2 framework (Appendix A, Schulte-Frankenfeld et al., 2024S-c). Consensus about deviant decisions was reached by unblinded discussion and consulting a senior author. A list of screened articles (Schulte-Frankenfeld et al., 2024S-b) and a PRISMA Flow Diagram (Figure 1) are available.

Figure 1

PRISMA Flow Diagram for the Systematic Review



Data Extraction and Statistical Analysis

Primary outcome was the difference in suicidality severity at post-treatment between ABFT and respective control conditions. Further measures of the patient's idiosyncratic condition (e.g. BDI-II, HAM-D) were collected if available. Demographic information (population type, age, country), study characteristics (*N*, design, conditions), outcome measures, and statistical factors necessary to calculate effect sizes were extracted by one author and independently checked by two others. All authors of included articles were contacted to request Individual Participant Data, and additionally summary statistics when reported outcome data was insufficient to calculate effect sizes. When

measurement data from multiple post-baseline timepoints was given, the timepoint most immediate to the end of treatment was used. Recalculating effect sizes based on means and standard deviations was preferred, and data transformations were applied following Cochrane guidelines where necessary (Higgins et al., 2023).

Calculations were performed using RStudio in version 2023.03.0+386 following guidelines by Harrer and colleagues (Harrer et al., 2021; RStudio Team, 2021). *Hedge's g* was used as effect size measure to correct for small sample bias, expressing the magnitude of difference between group means in units of pooled standard deviation. Although it is commonly interpreted with cut-offs (small (< 0.2), medium (~0.5), and large (> 0.8)) these thresholds are arbitrary and lack solid statistical and empirical foundations. Therefore, they should be interpreted only as rough guidelines within this review (Cuijpers et al., 2014). A random effects model using the Hartung-Knapp method was applied as considerable heterogeneity among studies was expected (van Aert & Jackson, 2019; Zeraatkar et al., 2020). Cochran's Q was used to assess homogeneity ($\alpha < .05$), and I^2 to quantify how much variability can be attributed to between-study heterogeneity (low: ≤ 25%, medium: 26%–50%, high: > 50%). A prediction interval (95% PI) was calculated to estimate the range in which true effect sizes of future replication studies are expected to fall (IntHout et al., 2016). Publication bias was assessed through visual inspection of colour-enhanced funnel plots. The {tidyverse} package v2.0.0 was used for data processing, and the {esc} package v0.5.1 for calculating effect sizes (Lüdecke, 2019; Wickham et al., 2019). Pooled effect sizes and 95% CIs, measures of heterogeneity, forest plots, and funnel plots were calculated using the {meta} package v6.5-0 (Balduzzi et al., 2019). Risk of bias plots were created using the {robvis} package v0.3.0 (McGuinness, 2019). All data and calculation scripts are publicly available (Schulte-Frankenfeld et al., 2024S-b).

Results

Study Selection

PubMed, PsycINFO, Embase, and Scopus were searched for prospective trials indexed for "attachment-based family therapy", measuring suicidality at post-treatment in a sample of adolescents and young adults aged ten years or older, up until November 06th, 2023. Out of 370 records identified, seven articles met the inclusion criteria (n = 332), of which four were included in the meta-analysis (n = 287) (G. M. Diamond et al., 2012; G. S. Diamond et al., 2002, 2010, 2019; Russon et al., 2022; van der Spek et al., 2024; Waraan et al., 2021). See Figure 1 for the selection process. Individual Participant Data were requested from all authors, though only provided by one team despite repeated outreach attempts. Thus, an Individual Participant Data Meta-Analysis was not performed. One study reported SIQ-JR data deviating between text, figures, and appendix, which only matched after multiplying reported summary scores by the number of items in the un-

derlying scale (Waraan et al., 2021). Another study reported confidence intervals instead of standard deviations, for which a *t*-distribution conversion following the Cochrane Handbook was applied (G. S. Diamond et al., 2010; Higgins et al., 2023).

Study Characteristics

Characteristics of included studies are depicted in Table 1. Included trials recruited a total of 332 participants (RCTs n = 287, Open Trials n = 45) with a mean age of 15.2 years (SD_{pooled} = 1.65, *range:* 12 – 25 years) and the majority identifying as female (~80%). Baseline levels of suicidal ideation, SIQ(-JR), M_{pooled} = 46.26, SD_{pooled} = 17.91, and depressive symptoms BDI(-II), M_{pooled} = 31.59, SD_{pooled} = 9.00, were high across trials. Heterogeneity between studies regarding populations, treatment protocols, study designs, and implemented control conditions was considerable. Four of the seven studies were randomized-controlled trials using the 12-week (k = 2) (G. S. Diamond et al., 2002, 2010) or 16-week (k = 2) (G. S. Diamond et al., 2019; Waraan et al., 2021) ABFT protocol, and a waitlist group (k = 1) (G. S. Diamond et al., 2002), Treatment as Usual (TAU) (Waraan et al., 2021), directly referred "enhanced" usual care (Enhanced Usual Care; EUC) (G. S. Diamond et al., 2010), or a nondirective supportive therapy augmented with psychoeducative sessions for primary caregivers (Family-Enhanced Nondirective Supportive Therapy; FE-NST) (G. S. Diamond et al., 2019) as control condition. TAU and EUC trials did not follow-up participants for the type of treatment received (G. S. Diamond et al., 2010; Waraan et al., 2021). The open non-controlled trials consisted of one study using the 16-week ABFT protocol, and two applying LGBTQ+ sensitive variants of ABFT to sexual minority youths (G. M. Diamond et al., 2012; Russon et al., 2022; van der Spek et al., 2024). Clinically significant scores of suicidal ideation were a recruitment criterion in five studies (RCTs k = 2; Open Trials k = 3), while two RCTs primarily recruited adolescents with depressive symptoms (G. S. Diamond et al., 2002; Waraan et al., 2021). Overall, trials were small, with samples ranging from 10 to 129 participants. A priori calculated recruitment targets were only met by one study (G. S. Diamond et al., 2019). One study presented a post-intervention attrition rate of ~80%, raising concerns about its validity (Waraan et al., 2021). After conducting a sensitivity analysis, we kept it for completeness. Risk of bias was high in most studies, with non-blinded assessment and selective data reporting as major issues (Appendix A and B, Schulte-Frankenfeld et al., 2024S-c).

Effect Sizes and Meta-Analyses

Suicide attempts and suicidal ideation, as assessed through the Suicidal Ideation Questionnaire (SIQ), were the only shared measure of suicidality between studies. Since the number of recorded attempts was too low for statistical analysis (see Table 1), SIQ scores were used to compare effect sizes.

Clinical Psychology in Europe
2024, Vol. 6(4), Article e13717
https://doi.org/10.32872/cpe.13717

Table 1

Characteristics of Studies Included in the Systematic Review

Study	Population	Inclusion criteria	*N*	Age, *M* (*SD*, *range*)	Gender, % Female	Location	Design	Conditions	Measure of effect	Secondary outcomes	Attrition rate	Adherence, *M* (*SD*, *range*)	Baseline Severity, *M* (*SD*)	Effect Size,g, [95% CI]	Clinical Recovery	Suicidal Behavior	Conclusion	Risk of Bias
Diamond et al., 2002[a]	Depressed youth, referred by schools and caregivers	MDD primary diagnosis according to DSM-III-R	32	14.9 (1.5, *range*: 13 - 17)	78%	USA	Pilot RCT	*ABFT:* 12 weeks ABFT *Control:* 6 weeks waitlist	SIQ	BDI, HAM-D, BHS, STAIC	3%	*NA*	*SIQ:* 32.10 (20.37) *BDI:* 25.9 (7.45)	*time*: 0.66 [-0.05, 1.38] *group*: 0.37 [-0.34, 1.08]	*NA*	*NA*	No effect over time, not superior to waitlist	High
Diamond et al., 2010[a]	Suicidal youth, referred by primary care and emergency rooms	SIQ-JR > 31, BDI-II > 20	66	15.1 (1.5, *range*: 12 - 17)	83%	USA	RCT	ABFT: 12 weeks ABFT *Control:* 12 weeks EUC	SIQ-JR	SSI, BDI-II	9%	*ABFT:* 9.71 (5.26) *Control:* 2.87 (3.3) Significant group difference (Z = -4.74, p < .001)*	*SIQ-JR:* 51.07 (13.45) *BDI-II:* 33.00 (9.03)	*time*: 3.91 [3.08, 4.74]* *group*: 0.83 [-0.30, -1.32]*	Remission (≤ 13): 87.1% *OR* = 6.30 [1.76 – 22.61]*	Suicide attempts (up to 6 months after baseline): *ABFT:* 11% (n = 4) *Control:* 23% (n = 7)	Effect over time, superior to EUC	High
G. M. Diamond et al., 2012	Suicidal LGB youth	LGB identification, SIQ-JR ≥ 31	10	15.1 (1.37, *range*: 14 - 18)	80%	USA	Open Trial	*ABFT:* 12 weeks ABFT-LGB	SIQ-JR	BDI-II	20%	*ABFT: NA* (*range*: 8 – 16)	*SIQ-JR:* 51.00 (13.00) *BDI-II:* 28.10 (13.63)	*time*: 3.86 [2.24, 5.47]*	Remission (≤ 13): 87.5%	*NA*	Effect over time	*NA*
Diamond et al., 2019[a]	Suicidal youth, referred by primary and emergency care, inpatient clinics, schools, and self-referral	SIQ-JR ≥ 31, BDI-II > 20	129	14.87 (1.68, *range*: 12 - 18)	81.9%	USA	RCT	*ABFT:* 16 weeks ABFT *Control:* 16 weeks FE-NST	SIQ-JR	BDI-II	18%	*ABFT:* 14.34 (7.58) *Control:* 12.67 (5.74) No significant group difference, t_{127} = -1.43, p = .16	*SIQ-JR:* 49.83 (15.08) *BDI-II:* 30.59 (7.94)	*time*: 1.91 [1.47, 2.34]* *group*: 0.17 [-0.21, 0.55]	Remission (< 12): 32.7% vs. 24.4%	Suicide attempts (during treatment): *ABFT:* 3% (n = 2) *Control:* 8% (n = 4)	Effect over time, not superior to FE-NST	Some concern
Waraan et al., 2021[a]	Depressed youth, referred to mental health clinics	Current MDE according to clinical interview, GRID-HAMD > 15	60	14.9 (1.35, *range*: 13 - 18)	87%	Norway	RCT	*ABFT:* 16 weeks ABFT *Control:* 16 weeks TAU	SIQ-JR	GRID-HAMD, BDI-II	80%	*NA*	*SIQ-JR:* 43.09 (22.96) *BDI-II:* 35.22 (8.66)	*time*: 0.64 [-0.32, 1.61] *group*: 0.10 [-0.66, 0.86]	*NA*	*NA*	No effect over time, not superior to TAU	High
Russon et al., 2022	Suicidal LGBTQ+ youth	LGBTQ+ identification, SIQ-JR ≥ 31, BDI-II ≥ 21	10	18.2 (*NA*, *range*: 15 - 25)	NA	USA	Open Trial	*ABFT:* 16 weeks ABFT-LGBTQ+	SIQ-JR	BDI-II	0%	*ABFT:* 15.3 (*range*: 11 – 23)	*SIQ-JR:* 52.70 (19.36) BDI-II: 35.40 (11.92)	*time*: 1.45 [0.40, 2.50]*	Remission (≤ 13): 11%	Suicide attempts (during treatment): *ABFT:* 11% (n = 1)	Effect over time	*NA*
van der Spek et al., 2024	Depressed and/or suicidal youth referred to an outpatient clinic	Clinical assessment, depressive and suicidal symptoms	25	17.1 (*NA*, *range*: 12 - 23)	74.2%	Netherlands	Open Trial	*ABFT:* 16 weeks ABFT	SIQ-JR	CDI-2	32%	*ABFT:* 89% completed 10 or more sessions	SIQ-JR: 36.31 (11.93) CDI-2: 32.07 (11.48)	*time*: 0.72 [0.09, 1.35]*	*NA*	Suicide attempts (during treatment): *ABFT:* 14% (n = 4)	Effect over time	*NA*

Note. Effect sizes for primary outcomes calculated as within-group effect over time for the ABFT group (*time*) and between groups at post-treatment (*group*). Sample size (*N*) represents the total number of participants included in the respective trial, the number of available datapoints per measure used to calculate effect sizes can deviate. *ABFT* indicates values for the ABFT intervention group, *Control* for the control group respectively. Attrition rate was calculated as the number of datapoints at post-treatment compared to baseline for the SIQ(-JR) scale as the primary outcome. Adherence was assessed as the amount of completed treatment sessions. Baseline severity represents outcome statistics for the total sample estimated from group-wise baseline scores. ABFT = Attachment-Based Family Therapy; ABFT-LGB = LGB sensitive variant of ABFT; ABFT-LGTBQ+ = LGBTQ+ sensitive variant of ABFT; BDI-III = Beck Depression Inventory-II; CDI-2 = Children's Depression Inventory 2nd Edition; EUC = Enhanced Usual Care; FE-NST = Family-Enhanced Nondirective Supportive Treatment; GRID-HAMD = GRID Hamilton Rating Scale for Depression; LGB = lesbian, gay and bisexual; MDD = Major Depressive Disorder; MDE = Major Depressive Episode; NA = not available; SIQ-JR = Suicidal Ideation Questionnaire-Junior; SSI = Scale for Suicidal Ideation; TAU = Treatment as Usual.

[a]Studies included in the meta-analysis.

*confidence interval excluding zero indicating significance or p < .05.

The SIQ scale is a commonly applied self-administered questionnaire appropriate for adolescents around Grade 10 – 12, consisting of 30 items measuring mental distress and suicidal intend. Participants receive a list of "thoughts that people sometimes have" and are instructed to "indicate which of these thoughts [they] have had in the past month" on a scale from 0 "I never had this thought" to 6 "almost every day" (Reynolds, 1987). The SIQ-JR scale is an adapted version of the SIQ, consisting of a 15-item subset appropriate for adolescents around Grade 7 – 9 (Reynolds & Mazza, 1999). Scores consist of summed item responses, thus ranging from 0 to 180 for the SIQ and 0 to 90 for the SIQ-JR respectively, with higher scores indicating higher levels of mental distress and suicidal intend.

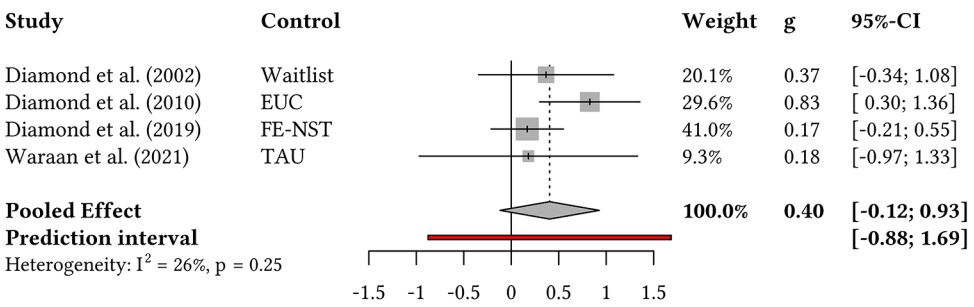
Out of seven studies, four trials presented moderate to large within-group changes in SIQ(-JR) scores over time, g = 0.72 to 3.91 (G. M. Diamond et al., 2012; G. S. Diamond et al., 2002, 2019; van der Spek et al., 2024). When compared with controls at post-treatment, a significant difference between treatment and comparator groups was present in one out of four RCTs, g = 0.83, 95% CI [0.30, 1.36], whereby the pooled effect size was large and reported to remain stable after 24 weeks (G. S. Diamond et al., 2010). Trials utilizing other comparators did not reveal significant group differences in SIQ(-JR) scores at post-treatment, and follow-up data was unavailable (G. S. Diamond et al., 2002, 2019; Waraan et al., 2021).

A meta-analysis was performed on SIQ(-JR) scores. Four trials (n = 287) were included in this synthesis, the results of which are depicted in Figure 2. The pooled effect size was g_{pooled} = 0.40 with a confidence interval including zero, 95% CI [-0.12, 0.93], indicating that SIQ(-JR) scores of participants in the ABFT conditions did not significantly differ from those in the associated control conditions at post-intervention. The prediction interval spanning zero, 95% PI [-0.88, 1.69], suggests that ABFT will likely not benefit some patients in future trials with similar conditions based on current evidence. Variation between effect sizes attributed to study heterogeneity was moderate, with I^2 = 26% and Cochran's Q being non-significant, p = 0.25. Visual inspection of the colour-enhanced funnel plot (Appendix C, Schulte-Frankenfeld et al., 2024S-c) did not indicate asymmetry. A sensitivity analysis restricted to active-control trials (EUC, FE-NST, TAU) yielded similar results, g_{pooled} = 0.42, 95% CI [-0.55, 1.39], I^2 = 51%, and Cochran's Q being non-significant, p = 0.13. Subgroup analyses were not performed due to the low number of included studies (Cuijpers et al., 2021).

Since all studies also measured the intervention's effect on depressive symptoms, performing a secondary meta-analysis on BDI(-II) scores was possible. The Beck's Depression Inventory (BDI) and its revised version, the BDI-II scale, are widely used 21-item self-report rating inventories measuring characteristic attitudes and symptoms of depression (Beck, 1961; Beck et al., 1996). Ratings per item range from 0 (not present) to 3 (strongly present) and scores are calculated by summing all item ratings, with higher scores indicating higher levels of depressive attitudes and symptoms. In this

Figure 2

Forest Plot on the Effect of ABFT Versus Control Treatments for Reducing SIQ(-JR) Scores



Note. Effect sizes for SIQ(-JR) scores calculated between groups at post-treatment as Hartung-Knapp corrected Hedges' *g* values. EUC, Enhanced Usual Care; FE-NST, Family-Enhanced Nondirective Supportive Therapy; TAU, Treatment as Usual.

meta-analysis of post-intervention BDI(-II) scores across trials, the pooled effect size was g_{pooled} = 0.33 with a confidence interval including zero, 95% CI [-0.18, 0.83], indicating that BDI(-II) scores of participants in the ABFT condition did not significantly differ from those in the associated control conditions. The prediction interval includes zero, 95% PI [-0.83, 1.49]. Heterogeneity was moderate, with I^2 = 24% and Cochran's Q non-significant, p = 0.27. A sensitivity analysis restricted to active-control trials (EUC, FE-NST, TAU) yielded similar results, g_{pooled} = 0.23, 95% CI [-0.43, 0.88], I^2 = 2.8%, Cochran's Q p = 0.36.

Discussion

This study is the first to systematically review the efficacy of Attachment-Based Family Therapy (ABFT), a psychotherapeutic treatment for pediatric patients focussing on intra-familial attachment dynamics, for treating youth who are suicidal. Out of 370 records, four randomized-controlled trials (n = 287) and three open trials (n = 45) measuring suicidality were identified. In the meta-analysis covering four RCTs, Attachment-Based Family Therapy was not superior in reducing suicidal ideation or depressive symptoms compared to investigated controls (waitlist, FE-NST, EUC, TAU), and these results remained stable when restricting the analysis to active control interventions only. Nonetheless, most studies (5 out of 7) reported significant reductions in suicidal ideation and depression within the ABFT treatment group over time.

Overall, experimental evidence is limited, with few available trials, small sample sizes, limited follow-up data, high sample heterogeneity, attrition rates, and risk of bias.

Currently, only four controlled trials including 287 participants exist, which might be insufficient to robustly establish or rule out superiority in a comparison with anticipated small effect sizes. Follow-up data was only available for one out of four RCTs, limiting the interpretability of clinical effects and their sustainability over time (G. S. Diamond et al., 2010). A priori defined recruitment targets were only met in one trial, indicating that most studies were underpowered (G. S. Diamond et al., 2019). Loss-to-follow-up was heterogeneous, with one study exhibiting a pre-post loss rate of ~80% for the primary outcome (Waraan et al., 2021). Attrition rates in the other randomized-controlled trials were moderate (3%–18%) and in line with estimates from other psychotherapy trials for depression in children and youth (Cooper & Conklin, 2015; Wright et al., 2021). SIQ(-JR) and BDI(-II) scales were used in most trials to assess suicidality and depressive symptoms, introducing some risk of instrument bias. Additionally, the SIQ(-JR) scale has recently been criticized for its insufficient psychometric properties (Courtney et al., 2024). Emerging evidence further suggests that severity of suicidal ideation strongly fluctuates short-term, for which multiple assessments might be necessary to measure state suicidality reliably (Czyz et al., 2019; Kleiman et al., 2017). There was also considerable design heterogeneity between trials. Two studies in the meta-analysis used the 12-week ABFT protocol, while others applied the treatment for 16 weeks. Comparator treatments differed strongly, with trials using waitlist control, (enhanced) Treatment as Usual (EUC, TAU), or Family-enhanced Nondirective Supportive Therapy (FE-NST), which interestingly has been found to be effective for adult depression (Cuijpers et al., 2012). Actual type of treatment delivered was not assessed and adequately described in EUC and TAU trials, which might contribute to unexplained heterogeneity (Burns, 2009; Witt et al., 2018). Further so, three out of four included RCTs were conducted by researchers associated with the treatment development group. Primary allegiance of care providing professionals was only reported in one RCT, potentially introducing allegiance bias. Unfortunately, extending the meta-analysis through subgroup analyses for heterogeneity factors was not feasible due to the limited number of included trials and lack of Individual Participant Data. As a result, it was not possible to assess the respective impact of these sources of bias on the potential underestimation or overestimation of the comparative effect.

Narrative reviews of previous trials suggested that Attachment-Based Family Therapy might be more effective in reducing suicidal ideation and depressive symptoms in youth than the current standard of care, which was not confirmed by this meta-analysis (G. S. Diamond et al., 2016, 2021; Ewing et al., 2015). Although restrained in validity due to the described limitations, these findings are in line with systematic reviews on the effect of other family-based interventions on depressive symptoms and a large meta-analysis on the effect of all-type treatments (e.g., medication, psychotherapy, combined) on several measures of suicidality (e.g., ideation, (non-)suicidal self-injury, death) in adolescents, which neither found significant treatment effects, regardless of outcome measure and

intervention type (Harris et al., 2022; van Aswegen et al., 2023; Waraan et al., 2023). Similar patterns can be observed with treatments for youth depression, which generally tend to yield lower effects than interventions for adults, and to which a substantial amount of patients do not respond within time (Cuijpers et al., 2020, 2023). For preventative approaches, data availability is currently strongly limited, and individual reports suggest that universal approaches might even lead to adverse effects on at-risk youth (Breedvelt et al., 2018; Montero-Marin et al., 2022). In contrast to these results, it should be noted that one other meta-analysis did identify a small positive effect of various family-based interventions on youth suicidal ideation, which upon closer inspection appears to be due to a discrepancy in data extraction for one specific trial (Waraan et al., 2023). As retrieving underlying data from this meta-analysis was not possible, we verified our extracted data with the author team of the trial in question to ensure correctness. Overall, effectiveness of current treatment options for youth suicidality is strongly limited and a major concern in pediatric healthcare.

Considering the high burden of disease implicated by continued suicidal ideation, attempts, and completed suicide in youth, developing more effective treatments is imperative. Previous discussions of this challenge suggested that integrating insights on the dynamics of youth, characterized by strong psychological, biological, and social volitions, might improve efficacy (Harris et al., 2022; Robinson et al., 2018). While Attachment-Based Family Therapy was, in contrast to more commonly applied psychological interventions (e.g., Cognitive Behavioral Therapy, Dialectical Behavior Therapy), developed explicitly for a pediatric population, it being non-superior to compared alternative treatments raises the question if other factors, such as patient baseline characteristics or therapeutic modalities, might influence treatment efficacy. Recent investigations on moderating factors in previous ABFT trials suggest, that patients with higher baseline levels of parent-teen conflict and underserved family backgrounds particularly benefit from treatment, and that more change in family cohesion during the intervention period was related to better treatment outcomes (Ibrahim et al., 2022; Zisk et al., 2019). Targeting Attachment-Based Family Therapy at high-yield subgroups, e.g., LGBTQI+ youth with non-accepting parents, could thus be beneficial (G. M. Diamond et al., 2022; Russon et al., 2023). Another area of improvement could be to intensify treatment components contributing to family cohesion, and disrupt negative feedback dynamics early on by e.g. implementing supervised exposure exercises to reduce patients' fear of the caregiver's emotional rejection (Bosmans et al., 2022). Other reviews pointed out that participants with higher baseline levels of depression, non-suicidal self-injury, perceived burdensomeness, and anxiety profited less from the treatment (Abbott et al., 2019; Herres et al., 2021). This might indicate that other interventions could have suited these patients better, or that sequential or multimodal approaches addressing their comorbidities and family dynamics simultaneously might have had better effects, e.g., combining Attachment-Based Family Therapy and Cognitive Behavioral Therapy for patients with anxiety (Herres

et al., 2023). Considering given evidence for moderating factors, it is thus conceivable that ABFT might potentially lead to better treatment outcomes compared to the current standard of care when applied to patients whose condition is strongly linked to dynamics of intrafamilial relations, and whose comorbid disorders are concurrently addressed – that is, when the right patient receives the right treatment at the right time.

Given the limitations of current evidence on the efficacy of Attachment-Based Family Therapy for treating youth who are suicidal, future research should focus on delivering high-quality evidence with more adequately powered samples, longer follow-up times, and more consistent and rigorous measurements. At the time of review, two adequately powered randomized-controlled clinical trials evaluating its performance and cost-effectiveness were registered at ClinicalTrials.gov (Bockting et al., 2023; Pachankis et al., 2023). Additionally, with the weight of evidence suggesting that ABFT yields varying degrees of benefit for different patient groups – on average not exceeding the effectiveness of alternative interventions such as TAU or FE-NST – a better understanding of underlying active ingredients and treatment-induced mechanisms of change might be helpful to dismantle contexts and conditions under which specific treatment components lead to better treatment outcomes (Cuijpers et al., 2019; Hofmann & Hayes, 2019). Previously, limited evidence for the role of therapeutic alliance as a common factor, and guided emotional processing, improved parenting, and reattachment as treatment-specific factors was found (G. S. Diamond et al., 2021). Future research could further focus on disentangling these specific factors from common factors also activated in alternative treatments such as TAU or FE-NST, understanding the underlying processes of change involved, identifying therapeutic components that activate these changes in a cost-efficient manner, and providing clinicians with informed guidelines on implementing these components within a process-based therapy framework (Hofmann & Hayes, 2019; Rief et al., 2024).

Finally, a key challenge is to establish predictive factors for when individual treatment components of Attachment-Based Family Therapy might be feasible for a patient. While tailoring therapy plans and manualized treatments to the needs of individual patients and their presenting condition is common practice among clinicians, the application of prediction and personalization models in research remains scarce (Meehan et al., 2022). Meanwhile, research indicates that such personalized medicine approaches could result in substantial benefits for patients, and it might therefore be valuable to investigate further 'what works for whom' when it comes to dynamicity-embracing treatments for suicidal youth (Deisenhofer et al., 2024; DeRubeis et al., 2014; Robinson et al., 2018). Ultimately, prioritizing investment into larger cohort studies and analyses of Individual Patient Data to develop and improve targeted preventative approaches in adolescence is crucial, as this special developmental stage provides a unique window of opportunity to stimulate long-lasting positive life trajectories (Breedvelt et al., 2018, 2024; Dahl et al., 2018).

Funding: CB is funded by the Centre for Urban Mental Health, University of Amsterdam, Amsterdam, The Netherlands. The funding source was not involved in any processes or decisions related to this study.

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: GB is head and founder of the KU Leuven Attachment-Based Family Therapy (ABFT) training centre. All other authors report no conflict of interest.

Preregistration: This systematic review was preregistered at the International Prospective Register of Systematic Reviews, PROSPERO: CRD42021271731.

Reporting Guidelines: This systematic-review and meta-analysis was conducted in accordance with the PRISMA guidelines. The Cochrane Rob-2 framework was applied to assess the risk of bias in randomized-controlled trials.

Author Contributions: CB is the guarantor. All authors contributed to the development of the research question and review design. PS performed the search, review, data extraction, analysis, and drafted the manuscript. MB provided expertise on systematic reviews, contributed to the review of articles, and verified extracted data. JB provided expertise on meta-analyses, verified extracted data and analysis, and contributed to the manuscript. NS and GB provided expertise on the intervention. All authors contributed by reading, providing feedback, and approving the final manuscript.

Social Media Accounts: @poschfeld, @JosefienUMH, @clbockting, @mebrouwer

Abbreviations:

ABFT – Attachment-Based Family Therapy
CBT – Cognitive Behavioral Therapy
DBT – Dialectical Behavior Therapy
EUC – Enhanced Usual Care
FE-NST – Family-Enhanced Nondirective Supportive Therapy
TAU – Treatment as Usual

Data Availability: All data and calculation scripts for this study are publicly available (see Schulte-Frankenfeld et al., 2024S-b).

Supplementary Materials

The Supplementary Materials contain the following items:

- The preregistration for the study (Schulte-Frankenfeld et al., 2024S-a).
- All collected summary statistics per study, analysis scripts, and a list of all articles screened (Schulte-Frankenfeld et al., 2024S-b)
- Online appendices: Additional information on the meta-analysis, namely detailed risk-of-bias scores per study and category, weighted overall risk-of-bias scores, and a color-enhanced funnel

plot on the distribution of effect sizes of included studies [Supplements 1] (Schulte-Frankenfeld et al., 2024S-c).

Index of Supplementary Materials

Schulte-Frankenfeld, P. M., Breedvelt, J. J. F., Brouwer, M. E., van der Spek, N., Bosmans, G., & Bockting, C. L. (2024S-a). *The effectiveness of attachment-based family therapy for suicidal adolescents and young adults: A systematic review and meta-analysis* [Preregistration]. PROSPERO. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021271731

Schulte-Frankenfeld, P. M., Breedvelt, J. J. F., Brouwer, M. E., van der Spek, N., Bosmans, G., & Bockting, C. L. (2024S-b). *Attachment-based family therapy for suicidal adolescents and young adults: A systematic review and meta-analysis* [Summary statistics, analysis scripts, list of screened articles]. OSF. https://doi.org/10.17605/OSF.IO/2SWE8

Schulte-Frankenfeld, P. M., Breedvelt, J. J. F., Brouwer, M. E., van der Spek, N., Bosmans, G., & Bockting, C. L. (2024S-c). *Supplementary materials to "Effectiveness of attachment-based family therapy for suicidal adolescents and young adults: A systematic review and meta-analysis"* [Online appendices]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.15566

References

Abbott, C. H., Zisk, A., Bounoua, N., Diamond, G. S., & Kobak, R. (2019). Predicting patterns of treatment response and outcome for adolescents who are suicidal and depressed. *Journal of the American Academy of Child and Adolescent Psychiatry, 58*(9), 897–906. https://doi.org/10.1016/j.jaac.2018.12.013

Bachmann, S. (2018). Epidemiology of suicide and the psychiatric perspective. *International Journal of Environmental Research and Public Health, 15*(7), Article 1425. https://doi.org/10.3390/ijerph15071425

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health, 22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Beck, A. T. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*(6), Article 561. https://doi.org/10.1001/archpsyc.1961.01710120031004

Beck, A. T., Steer, R. A., & Brown, G. (1996). *Beck Depression Inventory–II* [Database record]. APA PsycTests. https://doi.org/10.1037/t00742-000

Bockting, C., van der Spek, N., & Bosmans, G. (2023). *(Cost)-effectiveness of ABFT for suicidal youth (REPAIR).* ClinicalTrials.Gov. https://clinicaltrials.gov/study/NCT05965622

Bosmans, G., van Vlierberghe, L., Bakermans-Kranenburg, M. J., Kobak, R., Hermans, D., & van IJzendoorn, M. H. (2022). A learning theory approach to attachment theory: Exploring clinical applications. *Clinical Child and Family Psychology Review, 25*(3), 591–612. https://doi.org/10.1007/s10567-021-00377-x

Breedvelt, J. J. F., Kandola, A., Kousoulis, A. A., Brouwer, M. E., Karyotaki, E., Bockting, C. L. H., & Cuijpers, P. (2018). What are the effects of preventative interventions on major depressive disorder (MDD) in young adults? A systematic review and meta-analysis of randomized controlled trials. *Journal of Affective Disorders, 239*, 18–29. https://doi.org/10.1016/j.jad.2018.05.010

Breedvelt, J. J. F., Karyotaki, E., Warren, F. C., Brouwer, M. E., Jermann, F., Hollandare, F., Klein, N., de Jonge, M., Klein, D. N., Farb, N., Segal, Z., Biesheuvel Leliefeld, K. E. M., Jarrett, R., Vittengl, J., Thase, M., Ma, H., Kuyken, W., Shallcross, A. J., Van Heeringen, C., . . . Bockting, C. L. (2024). An individual participant data meta-analysis of psychological interventions for preventing depression relapse. *Nature Mental Health, 2*, 154–163. https://doi.org/10.1038/s44220-023-00178-x

Burns, T. (2009). End of the road for treatment-as-usual studies? *The British Journal of Psychiatry, 195*(1), 5–6. https://doi.org/10.1192/bjp.bp.108.062968

Busby, D. R., Hatkevich, C., McGuire, T. C., & King, C. A. (2020). Evidence-based interventions for youth suicide risk. *Current Psychiatry Reports, 22*(2), Article 5. https://doi.org/10.1007/s11920-020-1129-6

California Evidence-Based Clearinghouse for Child Welfare. (2020). *Attachment-Based Family Therapy*. https://www.cebc4cw.org/program/attachment-based-family-therapy/

Carballo, J. J., Llorente, C., Kehrmann, L., Flamarique, I., Zuddas, A., Purper-Ouakil, D., Hoekstra, P. J., Coghill, D., Schulze, U. M. E., Dittmann, R. W., Buitelaar, J. K., Castro-Fornieles, J., Lievesley, K., Santosh, P., Arango, C., & the STOP Consortium. (2020). Psychosocial risk factors for suicidality in children and adolescents. *European Child & Adolescent Psychiatry, 29*(6), 759–776. https://doi.org/10.1007/s00787-018-01270-9

Castelpietra, G., Knudsen, A. K. S., Agardh, E. E., Armocida, B., Beghi, M., Iburg, K. M., Logroscino, G., Ma, R., Starace, F., Steel, N., Addolorato, G., Andrei, C. L., Andrei, T., Ayuso-Mateos, J. L., Banach, M., Bärnighausen, T. W., Barone-Adesi, F., Bhagavathula, A. S., Carvalho, F., . . . Monasta, L. (2022). The burden of mental disorders, substance use disorders and self-harm among young people in Europe, 1990–2019: Findings from the Global Burden of Disease Study 2019. *The Lancet Regional Health – Europe, 16*, Article 100341. https://doi.org/10.1016/j.lanepe.2022.100341

Chiang, A., Paynter, J., Edlin, R., & Exeter, D. J. (2021). Suicide preceded by health services contact – A whole-of-population study in New Zealand 2013-2015. *PLoS One, 16*(12), Article e0261163. https://doi.org/10.1371/journal.pone.0261163

Chu, C., Buchman-Schmitt, J. M., Stanley, I. H., Hom, M. A., Tucker, R. P., Hagan, C. R., Rogers, M. L., Podlogar, M. C., Chiurliza, B., Ringer, F. B., Michaels, M. S., Patros, C. H. G., & Joiner, T. E. (2017). The interpersonal theory of suicide: A systematic review and meta-analysis of a decade of cross-national research. *Psychological Bulletin, 143*(12), 1313–1345. https://doi.org/10.1037/bul0000123

Cooper, A. A., & Conklin, L. R. (2015). Dropout from individual psychotherapy for major depression: A meta-analysis of randomized clinical trials. *Clinical Psychology Review, 40*, 57–65. https://doi.org/10.1016/j.cpr.2015.05.001

Courtney, D. B., Iseyas, N., Monga, S., Butcher, N. J., Krause, K., Besa, R., & Szatmari, P. (2024). Systematic review: The measurement properties of the Suicidal Ideation Questionnaire and Suicidal Ideation Questionnaire—Jr. *Journal of the American Academy of Child and Adolescent Psychiatry, 63*(9), 870–887. https://doi.org/10.1016/j.jaac.2023.10.013

Cuijpers, P., Driessen, E., Hollon, S. D., van Oppen, P., Barth, J., & Andersson, G. (2012). The efficacy of non-directive supportive therapy for adult depression: A meta-analysis. *Clinical Psychology Review, 32*(4), 280–291. https://doi.org/10.1016/j.cpr.2012.01.003

Cuijpers, P., Griffin, J. W., & Furukawa, T. A. (2021). The lack of statistical power of subgroup analyses in meta-analyses: A cautionary note. *Epidemiology and Psychiatric Sciences, 30*, Article e78. https://doi.org/10.1017/S2045796021000664

Cuijpers, P., Karyotaki, E., Ciharova, M., Miguel, C., Noma, H., Stikkelbroek, Y., Weisz, J. R., & Furukawa, T. A. (2023). The effects of psychological treatments of depression in children and adolescents on response, reliable change, and deterioration: A systematic review and meta-analysis. *European Child & Adolescent Psychiatry, 32*(1), 177–192. https://doi.org/10.1007/s00787-021-01884-6

Cuijpers, P., Karyotaki, E., Eckshtain, D., Ng, M. Y., Corteselli, K. A., Noma, H., Quero, S., & Weisz, J. R. (2020). Psychotherapy for depression across different age groups: A systematic review and meta-analysis. *JAMA Psychiatry, 77*(7), 694–702. https://doi.org/10.1001/jamapsychiatry.2020.0164

Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The role of common factors in psychotherapy outcomes. *Annual Review of Clinical Psychology, 15*(1), 207–231. https://doi.org/10.1146/annurev-clinpsy-050718-095424

Cuijpers, P., Turner, E. H., Koole, S. L., van Dijke, A., & Smit, F. (2014). What is the threshold for a clinically relevant effect? The case of major depressive disorders. *Depression and Anxiety, 31*(5), 374–378. https://doi.org/10.1002/da.22249

Czyz, E. K., Horwitz, A. G., Arango, A., & King, C. A. (2019). Short-term change and prediction of suicidal ideation among adolescents: A daily diary study following psychiatric hospitalization. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 60*(7), 732–741. https://doi.org/10.1111/jcpp.12974

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature, 554*(7693), 441–450. https://doi.org/10.1038/nature25770

Deisenhofer, A.-K., Barkham, M., Beierl, E. T., Schwartz, B., Aafjes-van Doorn, K., Beevers, C. G., Berwian, I. M., Blackwell, S. E., Bockting, C. L., Brakemeier, E.-L., Brown, G., Buckman, J. E. J., Castonguay, L. G., Cusack, C. E., Dalgleish, T., de Jong, K., Delgadillo, J., DeRubeis, R. J., Driessen, E., . . . Cohen, Z. D. (2024). Implementing precision methods in personalizing

psychological therapies: Barriers and possible ways forward. *Behaviour Research and Therapy, 172*, Article 104443. https://doi.org/10.1016/j.brat.2023.104443

DeRubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014). The Personalized Advantage Index: Translating research on prediction into individualized treatment recommendations: A demonstration. *PLoS One, 9*(1), Article e83875. https://doi.org/10.1371/journal.pone.0083875

Diamond, G. M., Boruchovitz-Zamir, R., Nir-Gotlieb, O., Gat, I., Bar-Kalifa, E., Fitoussi, P., & Katz, S. (2022). Attachment-based family therapy for sexual and gender minority young adults and their nonaccepting parents. *Family Process, 61*(2), 530–548. https://doi.org/10.1111/famp.12770

Diamond, G. M., Diamond, G. S., Levy, S., Closs, C., Ladipo, T., & Siqueland, L. (2012). Attachment-Based Family Therapy for suicidal lesbian, gay, and bisexual adolescents: A treatment development study and open trial with preliminary findings. *Psychotherapy, 49*(1), 62–71. https://doi.org/10.1037/a0026247

Diamond, G. S. (2022). Attachment-Based Family Therapy for youth at risk for suicide: The next evolution. *Journal of the American Academy of Child and Adolescent Psychiatry, 61*(10), S32–S33. https://doi.org/10.1016/j.jaac.2022.07.133

Diamond, G. S., Creed, T., Gillham, J., Gallop, R., & Hamilton, J. L. (2012). Sexual trauma history does not moderate treatment outcome in Attachment-Based Family Therapy (ABFT) for adolescents with suicide ideation. *Journal of Family Psychology, 26*(4), 595–605. https://doi.org/10.1037/a0028414

Diamond, G. S., Diamond, G. M., & Levy, S. A. (2014). *Attachment-based family therapy for depressed adolescents*. American Psychological Association. https://doi.org/10.1037/14296-000

Diamond, G. S., Diamond, G. M., & Levy, S. A. (2021). Attachment-based family therapy: Theory, clinical model, outcomes, and process research. *Journal of Affective Disorders, 294*(1), 286–295. https://doi.org/10.1016/j.jad.2021.07.005

Diamond, G. S., Kobak, R. R., Krauthamer Ewing, E. S., Levy, S. A., Herres, J. L., Russon, J. M., & Gallop, R. J. (2019). A randomized controlled trial: Attachment-based family and nondirective supportive treatments for youth who are suicidal. *Journal of the American Academy of Child and Adolescent Psychiatry, 58*(7), 721–731. https://doi.org/10.1016/j.jaac.2018.10.006

Diamond, G. S., Reis, B. F., Diamond, G. M., Siqueland, L., & Isaacs, L. (2002). Attachment-Based Family Therapy for depressed adolescents: A treatment development study. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(10), 1190–1196. https://doi.org/10.1097/00004583-200210000-00008

Diamond, G. S., Russon, J., & Levy, S. (2016). Attachment-Based Family Therapy: A review of the empirical support. *Family Process, 55*(3), 595–610. https://doi.org/10.1111/famp.12241

Diamond, G. S., Wintersteen, M. B., Brown, G. K., Diamond, G. M., Gallop, R., Shelef, K., & Levy, S. (2010). Attachment-Based Family Therapy for adolescents with suicidal ideation: A randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(2), 122–131. https://doi.org/10.1097/00004583-201002000-00006

Ewing, E. S. K., Diamond, G. S., & Levy, S. (2015). Attachment-based family therapy for depressed and suicidal adolescents: Theory, clinical model and empirical support. *Attachment & Human Development, 17*(2), 136–156. https://doi.org/10.1080/14616734.2015.1006384

Feder, M. M., & Diamond, G. M. (2016). Parent-therapist alliance and parent attachment-promoting behaviour in attachment-based family therapy for suicidal and depressed adolescents. *Journal of Family Therapy, 38*(1), 82–101. https://doi.org/10.1111/1467-6427.12078

Fergusson, D. M., Horwood, L. J., Ridder, E. M., & Beautrais, A. L. (2005). Suicidal behaviour in adolescence and subsequent mental health outcomes in young adulthood. *Psychological Medicine, 35*(7), 983–993. https://doi.org/10.1017/S0033291704004167

Fox, K. R., Huang, X., Guzmán, E. M., Funsch, K. M., Cha, C. B., Ribeiro, J. D., & Franklin, J. C. (2020). Interventions for suicide and self-injury: A meta-analysis of randomized controlled trials across nearly 50 years of research. *Psychological Bulletin, 146*(12), 1117–1145. https://doi.org/10.1037/bul0000305

Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., Musacchio, K. M., Jaroszewski, A. C., Chang, B. P., & Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin, 143*(2), 187–232. https://doi.org/10.1037/bul0000084

Franzoi, D., Bockting, C. L., Bennett, K. F., Odom, A., Lucassen, P. J., Pathania, A., Lee, A., Brouwer, M. E., Van De Schoot, R., Wiers, R. W., & Breedvelt, J. J. F. (2024). Which individual, social, and urban factors in early childhood predict psychopathology in later childhood, adolescence and young adulthood? A systematic review. *SSM – Population Health, 25*, Article 101575. https://doi.org/10.1016/j.ssmph.2023.101575

Frey, L. M., Hunt, Q. A., Russon, J. M., & Diamond, G. (2022). Review of family-based treatments from 2010 to 2019 for suicidal ideation and behavior. *Journal of Marital and Family Therapy, 48*(1), 154–177. https://doi.org/10.1111/jmft.12568

Gunn, J. F., Goldstein, S. E., & Gager, C. T. (2018). A longitudinal examination of social connectedness and suicidal thoughts and behaviors among adolescents. *Child and Adolescent Mental Health, 23*(4), 341–350. https://doi.org/10.1111/camh.12281

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide*. Chapman & Hall/CRC Press.

Harris, L. M., Huang, X., Funsch, K. M., Fox, K. R., & Ribeiro, J. D. (2022). Efficacy of interventions for suicide and self-injury in children and adolescents: A meta-analysis. *Scientific Reports, 12*(1), Article 12313. https://doi.org/10.1038/s41598-022-16567-8

Herres, J., James, K. M., Bounoua, N., Krauthamer Ewing, E. S., Kobak, R., & Diamond, G. S. (2021). Anxiety-related difficulties in goal-directed behavior predict worse treatment outcome among adolescents treated for suicidal ideation and depressive symptoms. *Psychotherapy, 58*(4), 523–532. https://doi.org/10.1037/pst0000391

Herres, J., Krauthamer Ewing, E. S., Levy, S., Creed, T. A., & Diamond, G. S. (2023). Combining attachment-based family therapy and cognitive behavioral therapy to improve outcomes for

adolescents with anxiety. *Frontiers in Psychiatry, 14*, Article 1096291. https://doi.org/10.3389/fpsyt.2023.1096291

Higgins, J., Li, T., & Deeks, J. (2023). Chapter 6: Choosing effect measures and computing estimates of effect. In *Cochrane handbook for systematic reviews of interventions* (Version 6.4; updated August 2023). Cochrane.

Hofmann, S. G., & Hayes, S. C. (2019). The future of intervention science: Process-based therapy. *Clinical Psychological Science, 7*(1), 37–50. https://doi.org/10.1177/2167702618772296

Holmes, E. A., Ghaderi, A., Harmer, C. J., Ramchandani, P. G., Cuijpers, P., Morrison, A. P., Roiser, J. P., Bockting, C. L. H., O'Connor, R. C., Shafran, R., Moulds, M. L., & Craske, M. G. (2018). The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *The Lancet Psychiatry, 5*(3), 237–286. https://doi.org/10.1016/S2215-0366(17)30513-8

Ibrahim, M., Jin, B., Russon, J., Diamond, G. S., & Kobak, R. (2018). Predicting alliance for depressed and suicidal adolescents: The role of perceived attachment to mothers. *Evidence-Based Practice in Child and Adolescent Mental Health, 3*(1), 42–56. https://doi.org/10.1080/23794925.2018.1423893

Ibrahim, M., King, A., Levy, S., Russon, J., & Diamond, G. (2022). Increased family cohesion mediates therapist adherence to the attachment task and depression outcomes in attachment-based family therapy. *Journal of Contemporary Psychotherapy, 52*(4), 303–310. https://doi.org/10.1007/s10879-022-09539-6

IntHout, J., Ioannidis, J. P. A., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open, 6*(7), Article e010247. https://doi.org/10.1136/bmjopen-2015-010247

Israel, P., & Diamond, G. S. (2013). Feasibility of Attachment-Based Family Therapy for depressed clinic-referred Norwegian adolescents. *Clinical Child Psychology and Psychiatry, 18*(3), 334–350. https://doi.org/10.1177/1359104512455811

Jackman, K. B., Caceres, B. A., Kreuze, E. J., & Bockting, W. O. (2021). Suicidality among gender minority youth: Analysis of 2017 Youth Risk Behavior Survey data. *Archives of Suicide Research, 25*(2), 208–223. https://doi.org/10.1080/13811118.2019.1678539

King, C. A., & Merchant, C. R. (2008). Social and interpersonal factors relating to adolescent suicidality: A review of the literature. *Archives of Suicide Research, 12*(3), 181–196. https://doi.org/10.1080/13811110802101203

Kleiman, E. M., Turner, B. J., Fedor, S., Beale, E. E., Huffman, J. C., & Nock, M. K. (2017). Examination of real-time fluctuations in suicidal ideation and its risk factors: Results from two ecological momentary assessment studies. *Journal of Abnormal Psychology, 126*(6), 726–738. https://doi.org/10.1037/abn0000273

Kothgassner, O. D., Robinson, K., Goreis, A., Ougrin, D., & Plener, P. L. (2020). Does treatment method matter? A meta-analysis of the past 20 years of research on therapeutic interventions for self-harm and suicidal ideation in adolescents. *Borderline Personality Disorder and Emotion Dysregulation, 7*(1), Article 9. https://doi.org/10.1186/s40479-020-00123-9

Liu, R. T., Walsh, R. F. L., Sheehan, A. E., Cheek, S. M., & Carter, S. M. (2020). Suicidal ideation and behavior among sexual minority and heterosexual youth: 1995–2017. *Pediatrics, 145*(3), Article e20192221. https://doi.org/10.1542/peds.2019-2221

Lüdecke, D. (2019). esc: Effect Size Computation for Meta Analysis [R package] [Computer software]. https://doi.org/10.5281/zenodo.1249218

McClelland, H., Evans, J. J., Nowland, R., Ferguson, E., & O'Connor, R. C. (2020). Loneliness as a predictor of suicidal ideation and behaviour: A systematic review and meta-analysis of prospective studies. *Journal of Affective Disorders, 274*, 880–896. https://doi.org/10.1016/j.jad.2020.05.004

McGuinness, L. A. (2019). *robvis: An R package and web application for visualising risk-of-bias assessments* [Computer software]. https://github.com/mcguinlu/robvis

Meehan, A. J., Lewis, S. J., Fazel, S., Fusar-Poli, P., Steyerberg, E. W., Stahl, D., & Danese, A. (2022). Clinical prediction models in psychiatry: A systematic review of two decades of progress and challenges. *Molecular Psychiatry, 27*(6), 2700–2708. https://doi.org/10.1038/s41380-022-01528-4

Montero-Marin, J., Allwood, M., Ball, S., Crane, C., De Wilde, K., Hinze, V., Jones, B., Lord, L., Nuthall, E., Raja, A., Taylor, L., Tudor, K., MYRIAD Team, Blakemore, S.-J., Byford, S., Dalgleish, T., Ford, T., Greenberg, M. T., Ukoumunne, O. C., … Kuyken, W. (2022). School-based mindfulness training in early adolescence: What works, for whom and how in the MYRIAD trial? *Evidence-Based Mental Health, 25*(3), 117–124. https://doi.org/10.1136/ebmental-2022-300439

National Institute for Health and Care Excellence. (2019). *Depression in children and young people: Identification and management (NICE Guideline).*

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews, 5*(1), Article 210. https://doi.org/10.1186/s13643-016-0384-4

Pachankis, J. E., Chiaramonte, D. M., & Diamond, G. M. (2023). *Affirmative family and individual psychotherapy for sexual and gender minority adults and their nonaccepting parents.* ClinicalTrials.Gov. https://clinicaltrials.gov/study/NCT05766592

Reynolds, W. M. (1987). *Suicidal Ideation Questionnaire (SIQ).* Psychological Assessment Resources.

Reynolds, W. M., & Mazza, J. J. (1999). Assessment of suicidal ideation in inner-city children and young adolescents: Reliability and validity of the Suicidal Ideation Questionnaire-JR. *School Psychology Review, 28*(1), 17–30. https://doi.org/10.1080/02796015.1999.12085945

Rief, W., Asmundson, G. J. G., Bryant, R. A., Clark, D. M., Ehlers, A., Holmes, E. A., McNally, R. J., Neufeld, C. B., Wilhelm, S., Jaroszewski, A. C., Berg, M., Haberkamp, A., & Hofmann, S. G. (2024). The future of psychological treatments: The Marburg Declaration. *Clinical Psychology Review, 110*, Article 102417. https://doi.org/10.1016/j.cpr.2024.102417

Robinson, J., Bailey, E., Witt, K., Stefanac, N., Milner, A., Currier, D., Pirkis, J., Condron, P., & Hetrick, S. (2018). What works in youth suicide prevention? A systematic review and meta-analysis. *EClinicalMedicine, 4–5*, 52–91. https://doi.org/10.1016/j.eclinm.2018.10.004

RStudio Team. (2021). *RStudio: Integrated Development Environment for R* [Computer software]. RStudio, PBC. http://www.rstudio.com/

Russon, J., Abbott, C. H., Jin, B., Rivers, A. S., Winston-Lindeboom, P., Kobak, R., & Diamond, G. S. (2023). Attachment-based family therapy versus nondirective supportive therapy for lesbian, gay, bisexual and questioning adolescents with depression, and suicidal ideation: An exploratory study. *Suicide & Life-Threatening Behavior, 53*(6), 958–967. https://doi.org/10.1111/sltb.12995

Russon, J., Morrissey, J., Dellinger, J., Jin, B., & Diamond, G. (2022). Implementing Attachment-Based Family Therapy for depressed and suicidal adolescents and young adults in LGBTQ+ Services: Feasibility, acceptability, and preliminary effectiveness. *Crisis: The Journal of Crisis Intervention and Suicide Prevention, 43*(6), 500–507. https://doi.org/10.1027/0227-5910/a000821

Shpigel, M. S., Diamond, G. M., & Diamond, G. S. (2012). Changes in parenting behaviors, attachment, depressive symptoms, and suicidal ideation in attachment-based family therapy for depressive and suicidal adolescents. *Journal of Marital and Family Therapy, 38*(s1), 271–283. https://doi.org/10.1111/j.1752-0606.2012.00295.x

Sullivan, S. R., Spears, A. P., Mitchell, E. L., Walsh, S., Love, C., & Goodman, M. (2023). Family treatments for individuals at risk for suicide. *Crisis: The Journal of Crisis Intervention and Suicide Prevention, 44*(1), 49–60. https://doi.org/10.1027/0227-5910/a000828

Tarrier, N., Taylor, K., & Gooding, P. (2008). Cognitive-behavioral interventions to reduce suicide behavior: A systematic review and meta-analysis. *Behavior Modification, 32*(1), 77–108. https://doi.org/10.1177/0145445507304728

van Aert, R. C. M., & Jackson, D. (2019). A new justification of the Hartung-Knapp method for random-effects meta-analysis based on weighted least squares regression. *Research Synthesis Methods, 10*(4), 515–527. https://doi.org/10.1002/jrsm.1356

van Aswegen, T., Samartzi, E., Morris, L., van der Spek, N., de Vries, R., Seedat, S., & van Straten, A. (2023). Effectiveness of family-based therapy for depressive symptoms in children and adolescents: A systematic review and meta-analysis. *International Journal of Psychology, 58*(6), 499–511. https://doi.org/10.1002/ijop.12926

van der Spek, N., Dekker, W., Peen, J., Santens, T., Cuijpers, P., Bosmans, G., & Dekker, J. (2024). Attachment-Based Family Therapy for adolescents and young adults with suicide ideation and depression: An open trial. *Crisis: The Journal of Crisis Intervention and Suicide Prevention, 45*(1), 48–56. https://doi.org/10.1027/0227-5910/a000916

Van Orden, K. A., Witte, T. K., Cukrowicz, K. C., Braithwaite, S. R., Selby, E. A., & Joiner, T. E. (2010). The interpersonal theory of suicide. *Psychological Review, 117*(2), 575–600. https://doi.org/10.1037/a0018697

Walby, F. A., Myhre, M. Ø., & Kildahl, A. T. (2018). Contact with mental health services prior to suicide: A systematic review and meta-analysis. *Psychiatric Services, 69*(7), 751–759. https://doi.org/10.1176/appi.ps.201700475

Waraan, L., Rognli, E. W., Czajkowski, N. O., Mehlum, L., & Aalberg, M. (2021). Efficacy of attachment-based family therapy compared to treatment as usual for suicidal ideation in

adolescents with MDD. *Clinical Child Psychology and Psychiatry, 26*(2), 464–474. https://doi.org/10.1177/1359104520980776

Waraan, L., Siqveland, J., Hanssen-Bauer, K., Czjakowski, N. O., Axelsdóttir, B., Mehlum, L., & Aalberg, M. (2023). Family therapy for adolescents with depression and suicidal ideation: A systematic review and meta–analysis. *Clinical Child Psychology and Psychiatry, 28*(2), 831–849. https://doi.org/10.1177/13591045221125005

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), Article 1686. https://doi.org/10.21105/joss.01686

Witt, K., De Moraes, D. P., Salisbury, T. T., Arensman, E., Gunnell, D., Hazell, P., Townsend, E., van Heeringen, K., & Hawton, K. (2018). Treatment as usual (TAU) as a control condition in trials of cognitive behavioural-based psychotherapy for self-harm: Impact of content and quality on outcomes in a systematic review. *Journal of Affective Disorders, 235*, 434–447. https://doi.org/10.1016/j.jad.2018.04.025

World Health Organization. (2020). *Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019* [Data set]. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death

World Health Organization. (2021). *Suicide worldwide in 2019: Global health estimates.*

Wright, I., Mughal, F., Bowers, G., & Meiser-Stedman, R. (2021). Dropout from randomised controlled trials of psychological treatments for depression in children and youth: A systematic review and meta-analyses. *Journal of Affective Disorders, 281*, 880–890. https://doi.org/10.1016/j.jad.2020.11.039

Zeraatkar, D., Han, M., Ge, L., Hanna, S., & Guyatt, G. (2020). A comparison of the Hartung-Knapp-Sidik-Jonkman method for meta-analysis with conventional frequentist methods: A systematic review of simulation and empirical studies. *Cochrane Database of Systematic Reviews, Abstracts of the 26th Cochrane Colloquium, Santiago, Chile.* https://doi.org/10.1002/14651858.CD201901

Zisk, A., Abbott, C. H., Bounoua, N., Diamond, G. S., & Kobak, R. (2019). Parent–teen communication predicts treatment benefit for depressed and suicidal adolescents. *Journal of Consulting and Clinical Psychology, 87*(12), 1137–1148. https://doi.org/10.1037/ccp0000457

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

The Relation Between Social Anxiety and Perceptions of Likeability and Friendship in Adolescents

Jeanine M. D. Baartmans[1,2,3], Bonny F. J. A. van Steensel[4], J. Loes Pouwels[5], Tessa A. M. Lansu[5], Reinout W. H. J. Wiers[2], Susan M. Bögels[4], Anke M. Klein[2,6]

[1] *UvA Minds: Academic Treatment Centre, Amsterdam, The Netherlands.* **[2]** *Developmental Psychology, University of Amsterdam, Amsterdam, The Netherlands.* **[3]** *Education and Child Studies, Leiden University, Leiden, The Netherlands.* **[4]** *Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands.* **[5]** *Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands.* **[6]** *Developmental and Educational Psychology, Leiden University, Leiden, The Netherlands.*

Clinical Psychology in Europe, 2024, Vol. 6(4), Article e10705, https://doi.org/10.32872/cpe.10705

Received: 2022-11-20 • **Accepted:** 2024-07-02 • **Published (VoR):** 2024-12-20

Handling Editor: Cornelia Weise, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Corresponding Author: Anke M. Klein, Wassenaarseweg 53, 2333 AK Leiden, The Netherlands. Phone number: 0031 71 527 6673. E-mail: a.m.klein@fsw.leidenuniv.nl

Abstract

Background: This study investigated how different social anxiety symptoms (i.e., worrying about negative evaluation versus avoidance tendencies) in adolescents are related to the perception accuracy of likeability by peers and friendships with peers.

Method: A community sample of 263 adolescents between 12 and 15 years old reported on their social anxiety symptoms. In addition, they estimated how much their peers liked them, indicated how much they liked their peers, and who their friends were in their classroom.

Results: Results showed that socially anxious adolescents who mainly worried about negative evaluations, underestimated their likeability by peers. Adolescents with strong social avoidance tendencies had a more accurate perception of their likeability and friendships; they were less liked by their peers and had fewer friends.

Conclusion: The results emphasize the importance of treating avoidance behavior in social anxiety since avoidance tendencies may not only maintain the social anxiety symptoms but are also related to a more negative judgment by others.

Keywords

social anxiety, adolescents, worry, avoidance, likeability, cognitive bias, friendship

Highlights

- Adolescents who worry about their social evaluation tend to underestimate their own likeability.
- Adolescents who tend to avoid social situations are less liked and have fewer friends.
- It is important to be aware of different symptoms of social anxiety.

Background

The beginning of adolescence is characterized by an increased sensitivity to others' judgment. Adolescents find it important to be liked by their peers and to be accepted in the peer group and often fear being rejected or negatively evaluated. Therefore, fear of negative evaluation is part of normal development. Still, it can become problematic when these fears hinder adolescents in their daily lives and development (American Psychiatric Association, 2013). Several studies indeed found that a strong fear of negative evaluation peaks in early adolescence due to challenges in the peer context and social-cognitive development. This heightened fear of negative evaluation, one of the key cognitive symptoms of social anxiety, can amplify peer relationship stress (Chavira & Stein, 2005; Erath et al., 2008; Ranta et al., 2014).

However, fear of negative evaluation is only one of the symptoms associated with social anxiety, with other symptoms including attentional, emotional, behavioral, and physical symptoms also being present in social anxiety. In this study we focused on cognitive symptoms (fear of negative evaluation) and behavioral symptoms (avoidance). With regards to cognitive symptoms, adolescents with social anxiety (disorder) have strong assumptions about being judged or rejected by others (American Psychiatric Association, 2013; Bögels et al., 2010); they have fears of saying something 'wrong,' behaving 'inappropriate' or making a fool out of oneself in social situations. In addition, they expect that their social performances have disastrous consequences, and as a result, they are likely to avoid social situations (Leigh & Clark, 2018; Ranta et al., 2014). Thus, cognitive symptoms often lead to behavioral symptoms: i.e., avoidance of feared (social) situations and/or the use of safety behaviors to avoid the perceived anxious scenario from happening (e.g., extremely preparing and/or memorizing what to say prior to giving a presentation because the child believes this will prevent that other children will make fun of him). The avoidance of threatening stimuli results in a decrease in anxiety in the short term. However, in the long term, avoidance may maintain social anxiety by preventing habituation and by disconfirmation of biased perceptions (Miers et al., 2014; Rapee & Spence, 2004). For social anxiety specifically, social avoidance may be reflected in withdrawal from peer interactions. The current study aims to examine the associations

of (the subtypes of symptoms of) social anxiety with social functioning and (biased) perceptions of social functioning.

There is a growing body of evidence showing that adults with social anxiety (disorder) also suffer from actual deficits in social interactions (e.g., Voncken & Bögels, 2008). In addition, social withdrawal has been associated with social anxiety in children and can be defined as abstaining from social activities in the presence of peers (Erath et al., 2007). This withdrawn behavior prevents adolescents from practicing their social skills (Blöte et al., 2014; Clark & Wells, 1995). However, instead of, or in addition to actual lesser social performance, individuals high in social anxiety might also underestimate their social performance. Cognitive theories state that people with high levels of social anxiety or social anxiety disorder have a negatively biased perception of their social performance (e.g., Clark & Wells, 1995; Hofmann & DiBartolo, 2014; Rapee & Heimberg, 1997). As a result, it could be that the relation between social anxiety and worse social performance in self-report measures is affected by negative cognitive biases and might not be in line with others' opinions. Regarding youths, research shows considerable evidence for the relation between social anxiety, negative expectations of one's own social performance, and negative ratings of one's own social competence (Kingery et al., 2010).

In previous studies, we tested the association between social anxiety and the accuracy of being disliked by peers in children and adolescents (Baartmans et al., 2019, 2020; Klein et al., 2018). For example, Baartmans and colleagues (2019) examined the extent to which children between 7 and 13 years old had an accurate or biased perception of their general likeability among classroom peers. Results showed that in children, higher levels of social anxiety were associated with underestimating one's likeability among classroom peers. In line with this study, Baartmans and colleagues (2020) examined these questions in a different sample, using the same age group (7 to 13 years old) but investigated estimations of likeability among each peer individually instead of likeability among all classroom peers. Results again showed that when children had higher levels of social anxiety, they more strongly underestimated their own likeability. In addition, it was found that older children with social anxiety symptoms were more likely to underestimate their likeability by peers than younger children (Baartmans et al., 2020). Finally, Klein and colleagues (2018) followed a similar procedure to study adolescents aged between 12 and 19 with a mild intellectual disability. They also found that social anxiety symptoms were linked to a biased perception of likeability; adolescents with mild intellectual disability and with symptoms of social anxiety underestimated their likeability by peers (Klein et al., 2018). Previous research thus showed preliminary evidence for a relation between social anxiety and a negatively biased perception of likeability. In order to be able to compare the findings of the current study to previous findings, the current study examines a group of typically developing adolescents with an age range in between the studies mentioned above (10-15 years).

Whereas we already have some insight into how social anxiety is related to youth's (accuracy of the perception of) likeability among peers, we know less about their functioning and particularly accuracy regarding friendships. In addition to likeability, friendship plays a vital role in the lives of adolescents. Although both constructs are closely related (Cantin et al., 2019), they each tap into different components of social competence. Whereas likeability reflects a relatively general affective evaluation, friendships, conversely, are dyadic relationships requiring mutuality (Greco & Morris, 2005). Research on the relation between social anxiety and friendships in youth showed that social anxiety negatively influences companionship and intimacy in friendships (Vernberg et al., 1992) and that higher social anxiety is related to negative interactions in best friendships (La Greca & Harrison, 2005). Studies that focused on the number of friends about social anxiety symptoms suggested that especially girls with high levels of social anxiety report fewer friendships and less intimacy and support within close friendships (La Greca & Lopez, 1998). Socially anxious adolescents also indicate to have fewer friends, and they more often choose other socially anxious youth as their friends (Van Zalk et al., 2011; but see Karkavandi et al., 2022, who find that socially anxious girls nominate as many friends as socially non-anxious girls).

However, important to note is that most findings are based on self-report measures. This raises the question of to what extent these findings reflect actual friendship functioning or are driven by negative cognitive biases related to social anxiety. There are some indications that socially anxious youth may indeed be less likely to be selected as a friend (Karkavandi et al., 2022; Van Zalk et al., 2011) and that they have fewer reciprocated friendships (Erath et al., 2010). However, it could also be the case that adolescents do not always recognize it when peers consider them to be friends (an underestimation of friendship). Therefore, the second goal of this study was to investigate the association between social anxiety symptoms in adolescents and the accuracy of their estimations of friendships within the classroom.

The Present Study

The overall aim of this study was to examine the relations of social anxiety level with cognitive biases and social performance in the peer group among a large group of adolescents (10-15 years old). We examined this aim in the context of likeability and friendships within a classroom setting. In addition, we paid special attention to the subcategories of social anxiety symptoms; i.e., cognitive symptoms (fear of negative evaluation) and behavioral symptoms (avoidance).

To answer the first research question – how social anxiety symptoms in adolescents are related to self-perceived and peer-perceived likeability and the perception accuracy of likeability by peers - we first studied how both self-perceived and peer-perceived likeability were related to social anxiety in adolescents (Goal 1a). We expected that social anxiety would be negatively related to self- and peer-perceived likeability. Regarding the

subcategories of social anxiety, we expected that especially the behavioral symptoms (avoidance) would play an important role in actual likeability among peers because avoidance prevents adolescents from practicing their social skills, and adolescents with social skills deficits can be less likeable (Blöte et al., 2012, 2014; Clark & Wells, 1995; Miers et al., 2010, 2011). Moreover, when socially anxious adolescents avoid social interactions with peers, these peers are less likely to get to know them. They may, therefore, be less likely to develop a positive opinion about them. This is also in line with the findings of Henricks and colleagues (2021, 2023), who find a negative association between the behavioral social anxiety component 'avoidance' and likeability among peers, but not between the cognitive 'fear of negative evaluation' component and likeability. In Goal 1b, we focused on the discrepancy between the two reporters and its relation with social anxiety symptoms in adolescents. Based on previous research in children, we hypothesized that higher levels of (cognitive symptoms of) social anxiety in adolescents would be associated with stronger underestimations of likeability (La Greca & Lopez, 1998; Van Zalk et al., 2011).

The second research question focused on how social anxiety symptoms in adolescents relate to self- and peer-reported friendships and the perception accuracy of friendships within the classroom. To answer this question, we first focused on the association of social anxiety with the number of self-nominated, peer-nominated, and reciprocal friends (Goal 2a). We computed discrepancies between these three scores as indicators of underestimation and overestimation of the number of friendships. Subsequently, we examined how these discrepancies were related to social anxiety symptoms (Goal 2b). We expected that higher levels of social anxiety would be related to fewer self-reported friendships (La Greca & Lopez, 1998; Van Zalk et al., 2011), fewer peer-reported friendships, and fewer reciprocal friendships. We also expected that adolescents would underestimate the number of friendships (i.e., them 'being blind' to the friendship offered by others). With regards to the subcategories of social anxiety, we again expected that especially the behavioral symptoms (avoidance) would be associated with the number of self-reported and peer-reported friends. In contrast, the cognitive symptoms would be associated with underestimating the number of friendships.

Method

Participants and Procedure

A total of 263 adolescents (49.4% boys) between 10 and 15 years old from 12 classrooms in Grades 7 and 8 from two secondary schools (M = 13.64, SD = 0.65) participated in the study. In total, 168 adolescents (63.9%) attended the school level 'senior general secondary education / pre-university education,' and 95 adolescents (36.1%) attended the school level 'pre-vocational secondary education / senior general secondary education.'

The adolescents completed digital questionnaires in a classroom setting. Tables were set up in a 'test' setting so that participants could not see each other's answers. Also, the project coordinator stressed that there were no right or wrong answers and that the individual results would not be shared with anyone. The study was part of a larger study on social development and bullying (Henricks et al., 2021; Pouwels et al., 2018, 2019). The Ethical Committee of the Behavioural Science Institute of Radboud University Nijmegen, the Netherlands, approved this study.

Materials

Social Anxiety Symptoms

Social anxiety symptoms were measured with the shortened version of the *Social Anxiety Scale for Adolescents* (SAS-A; La Greca & Lopez, 1998; Kärnä et al., 2010). This questionnaire consists of nine items divided into two subscales, 'Fear of Negative Evaluation' (FNE) and 'Social Avoidance and Distress' (SAD; derived from the original SAD-general subscale of the SAS-A). Adolescents were asked to indicate to what extent the questions applied to them, ranging from 0 (never) to 4 (always). The shortened version of the SAS-A has a good internal consistency, $\alpha = .88$ (Kärnä et al., 2010). In our sample, the SAS-A also had a good internal consistency, $\alpha = .88$. The subscale FNE had an excellent internal consistency, $\alpha = .93$, and the subscale SAD-general had a good internal consistency, $\alpha = .84$. Both the total scale and the two subscales were included in the analyses.

Likeability

The adolescents were asked to answer on a Likert scale (1 = not at all, 7 = a lot), "How much do you think your classmates like you?". This score is indicated as *like-self*. All adolescents were also given a list of their classmates and asked to answer on a Likert scale (1 = not liked at all, 7 = very much liked): "How much do you like classmate X?". The average received score for each participating adolescent was computed and indicated as *like-peer*.

Discrepancy scores were used as a measure of perception bias. The discrepancy scores were computed by subtracting the *like-peer*-score from the *like-self*-score. This resulted in the *like-discr*-score. Positive values correspond with overestimating one's own likeability, and negative scores with underestimating one's likeability.

Friendship

First, adolescents were asked to name up to five best friends among their classmates. In a second question, adolescents were asked to name who were other good friends among their classmates. *Best-friend*-scores were derived from the first question and *total-friend*-scores were derived by computing the sum of the first and second question. *Given-friend*-scores were derived by computing the number of indicated best friend and

good friend nominations by each participant. *Received-friend*-scores were derived from the number of received nominations for each participant (i.e., how often a participant was indicated as a friend) from their classmates. Furthermore, for both the *best-friend*- and *total-friend*-scores the number of *reciprocal* nominations was determined by examining how often the participant and a classmate nominated each other as friends. This resulted in six friendship scores: *best-friend-received, best-friend-given, best-friend-reciprocal, total-friend-received, total-friend-given, total-friend-reciprocal.*

Over- and underestimation scores were computed to obtain information on the extent to which adolescents' self-indicated friendships were in line with the friendships indicated by their classmates. These two types of scores were used as indicators for perception accuracy. The first type of perception bias indicator on friendship was the *underestimate*-scores. *Underestimate*-scores were computed by subtracting the *reciprocal*-scores from the *received-friend*-scores ($N_{received} - N_{reciprocal}$). These scores indicate to what extent an adolescent recognizes the friendship-nominations that they receive. Higher *underestimate*-scores correspond with higher underestimation of the number of friends in the class. The second type of perception bias scores on friendship are the *overestimate*-scores. These were computed by subtracting the *reciprocal*-scores from the *given-friend*-scores ($N_{given} - N_{reciprocal}$). These scores indicate the extent to which peers indicate the same friendships as the participants themselves nominate. Higher *overestimate*-scores correspond with higher levels of overestimation of the number of friends in the class. The *overestimate*- and *underestimate*-scores were both computed for the number of best friends and the number of total friends. In total, this resulted in four friendship accuracy scores: overest-best-friend, underest-best-friend, overest-total-friend, and *underest-total-friend.*

Data Analysis

Before answering the research question, a Pearson correlation between the two subscales of the SAS-A (FNE and SAD) was computed. In addition, we tested if the *like-discr*-score deviated significantly from zero with a one-sample *t*-test to test if the group of adolescents overestimated or underestimated their likeability by peers on average.

To answer the first research question – how social anxiety symptoms in adolescents relate to the perception accuracy of likeability by peers – partial correlations were computed between social anxiety symptoms (SAS-A) and *like-self*, *like-peer*, and *like-discr*, while controlling for gender and age. We controlled for age and gender in the analyses since previous studies found evidence for the relation between age, gender, social anxiety symptoms, and biased perceptions (Baartmans et al., 2019, 2020). The partial correlation between the social anxiety symptoms and *like-self* and *like-peer* provides information about how the self-estimates and the opinions of peers about adolescents' likeability are related to social anxiety symptoms. The partial correlation between *like-discr* and social anxiety symptoms indicates if adolescents with higher levels of social anxiety are more likely to underestimate their likeability by peers.

To answer the second research question – how social anxiety symptoms in adolescents relate to the perception accuracy of their friendships within the classroom – we first tested how social anxiety symptoms were related to the self- and peer-nominations of friendships within the class. Therefore, partial correlations were computed between social anxiety symptoms and the *best-friend-received, best-friend-given, best-friend-reciprocal, total-friend-received, total-friend-given,* and *total-friend-reciprocal-scores* while controlling for gender and age. In the next step, partial correlations were computed between social anxiety symptoms and *overestimate-best-friend, underestimate-best-friend, overestimate-total-friend,* and *underestimate-total-friend* to test if social anxiety symptoms in adolescents were related to over- and/or underestimation of the number of friendships, again while controlling for age and gender.

In order to discriminate between sub-symptoms of social anxiety in their relation to perception accuracy of likeability and friendship – we conducted all analyses described above for the total social anxiety symptoms, as well as for the FNE- and SAD-scale separately.

Results

Descriptives

The Pearson correlation between the FNE and SAD subscales was .35 ($p < .001$). The *like-discr*-score had a mean of 0, indicating that adolescents had, on average, an accurate perception of their own likeability as rated by their peers. However, the standard deviation suggests that some adolescents underestimated or overestimated their likeability (see Table 1).

Research Question 1: Social Anxiety and Likeability

As expected, higher levels of social anxiety were significantly related to lower self-estimates of likeability (see Table 2). In addition, higher levels of total anxiety were also significantly related to being less likeable according to peers and an underestimation of likeability by peers. Adolescents with higher levels of social anxiety thus were less liked by their peers and also indicated themselves that their peers liked them less (Goal 1A). However, adolescents with higher social anxiety scores were, on average, too pessimistic about their low likeability, as indicated by a tendency to underestimate their actual likeability level (Goal 1B). As expected, both higher levels of the subscales Fear of Negative Evaluation and Social Avoidance and Distress were significantly related to lower self-estimates of likeability. In accordance with our hypothesis, only the avoidance subscale was significantly related to lower peer ratings of likeability. This resulted in a significant negative relation between the negative evaluation scale and the discrepancy score but not between the avoidance subscale and the discrepancy score. These results

suggest that a higher fear of negative evaluation was related to the underestimation of likeability by peers. In contrast, higher levels of social avoidance were related to both lower peer- and self-ratings of likeability, resulting in unbiased perceptions of one's likeability.

Table 1

Means (M) and Standard Deviations (SD) of the Social Anxiety Symptoms, Likeability Measures, and Friendship Measures

Variable	*M*	*SD*
SAS-total	1.46	0.67
SAS-SAD	1.78	0.85
SAS-FNE	1.07	0.75
Like-self	4.34	0.81
Like-peer	4.29	0.57
Like-discr	0.00	1.28
Best-friend-given	4.14	1.24
Total-friend-given	7.98	3.98
Best-friend-received	4.00	2.09
Total-friend-received	7.68	3.18
Best-friend-reciprocal	2.79	1.36
Total-friends-reciprocal	5.62	2.63
Overest-best-friend	1.35	1.36
Overest-total-friend	2.36	2.80
Underest-best-friend	1.21	1.31
Underest-total-friend	2.06	1.95

Note. SAS-total = total anxiety symptoms; SAS-SAD = social avoidance and distress symptoms; SAS-FNE = fear of negative evaluation symptoms; like-self = self-perceived likeability; like-peer = peer rated likeability; like-discr = difference between self- and peer-perceived likeability; best-friend-given = self-nominated number of best friends; total-friend-given = self-nominated number of best and good friends; best-friend-received = peer-indicated nominations of best friend; total-friend-received = peer-indicated nominations of total friend; best-friend-reciprocal = reciprocal number of best friends; total-friend-reciprocal = reciprocal number of best and good friends; overest-best-friend = overestimation of number of best friends; overest-total-friend = overestimation of number of best and good friends; underest-best-friend = underestimation of number of best friends; underest-total-friend = underestimation of number of best and good friends.

Table 2

Partial Correlations Between Social Anxiety Symptoms (Total, FNE, and SAD), the Likeability Measures, the Friendship Measures, and the Perception Bias Measures on Friendship While Controlling for Gender and Age

	Social anxiety symptoms		
Variable	**Total**	**FNE**	**SAD**
Likeability			
Like-self	-.31***	-.22***	-.32***
Like-peer	-.14*	-.04	-.23*
Like-discr	-.13*	-.14*	-.07
Friendship quantity			
Best-friend-given	-.19*	-.14	-.18*
Total-friend-given	-.11**	-.07	-.13*
Best-friend-received	-.16*	-.04	-.25***
Total-friend-received	-.10	.02	-.22***
Best-friend-reciprocal	-.17*	-.06	-.25***
Total-friends-reciprocal	-.17*	-.06	-.24***
Friendship accuracy			
Overest-best-friend	.00	-.07	.08
Overest-total-friend	.00	-.04	.04
Underest-best-friend	-.08	.00	-.14*
Underest-total-friend	.07	.11**	-.02

Note. Total = total anxiety symptoms; FNE = fear of negative evaluation symptoms; SAD = social avoidance and distress symptoms; like-self = self-perceived likeability; like-peer = peer rated likeability; like-discr = difference between self- and peer-perceived likeability; best-friend-given = self-nominated number of best friends; total-friend-given = self-nominated number of best and good friends; best-friend-received = peer-indicated nominations of best friend; total-friend-received = peer-indicated nominations of total friend; best-friend-reciprocal = reciprocal number of best friends; total-friend-reciprocal = reciprocal number of best and good friends; overest-best-friend = overestimation of number of best friends; overest-total-friend = overestimation of number of best and good friends; underest-best-friend = underestimation of number of best friends; underest-total-friend = underestimation of number of best and good friends.
*$p < .05$. **$p < .01$. ***$p < .001$.

Research Question 2: Social Anxiety and Friendship

As expected, total social anxiety symptoms were significantly related to self-nominating fewer best friends, being less often nominated as best friends, and fewer reciprocal best friends and reciprocal total friend nominations (Goal 2a). No significant relations existed between the Fear of Negative Evaluation subscale and friendship quantity measures. In contrast, the Social Avoidance and Distress subscale was significantly related to lower scores on all friendship quantity measures. Thus, higher levels of avoidance and distress were related to less self-nominated, peer-nominated, and mutual best friends and total friends. In contrast, fear of negative evaluation was not (see Table 2).

The results of the partial correlations between the over- and underestimation scores and social anxiety symptoms are displayed in Table 2 (Goal 2b). Total social anxiety symptoms and fear of negative evaluation were not significantly related to over- or underestimation of best friends and total friends. Only the subscale of social avoidance and distress was negatively related to the *underestimate-best-friend*-score, which suggests that adolescents with high levels of social avoidance and distress were less likely to "miss" best friends they had within their class than adolescents with low levels of social avoidance and distress. In other words, when adolescents show higher levels of avoidance and distress, they are more likely to recognize a friendship.

Discussion

The overall aim of the current study was to investigate how different symptoms of social anxiety in adolescents relate to adolescents' perception accuracy of likeability by peers and friendships with their peers. The first goal was to examine how social anxiety symptoms in adolescents are related to self-perceived likeability, peer-perceived likeability, and the perception accuracy in likeability. The second goal was to examine how social anxiety symptoms relate to self-perceived friendships, peer-perceived friendships, and perception accuracy of friendships within the classroom. As social anxiety can be divided into cognitive factors or symptoms (i.e., worry about negative evaluation) and behavioral symptoms (i.e., avoidance; American Psychiatric Association, 2013; Chavira & Stein, 2005; Clark & Wells, 1995; Ranta et al., 2014), we differentiated between cognitive and behavioral anxiety symptoms in the examination of both research goals.

Social Anxiety and Likeability

We first examined how overall levels of social anxiety were related to self-rated and peer-rated likeability (Goal 1a). As hypothesized, adolescents with higher levels of social anxiety were less liked by their peers according to themselves and their classmates. This finding is in line with other studies that also found that higher social anxiety symptoms were related to lower peer-rated likeability (Baartmans et al., 2020; Verduin & Kendall, 2008) and with studies finding that more social anxiety symptoms were related to decreased peer acceptance and victimization in adolescents (Henricks et al., 2021; Tillfors et al., 2012; Verduin & Kendall, 2008). We also examined the discrepancy between adolescents' own likeability perception and the likeability perception of their peers (Goal 1b). As expected, socially anxious adolescents tended to underestimate their likeability. These findings are in line with previous findings showing that social anxiety symptoms in pre-adolescents and adolescents with mild intellectual disability are related to the underestimation of likeability by peers (Baartmans et al., 2019; Klein et al., 2018). Therefore, by studying a sample of 10- to 15-year-old typically developing adolescents,

the current study found that adolescents' overall levels of anxiety were not only related to actual lower likeability according to peers but also to an underestimation of their likeability.

An important note is that most previous studies did not discriminate between worrying about negative judgment by others and the tendency to avoid social situations as subtypes of social anxiety when studying the accuracy of perceived likeability among peers. In order to overcome this limitation, we also examined how the cognitive and behavioral social anxiety symptoms were uniquely associated with self-perceived likeability, peer-perceived likeability, and the accuracy of perceived likeability. As expected, a higher *tendency to avoid* social situations was associated with being less liked by peers and perceiving oneself as less liked. Therefore, adolescents with a higher tendency to avoid social situations accurately perceive that their peers like them less than other classmates. In contrast, although adolescents with a *tendency to worry* about social situations (fear of negative evaluation) had the perception that they were less liked among their peers, their social anxiety symptoms were unrelated to their actual likeability among peers. Adolescents with cognitive and social anxiety symptoms, therefore, have an inaccurate perception (i.e., underestimation) of one's likeability among peers.

Like Henricks and colleagues (2021, 2023), this study shows that especially avoiding social situations is associated with low likeability among peers. It is known that the innate need to belong becomes more important in adolescence and that adolescents might prefer peers with less strong avoidance tendencies (Baumeister & Leary, 1995; Schoch et al., 2015). At the same time, avoidance tendencies become stronger and more influential in adolescence (Miers et al., 2014), which could explain our finding that the tendency to avoid is related to less likeability by peers. In addition, worrying mainly plays a role in the mind, while others can notice avoidance tendencies, which may lead to lower ratings on likeability.

Social Anxiety and Friendships

Regarding friendships (Goal 2), the results of the current study confirmed the hypothesis and previous research findings that social anxiety is related to fewer friendships (La Greca & Lopez, 1998; Van Zalk et al., 2011). We extended these findings by showing that like for likeability, these findings depend on the type of social anxiety symptoms. In line with our hypothesis, the tendency to avoid social situations and/or experiencing social distress was related to fewer (self-reported, peer-reported, and reciprocal) friendships. In contrast, worrying about negative judgment by others was unrelated to self-perceived and peer-perceived and reciprocal friendships.

In contrast to the findings for likeability and the hypothesis, total levels of social anxiety symptoms were *not* significantly related to more over- or underestimation of the number of best friends and total friends. Thus, adolescents with higher levels of social anxiety did not seem to have a biased perception of the number of friendships within the

class compared to adolescents with lower levels of social anxiety. Surprisingly, a stronger tendency to avoid social situations and/or experiencing social distress was related to less underestimation of the number of best friends within the class, whereas worry was not associated with a biased friendship perception.

Based on theories about cognitive biases in social anxiety, we expected that socially anxious adolescents might underestimate their number of friendships (Hofmann & DiBartolo, 2014; Morrison & Heimberg, 2013), but the results did not show this. The lack of a significant relation between social anxiety symptoms and underestimating the number of one's friendships might be explained by the fact that friendship is a social construct that requires reciprocal liking and, for instance, engaging in joint activities (Bukowski & Hoza, 1989; Demir & Urberg, 2004). These requirements could make it easier and clearer for adolescents to be sure whether or not someone can be considered as a friend. Thus, the fact that friendship quantity might be easier to perceive than liking could possibly explain why adolescents with higher levels of (cognitive) social anxiety seem to have more problems with accurately estimating a more general concept, such as their likeability, and not with a more concrete concept, such as friendship.

Strengths, Limitations, and Suggestions for Future Research

A strength of the current study was the differentiation of the social anxiety symptoms into the subscales of fear of negative evaluation and the tendency to avoid social situations. This provides further information about the specific relation between social anxiety symptoms and social functioning measures. An additional strength was using a multiple-informant approach and including different measures of likeability and friendship. Limitations also need to be mentioned.

A first limitation of the current study was that we included a typically developing community sample only and no clinical sample. Even though the levels of social anxiety were comparable to other community sample studies (Kärnä et al., 2010), the scores of social anxiety were relatively low. This limits the generalization of the results to children with high levels of social anxiety or clinical samples. Clearly, more research is needed on children with high levels of social anxiety and clinical samples to draw more robust conclusions and to recommend implications for treatment. Second, the current study used a cross-sectional design. Therefore, no conclusions can be drawn about the causality or longitudinal effects of social anxiety symptoms and perception accuracy of social functioning. Third, using self-reports in a classroom setting could have led to socially desirable answers. Adolescents filled in the questionnaires in a 'test setting' so that others could not see their answers, and we stressed that there were no right or wrong answers and that the results would not be shared with anyone. Still, we cannot rule out the fact that social desirability might have influenced our findings. Fourth, there were some limitations regarding our friendship measures. We only measured friendships within the classrooms. Therefore, we cannot generalize these results to friendships in

general. It could, for instance, be that socially anxious adolescents have more friendships outside the classroom (e.g., friends from primary school or sports).

The findings of the current study lead to some suggestions for future research. First, as we only included adolescents between the ages of 10-15 in the current study, little is known about how the associations between social anxiety, likeability, and friendships vary by age. A broader age range would allow drawing further conclusions regarding the relation between social anxiety symptoms and perceptions of social functioning across ages. Second, the current study extended previous research using a multi-method approach using self- and peer-reported measures. This approach could be further strengthened by including measures of social anxiety and social functioning by other raters, like teachers and parents. Third, a suggestion for future research would be to include measures to examine depression and externalizing problems. As depression and externalizing disorders are found to influence likeability and friendship and often co-occur with social anxiety disorder (Mohammadi et al., 2020), it would be especially important to include both variables when testing the associations between social anxiety, likeability, and friendships, to get a better understanding of the complex constructs of likeability and friendships in relation to mental health. Fourth, further research could be conducted to get more fine-grained insights into the accuracy of perceptions of adolescents' friendships. In this study, we determined over- and underestimating friendships by studying discrepancies in nominations of friendships within the classroom. This differed from the likeability accuracy scores since we did not ask who the adolescent thought would indicate them as friends. This would have allowed us to distinguish adolescents not noticing others wanting to be friends with them from adolescents noticing a peer wanting to be friends with them but choosing not to engage in a friendship with that peer. Including these measures in future research would allow for drawing more precise conclusions on the accuracy of estimations.

Practical Implications

The current study stresses the importance of targeting avoidance behavior in treating social anxiety symptoms. When adolescents with higher levels of social anxiety report the tendency to avoid, they are more likely to be disliked by peers, which could make their socially anxious thoughts about being negatively evaluated by others warranted. If replicated in a sample with high levels of social anxiety, these results might suggest that in addition to the cognitive training in CBT, encouraging socially anxious adolescents to approach and practice social situations rather than avoid them appears to be an important aspect of treatment. In addition, following the Extended Process Model (Gross, 2015), it may be interesting to examine which of the five ways (situation selection, situation modification, attentional deployment, cognitive change, and/or response modulation) is the most relevant (or best) target for treating social anxiety.

Conclusion

In conclusion, the results of the current study provide evidence that even though social anxiety symptoms are related to both lower self-estimates and peer-estimates of likeability, adolescents with higher levels of social anxiety still seem to underestimate their likeability among classroom peers. In addition, we found that adolescents who worried more about negative evaluations underestimated their likeability by peers; they think they are less liked than they are. No such effects occurred for worrying about negative evaluations and perceptions of friendship. Further, adolescents with strong social avoidance tendencies accurately perceived their likeability and friendship; they were less liked by their peers and had fewer friendships, and they also perceived their situation as such. Even though these findings must be reviewed in light of the current study's limitations, these results highlight the possible importance of disentangling different subtypes of social anxiety and using multiple informants of likeability and friendships in future research and clinical practice. If replicated in a group of children with high (clinical) levels of social anxiety, these results emphasize the importance of explicitly treating avoidance behavior in social anxiety since this may not only maintain the social anxiety symptoms but were also found to be related to more negative judgments by others.

Funding: There are no significant sources of funding to declare.

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: There are no financial or personal competing interests for any author in the making of the study.

Ethics Statement: The authors confirm that they complied with the recognized ethical standards and local guidelines. All participants and parents were properly instructed and gave informed consent. The Ethical Committee of the Behavioural Science Institute of Radboud University Nijmegen, the Netherlands, approved this study.

Preregistration: There was no preregistration of the study.

Reporting Guidelines: We followed APA7 reporting guidelines and JARS-Quant guidelines for writing the manuscript.

Social Media Accounts: @ReinoutWiers, @SusanBogels

Related Versions: The manuscript was based on Chapter 8 of the dissertation of the first author:
Baartmans, J. M. D. (2021). *Childhood social anxiety: What's next? Exploring the role of cognitions, depression, parents, and peers.* University of Amsterdam. URL: https://pure.uva.nl/ws/files/54463245/Thesis_complete_.pdf

Data Availability: Data from our study is available by contacting the corresponding author.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Baartmans, J. M. D., Rinck, M., Hudson, J. L., Lansu, T. A., van Niekerk, R. E., Bögels, S. M., & Klein, A. M. (2019). Are socially anxious children really less liked, or do they only think so? *Cognitive Therapy and Research, 43*(6), 1043–1050. https://doi.org/10.1007/s10608-019-10028-9

Baartmans, J. M. D., van Steensel, F. J., Mobach, L., Lansu, T. A., Bijsterbosch, G., Verpaalen, I., Rapee, R. M., Magson, N., Bögels, S. M., Rinck, M., & Klein, A. M. (2020). Social anxiety and perceptions of likeability by peers in children. *British Journal of Developmental Psychology, 38*(2), 319–336. https://doi.org/10.1111/bjdp.12324

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Bukowski, W. M., & Hoza, B. (1989). Popularity and friendship: Issues in theory, measurement, and outcome. In T. J. Berndt & G. W. Ladd (Eds.), *Peer relationships in child development* (pp. 15–45). John Wiley & Sons.

Blöte, A. W., Bokhorst, C. L., Miers, A. C., & Westenberg, P. M. (2012). Why are socially anxious adolescents rejected by peers? The role of subject-group similarity characteristics. *Journal of Research on Adolescence, 22*(1), 123–134. https://doi.org/10.1111/j.1532-7795.2011.00768.x

Blöte, A. W., Miers, A. C., Heyne, D. A., Clark, D. M., & Westenberg, P. M. (2014). The relation between social anxiety and audience perception: Examining Clark and Wells' (1995) model among adolescents. *Behavioural and Cognitive Psychotherapy, 42*(5), 555–567. https://doi.org/10.1017/S1352465813000271

Bögels, S. M., Alden, L., Beidel, D. C., Clark, L. A., Pine, D. S., Stein, M. B., & Voncken, M. (2010). Social anxiety disorder: Questions and answers for the DSM-V. *Depression and Anxiety, 27*(2), 168–189. https://doi.org/10.1002/da.20670

Cantin, S., Brendgen, M., Dussault, F., & Vitaro, F. (2019). Transactional links between adolescents' and friends' victimization during the first two years of secondary school: The mediating role of likeability and friendship involvement. *Social Development, 28*(3), 743–757. https://doi.org/10.1111/sode.12355

Chavira, D. A., & Stein, M. B. (2005). Childhood social anxiety disorder: From understanding to treatment. *Child and Adolescent Psychiatric Clinics of North America, 14*(4), 797–818. https://doi.org/10.1016/j.chc.2005.05.003

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

Demir, M., & Urberg, K. A. (2004). Friendship and adjustment among adolescents. *Journal of Experimental Child Psychology, 88*(1), 68–82. https://doi.org/10.1016/j.jecp.2004.02.006

Erath, S. A., Flanagan, K. S., & Bierman, K. L. (2007). Social anxiety and peer relations in early adolescence: Behavioral and cognitive factors. *Journal of Abnormal Child Psychology, 35*(3), 405–416. https://doi.org/10.1007/s10802-007-9099-2

Erath, S. A., Flanagan, K. S., Bierman, K. L., & Tu, K. M. (2010). Friendships moderate psychosocial maladjustment in socially anxious early adolescents. *Journal of Applied Developmental Psychology, 31*(1), 15–26. https://doi.org/10.1016/j.appdev.2009.05.005

Erath, S. A., Flanagan, K. S., & Bierman, K. L. (2008). Early adolescent school adjustment: Associations with friendship and peer victimization. *Social Development, 17*(4), 853–870. https://doi.org/10.1111/j.1467-9507.2008.00458.x

Greco, L. A., & Morris, T. L. (2005). Factors influencing the link between social anxiety and peer acceptance: Contributions of social skills and close friendships during middle childhood. *Behavior Therapy, 36*(2), 197–205. https://doi.org/10.1016/S0005-7894(05)80068-1

Gross, J. J. (2015). The extended process model of emotion regulation: Elaborations, applications, and future directions. *Psychological Inquiry, 26*(1), 130–137. https://doi.org/10.1080/1047840X.2015.989751

Henricks, L. A., Lange, W. G., Luijten, M., van den Berg, Y. H. M., Stoltz, S. E. M. J., Cillessen, A. H. N., & Becker, E. S. (2023). The longitudinal link between popularity, likeability, fear of negative evaluation and social avoidance across adolescence. *Journal of Research on Adolescence, 33*(3), 720–734. https://doi.org/10.1111/jora.12833

Henricks, L. A., Pouwels, J. L., Lansu, T. A. M., Lange, W., Becker, E. S., & Klein, A. M. (2021). Prospective associations between social status and social anxiety in early adolescence. *British Journal of Developmental Psychology, 39*(3), 462–480. https://doi.org/10.1111/bjdp.12374

Hofmann, S. G., & DiBartolo, P. M. (2014). *Social anxiety: Clinical, developmental, and social perspectives*. Academic Press.

Kärnä, A., Voeten, M., Poskiparta, E., & Salmivalli, C. (2010). Vulnerable children in varying classroom contexts: Bystanders' behaviors moderate the effects of risk factors on victimization. *Merrill-Palmer Quarterly, 56*(3), 261–282. https://doi.org/10.1353/mpq.0.0052

Karkavandi, M. A., Wang, P., Lusher, D., Bastian, B., McKenzie, V., & Robins, G. (2022). Perceived friendship network of socially anxious adolescent girls. *Social Networks, 68*, 330–345. https://doi.org/10.1016/j.socnet.2021.08.011

Klein, A. M., Houtkamp, E. O., Salemink, E., Baartmans, J. M. D., Rinck, M., & van der Molen, M. J. (2018). Differences between self-and peer-rated likability in relation to social anxiety and depression in adolescents with mild intellectual disabilities. *Research in Developmental Disabilities, 80*, 44–51. https://doi.org/10.1016/j.ridd.2018.05.016

Kingery, J. N., Erdley, C. A., Marshall, K. C., Whitaker, K. G., & Reuter, T. R. (2010). Peer experiences of anxious and socially withdrawn youth: An integrative review of the developmental and clinical literature. *Clinical Child and Family Psychology Review, 13*(1), 91–128. https://doi.org/10.1007/s10567-009-0063-2

La Greca, A. M., & Harrison, H. M. (2005). Adolescent peer relations, friendships, and romantic relationships: Do they predict social anxiety and depression? *Journal of Clinical Child and Adolescent Psychology, 34*(1), 49–61. https://doi.org/10.1207/s15374424jccp3401_5

La Greca, A. M., & Lopez, N. (1998). Social anxiety among adolescents: Linkages with peer relations and friendships. *Journal of Abnormal Child Psychology, 26*(2), 83–94. https://doi.org/10.1023/A:1022684520514

Leigh, E., & Clark, D. M. (2018). Understanding social anxiety disorder in adolescents and improving treatment outcomes: Applying the Cognitive Model of Clark and Wells (1995). *Clinical Child and Family Psychology Review, 21*(3), 388–414. https://doi.org/10.1007/s10567-018-0258-5

Miers, A. C., Blöte, A. W., Heyne, D. A., & Westenberg, P. M. (2014). Developmental pathways of social avoidance across adolescence: The role of social anxiety and negative cognition. *Journal of Anxiety Disorders, 28*(8), 787–794. https://doi.org/10.1016/j.janxdis.2014.09.008

Miers, A. C., Blöte, A. W., & Westenberg, M. M. (2011). Negative social cognitions in socially anxious youth: Distorted reality or a kernel of truth? *Journal of Child and Family Studies, 20*(2), 214–223. https://doi.org/10.1007/s10826-010-9423-2

Miers, A. C., Blöte, A. W., & Westenberg, P. M. (2010). Peer perceptions of social skills in socially anxious and nonanxious adolescents. *Journal of Abnormal Child Psychology, 38*(1), 33–41. https://doi.org/10.1007/s10802-009-9345-x

Mohammadi, M. R., Salehi, M., Khaleghi, A., Hooshyari, Z., Mostafavi, S. A., Ahmadi, N., Hojjat, S. K., Safavi, P., & Amanat, M. (2020). Social anxiety disorder among children and adolescents: A nationwide survey of prevalence, socio-demographic characteristics, risk factors and co-morbidities. *Journal of Affective Disorders, 263*, 450–457. https://doi.org/10.1016/j.jad.2019.12.015

Morrison, A. S., & Heimberg, R. G. (2013). Social anxiety and social anxiety disorder. *Annual Review of Clinical Psychology, 9*, 249–274. https://doi.org/10.1146/annurev-clinpsy-050212-185631

Pouwels, J. L., van Noorden, T. H. J., & Caravita, S. C. S. (2019). Defending victims of bullying in the classroom: The role of moral responsibility and social costs. *Journal of Experimental Social Psychology, 84*, Article 103831. https://doi.org/10.1016/j.jesp.2019.103831

Pouwels, J. L., van Noorden, T. H. J., Lansu, T. A. M., & Cillessen, A. H. N. (2018). The participant roles of bullying in different grades: Prevalence and social status profiles. *Social Development, 27*(4), 732–747. https://doi.org/10.1111/sode.12294

Ranta, K., Tuomisto, M. T., Kaltiala-Heino, R., Rantanen, P., & Marttunen, M. (2014). Cognition, imagery and coping among adolescents with social anxiety and phobia: Testing the Clark and Wells model in the population. *Clinical Psychology & Psychotherapy, 21*(3), 252–263. https://doi.org/10.1002/cpp.1833

Rapee, R. M., & Spence, S. H. (2004). The etiology of social phobia: Empirical evidence and an initial model. *Clinical Psychology Review, 24*(7), 737–767. https://doi.org/10.1016/j.cpr.2004.06.004

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy, 35*(8), 741–756. https://doi.org/10.1016/S0005-7967(97)00022-3

Schoch, S., Nikitin, J., & Freund, A. M. (2015). Why do(n't) you like me? The role of social approach and avoidance motives in attributions following social acceptance and rejection. *Motivation and Emotion, 39*(5), 680–692. https://doi.org/10.1007/s11031-015-9482-1

Tillfors, M., Persson, S., Willén, M., & Burk, W. J. (2012). Prospective links between social anxiety and adolescent peer relations. *Journal of Adolescence, 35*(5), 1255–1263. https://doi.org/10.1016/j.adolescence.2012.04.008

Verduin, T. L., & Kendall, P. C. (2008). Peer perceptions and liking of children with anxiety disorders. *Journal of Abnormal Child Psychology, 36*(4), 459–469. https://doi.org/10.1007/s10802-007-9192-6

Vernberg, E. M., Abwender, D. A., Ewell, K. K., & Beery, S. H. (1992). Social anxiety and peer relationships in early adolescence: A prospective analysis. *Journal of Clinical Child Psychology, 21*(2), 189–196. https://doi.org/10.1207/s15374424jccp2102_11

Voncken, M. J., & Bögels, S. M. (2008). Social performance deficits in social anxiety disorder: Reality during conversation and biased perception during speech. *Journal of Anxiety Disorders, 22*(8), 1384–1392. https://doi.org/10.1016/j.janxdis.2008.02.001

Van Zalk, N., van Zalk, M., Kerr, M., & Stattin, H. (2011). Social anxiety as a basis for friendship selection and socialization in adolescents' social networks. *Journal of Personality, 79*(3), 499–526. https://doi.org/10.1111/j.1467-6494.2011.00682.x

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

Assessing Diagnostic Precision: Adaptations of the Hopkins Symptom Checklist (HSCL-5/10/25) Among Tertiary-Level Students in Norway

Børge Sivertsen[1,2], Jens C. Skogen[1,3,4], Anne Reneflot[5], Marit Knapstad[1], Otto Robert Frans Smith[1,4,6], Leif Edvard Aarø[1], Benedicte Kirkøen[4], Bengt Oscar Lagerstrøm[7], Ann Kristin Skrindo Knudsen[8]

[1] *Department of Health Promotion, Norwegian Institute of Public Health, Bergen, Norway.* **[2]** *Department of Research & Innovation, Helse-Fonna HF, Haugesund, Norway.* **[3]** *Center for Alcohol & Drug Research, Stavanger University Hospital, Stavanger, Norway.* **[4]** *Centre for Evaluation of Public Health Measures, Norwegian Institute of Public Health, Oslo, Norway.* **[5]** *Department of Mental Health and Suicide, Norwegian Institute of Public Health, Oslo, Norway.* **[6]** *Department of Teacher Education, NLA University College, Bergen, Norway.* **[7]** *Department for Methodology and Data Collection, Statistics Norway, Oslo, Norway.* **[8]** *Department of Disease Burden, Norwegian Institute of Public Health, Bergen, Norway.*

Clinical Psychology in Europe, 2024, Vol. 6(4), Article e13275, https://doi.org/10.32872/cpe.13275

Received: 2023-11-27 • **Accepted:** 2024-08-01 • **Published (VoR):** 2024-12-20

Handling Editor: Winfried Rief, Philipps-University of Marburg, Marburg, Germany

Corresponding Author: Børge Sivertsen, Department of Health Promotion, Norwegian Institute of Public Health, Zander Kaaes gate 7, 5018 Bergen, Norway. Phone: +47 99 76 92 62. E-mail: borge.sivertsen@fhi.no

Abstract

Background: Universities worldwide are witnessing a surge in mental health problems among students, particularly in anxiety and depression. The Hopkins Symptom Checklist (HSCL) is a popular screening tool, but its reliability in identifying mental disorders remains debated. The aim of this study was to evaluate the criterion validity of the HSCL-25, HSCL-10, and HSCL-5 using 30-day prevalence of major depressive episode (MDE) and generalized anxiety disorder (GAD) from a self-administered electronic version of the Composite International Diagnostic Interview, fifth version (CIDI 5.0), as the benchmark.

Method: Data stem from a national survey targeting students in higher education in Norway. In a 2023 follow-up study on mental disorders, 5,568 participants completed both the HSCL-25 and the CIDI. Sex-specific optimal thresholds for all HSCL versions in relation to MDE and GAD (from CIDI) were determined using the Youden Index maximization.

Results: The optimal cut-off values for detecting MDE or GAD with the HSCL-25 were 1.96 for males and 2.20 for females, displaying a good balance between sensitivity and specificity. Similar high and balanced sensitivity and specificity patterns were found for both the HSCL-10 and HSCL-5. However, all HSCL versions overestimated prevalence rates compared to the self-administered CIDI.
Conclusions: All three HSCL versions showed high criterion validity. The data indicate that HSCL may be better as a screening tool than for precise estimation of MDE and GAD prevalence. For improved diagnostic accuracy, future HSCL versions should incorporate functional impairment assessment. This update would bring the HSCL into closer alignment with clinical diagnostic standards.

Keywords

depression, anxiety, students, young adults, questionnaires, psychometrics

Highlights

- This study assessed HSCL's ability to detect anxiety and depression in Norwegian college students.
- A self-administered version of CIDI was used as the gold standard for diagnostic accuracy.
- All three HSCL versions showed high criterion validity and good diagnostic precision.
- HSCL may be better as a screening tool than for precise estimation of mental disorders.

Universities and colleges worldwide are confronting a concerning rise in the incidence of mental health issues among their students, with projections significantly surpassing those observed in the general population (Ibrahim et al., 2013; McCloud et al., 2023). Recent systematic reviews on anxiety and depression have yielded a cumulative annual prevalence range of 25-30% among tertiary education students (Chi et al., 2023; Ibrahim et al., 2013; Kou et al., 2012; Sheldon et al., 2021).

Although diagnostic interviews maintain their status as the benchmark for mental disorder diagnosis (Nordgaard et al., 2013; Rettew et al., 2009), their demanding time and resource requirements hinder most researchers from incorporating comprehensive psychiatric interviews into their assessment battery. Consequently, a substantial portion of investigations are constrained to incorporating brief survey questionnaires when evaluating mental health issues (Auerbach et al., 2016). However, in order to obtain prevalence estimates that are as accurate as possible, it is important that the case-detection capabilities of these briefer questionnaires are thoroughly evaluated. Also, considering the changing trends and disparities observed across age cohorts and research populations, it is imperative to re-examine previously established cut-off values to ensure that the case-detection capabilities are effectively validated for each distinct study population.

In this context, a frequently employed survey instrument is the Hopkins Symptom Checklist (HSCL), initially devised during the 1950s as a clinical tool for assessing symptoms of several mental disorders. Although the original iteration encompassed an extensive array of mental disorders, one of the most commonly used versions today is the HSCL-25. This abbreviated version focuses on two symptom dimensions, anxiety and depression. However, the capacity of the HSCL-25 to differentiate between these two conditions across different sexes is not fully established (Sandanger et al., 1998; Skogen et al., 2017). Since the 1990s, the efficacy of HSCL-25 in identifying cases has been examined a few times through comparisons with structured diagnostic interviews, suggesting that while HSCL-25 performs adequately in detecting depression, the results for anxiety are more variable (Sandanger et al., 1998; Veijola et al., 2003).

The HSCL-25 uses a scoring range of 1 to 4, and to determine the commonly used mean HSCL-25 score, the total score is divided by the item count. A traditional cut-off value of 1.75 is commonly employed to indicate major depressive disorder (Glaesmer et al., 2014; Sandanger et al., 1998; Veijola et al., 2003). However, a recent Spanish study has challenged this one-size-fits-all approach, proposing distinct optimal cut-off values for women (1.76) and men (1.84) (Rodríguez-Barragán et al., 2021). This suggests that sex-specific thresholds might be necessary for more accurate diagnosis, considering the different ways in which men and women experience and report mental health symptoms.

Further abbreviated iterations of the HSCL have been developed, with both HSCL-5 and HSCL-10 now in widespread use. However, few studies have investigated the case-detection capability of these versions, and none have done so for males and females separately. To the best of our knowledge, just one study has directly contrasted HSCL-5 and HSCL-10 against a structured diagnostic interview. A recent Spanish study demonstrated good reliability and validity of both tools for detecting depression (Rodríguez-Barragán et al., 2023). Additionally, a recent Norwegian study assessed HSCL-5's capacity for identifying cases in the general population, and its findings indicate that HSCL-5 effectively identifies people with generalized anxiety disorder or major depressive disorder within this context (Kirkøen et al., Manuscript in preparation).

Based on these considerations, the objective of this current study is to assess the proficiency of HSCL-25, HSCL-10, and HSCL-5 in detecting cases of major depressive episode (MDE) and generalized anxiety disorder (GAD) in a national sample of college and university students. Our focus is on potential sex-specific cut-off values and differences, using a recently developed self-administered electronic version of the Composite International Diagnostic Interview (CIDI) 5.0. This approach aims to enhance the accuracy of case-detection and ensure that the nuanced mental health experiences of both sexes are adequately captured.

Method

Setting and Participants

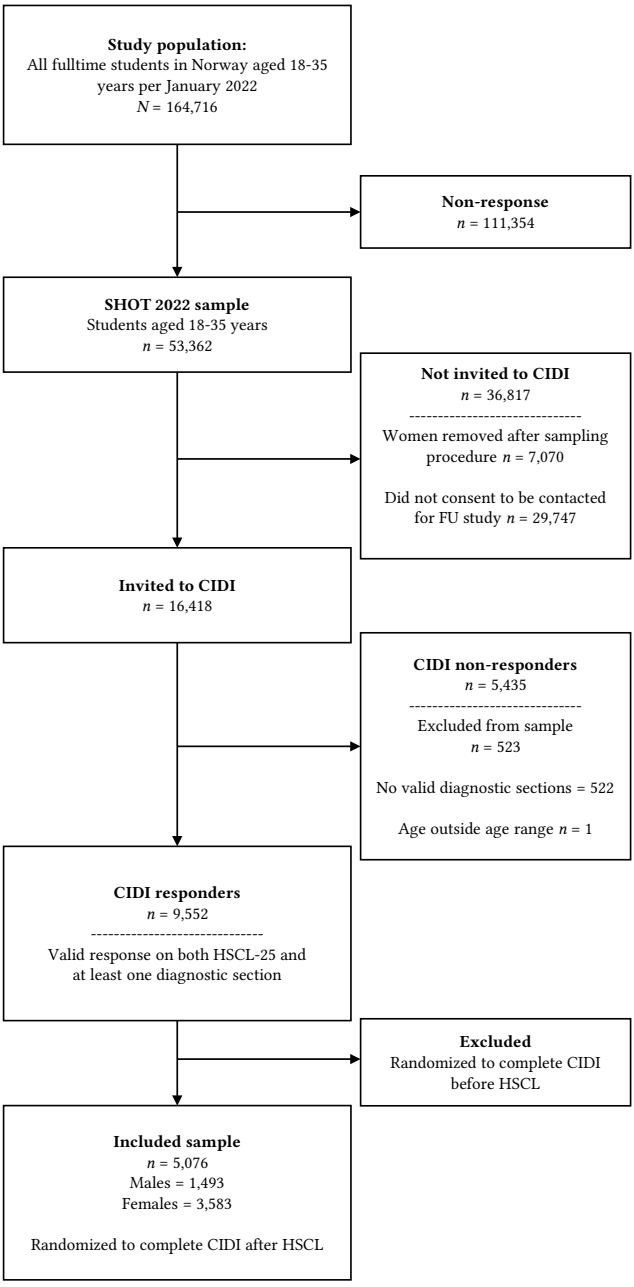
The primary population for this study is derived from the Norwegian SHOT study (Students' Health and Wellbeing Study), a nationwide survey that centres on students pursuing higher education. Since 2010, four major surveys have been conducted, with the latest wave conducted in 2022. The SHOT2022 survey comprehensively explored a multitude of dimensions encompassing health and lifestyle. These dimensions included psychological distress, suicidality, life satisfaction, loneliness, sleep problems, sexual harassment, pain, physical exercise, alcohol and drug use, as well as demographic and educational parameters. Comprehensive information concerning the SHOT study has been previously documented (Sivertsen et al., 2019).

During the survey period, SHOT2022 was distributed electronically via a web-based platform and was open for submissions from February 8 to April 19, 2022. Invitations for participation were sent to all full-time Norwegian students engaged in higher education, both within the country and abroad. Extensive efforts were undertaken to increase awareness about the study through channels such as email, SMS, and informational campaigns conducted by welfare organizations and educational institutions. A total of 169,572 students met the study's inclusion criteria, which required them to be full-time college or university students and hold Norwegian citizenship. Of these, 59,544 students completed the online questionnaire after receiving two reminders. This resulted in a response rate of 35.1% (which did not differ between geographical regions). For the present study, the inclusion criteria specified that participants be between the ages of 18 and 35 years. Consequently, a subset of 53,362 students within this age range was selected for analysis. When consenting to participate in the SHOT2022, students were given the option to express their interest in participating in a follow-up study on mental disorders. Out of the total participants, 26,311 students consented to be a part of this follow-up study. To better reflect the sex distribution of the base study population, more male students were invited to participate in the CIDI study. Consequently, 16,418 students, officially registered as of January 2023, were invited. However, fewer male students consented to follow-up contact compared to females. This led to females comprising a higher proportion (70.4%) of the invitations for the CIDI study.

Figure 1 illustrates the participation process for the current study. Out of 9,552 students who provided valid responses on both the HSCL-25 and at least one of the CIDI diagnostic sections, half were randomly selected to complete the HSCL-25 before the CIDI. This resulted in a subset of 5,076 participants who provided valid scores on the HSCL-25 and then completed the CIDI. The remaining half completed the assessments in the reverse order. This deliberate sequencing was chosen to facilitate future investigations into the potential influences of the order of questionnaire administration on the reported results from the HSCL-25.

Figure 1

Participant Flow in the Study



For the purposes of the present study, we only used observations where the HSCL-25 was administered before the CIDI, as screening instruments are normally not preceded by full diagnostic assessments in population-based surveys. The CIDI study took place between January 24 and February 6, 2023, approximately 12 months after the SHOT2022 survey was conducted. More detailed information on the participation process has been published elsewhere (Sivertsen et al., 2023).

Instruments

Sociodemographic Information

Participants' age and sex information were derived from their 11-digit Norwegian national identity numbers.

Depression and Anxiety Disorders: The CIDI

The data collection utilized a recently developed electronic, self-administered version of the Composite International Diagnostic Interview, fifth version (CIDI 5.0), developed for the World Health Organization (WHO) World Mental Health (WMH) Surveys (Kessler & Üstün, 2004). Both the original CIDI and this self-administered version maintained the same sequence of CIDI modules, and the wording of the questions remained nearly identical, though instructions were slightly adapted for self-administration. To reduce participant burden and enhance response rates, the current version excluded certain diagnostic categories, namely bipolar, obsessive-compulsive, PTSD, ADHD, and personality disorders. In short, CIDI 5.0 is a standardized interview that assesses 30-days, 12 months and lifetime prevalence for several mental and substance use disorders according to diagnostic criteria in the Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5) (American Psychiatric Association, 2013). The interview version of CIDI 3.0 has shown good concordance with diagnostic instruments such as the Structured Clinical Interview for DSM-IV (SCID) (Haro et al., 2006) and Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (Jordanova et al., 2004).

Current mental disorder was defined as the presence of MDE or GAD during the 30 days before study participation. We also estimated prevalence rates for mental disorders spanning a 12-month period and over the lifetime, but these estimations were not incorporated into the present study, aligning with its specific emphasis on assessing current mental disorders in comparison with the HSCL. The operationalization of these diagnoses was based on algorithms developed for CIDI 5.0 in the WMH Surveys Initiative.

The Hopkins Symptom Checklist (HSCL)

Mental health problems in the 14 days before the survey were assessed by the widely used *Hopkins Symptom Checklist* (HSCL-25) (Derogatis et al., 1974), derived from the 90-item Symptom Checklist (SCL-90). The score for the HSCL is determined by taking

the sum of the item scores and dividing it by the total number of items responded to, yielding a potential range of 1 to 4. An investigation of the factor structure based on the SHOT2014 dataset showed that a unidimensional model had the best psychometric properties in the student population and not the original subscales of anxiety and depression (Skogen et al., 2017). Details on the development of mental health problems assessed with the HSCL-25 in the SHOT waves were recently published by Knapstad and colleagues (Knapstad et al., 2021). The items included in the three different versions of the HSCL are listed in Table 1.

Table 1

Items Included in the Different HSCL Iterations[a]

Item content	HSCL-25	HSCL-10	HSCL-5
1. Suddenly scared for no reason	◯	◯	
2. Feeling fearful	◯	◯	◯
3. Faintness, dizziness, or weakness	◯	◯	
4. Nervousness or shakiness inside	◯		◯
5. Heart pounding or racing	◯		
6. Trembling	◯		
7. Feeling tense or keyed up	◯	◯	
8. Headaches	◯		
9. Spells of terror or panic	◯		
10. Feeling restless, can't sit still	◯		
11. Feeling low in energy, slowed down	◯		
12. Blaming yourself for things	◯	◯	
13. Crying easily	◯		
14. Loss of sexual interest or pleasure	◯		
15. Poor appetite	◯		
16. Difficulty falling asleep, staying asleep	◯	◯	
17. Feeling hopeless about the future	◯	◯	◯
18. Feeling blue	◯	◯	◯
19. Feeling lonely	◯		
20. Feeling trapped or caught	◯		
21. Worrying too much about things	◯		◯
22. Feeling no interest in things	◯		
23. Thoughts of ending your life	◯		
24. Feeling everything is an effort	◯	◯	
25. Feelings of worthlessness	◯	◯	

[a]Each symptom is rated on a four-point frequency scale (1 = not at all, 2 = a little, 3 = quite a lot, 4 = extremely).

Statistical Analyses

In the present study, we first present summary statistics for HSCL scores and the prevalence of MDE and GAD, stratified by sex (Table 3 and Table 4). Subsequently, we proceeded to determine sex-specific optimal cut-off points for HSCL-25, HSCL-10, and HSCL-5 in relation to either MDE or GAD combined, and for MDE and GAD separately, as identified through CIDI. This estimation was grounded in the maximization of the Youden Index, which is a commonly used metric for binary classification in validation studies aimed at striking a balance between sensitivity and specificity. The formula for the Youden Index, denoted as ('sensitivity' + 'specificity') - 1, yields a scale ranging from 0 to 1. Higher values signify better discriminative capacity, where 0 denotes no discrimination, and 1 reflects perfect discrimination. Although rules of thumb always must be considered in conjunction with other aspects, a score below 0.5 on the Youden Index indicates that the test in question may not be useful as a classification tool, whereas a score surpassing 0.5 can be construed as indicating a valuable test. Alongside the Youden Index, we also present a comprehensive view of overall accuracy, sensitivity, specificity, positive and negative predictive values (PPV and NPV), and the area under the curve (AUC). All analyses were performed using R 4.2.2.

Results

As detailed in Table 2, the sample for the CIDI study primarily consisted of female students of Norwegian ethnicity, with an average age of 24 years, and about half were single. Most participants had parents with high educational levels. Compared with the overall SHOT2022 study, the sociodemographic characteristics were similar, except for a slightly higher proportion of females in the CIDI study (70%) compared to the SHOT2022 study (66%), as shown in Table 2. Non-respondents to the CIDI study, who were invited but did not participate, differed mainly in having parents with lower education levels.

The level of mental health problems, measured by the HSCL-25 in the SHOT2022 study, was marginally lower in CIDI respondents (M = 1.88, SD = 0.61) compared to non-respondents (M = 1.90, SD = 0.61, Cohen's d = 0.03). However, CIDI respondents had a slightly higher HSCL-25 score than the overall SHOT2022 sample (M = 1.86, SD = 0.57, Cohen's d = 0.03).

Figure 1 details the participation process of the current study. A total of 5,076 participants completed both the CIDI and HSCL-25 in advance. The mental health characteristics of the sample are detailed in Table 3. Females reported considerably higher average scores on all HSCL iterations, and the prevalences of 30-day MDE and GAD were much higher among females as well, as detailed in Table 3.

Table 2

Demographical Characteristics and Representativeness of the CIDI Responders, CIDI Non-Responders and the Overall Sample Based on Data From 2022

Characteristic	CIDI responders (*n* = 9,552)	CIDI non-responders (*n* = 6,993)	*p*[a]	SHOT2022[b] (*n* = 53,362)	*p*[a]
Age, mean (*SD*)	24.03 (3.28)	23.97 (3.24)	.24	23.98 (1.85)	.14
Sex, % (*n*)			.35		< .001
Women	70.0 (6,686)	71.0 (4,968)		66.4 (35,423)	
Men	30.0 (2,866)	29.0 (2,025)		33.6 (17,939)	
Marital status, % (*n*)			.20		.81
Single	51.3 (4,904)	50.4 (3,526)		51.0 (27,197)	
Boy-/girlfriend	22.5 (2,152)	23.7 (1,659)		22.8 (12,152)	
Cohabitant	22.6 (2,156)	22.4 (1,563)		22.6 (12,058)	
Married/registered partner	3.2 (308)	3.0 (207)		3.1 (1,667)	
Missing	*0.3 (32)*	*0.5 (38)*		*0.5 (288)*	
Maternal education, % (*n*)			.01		.27
Primary	4.3 (407)	5.3 (369)		4.5 (2,407)	
Secondary	27.2 (2,601)	27.6 (1,931)		27.6 (14,707)	
College/university	65.9 (6,290)	64.2 (4,488)		64.3 (34,326)	
Missing	*2.7 (254)*	*2.9 (205)*		*3.6 (1,992)*	
Paternal education, % (*n*)			.02		.39
Primary	5.7 (544)	6.8 (473)		6.0 (3,182)	
Secondary	34.9 (3,335)	35.0 (2,449)		35.1 (18,735)	
College/university	54.6 (5,211)	52.7 (3,687)		53.3 (28,446)	
Missing	*4.8 (462)*	*5.5 (384)*		*5.6 (2,999)*	
HSCL-25, Mean (*SD*)	1.88 (0.61)	1.90 (0.61)	.03	1.86 (0.59)	< .001
Missing, % (n)	*0.2 (17)*	*0.3 (24)*		*0.4 (214)*	

Note. SHOT2022 = Students' Health and Wellbeing Study 2022; CIDI = Composite International Diagnostic Interview; HSCL-25 = Hopkins Symptoms Checklist – 25 items version.
[a]compared with the CIDI responders group (*p*-values based on Chi-squared test [categorical variables] or *t*-test [continuous variables]). [b]grand mean for the SHOT2022 sample aged 18-35.

Table 3

Mental Health Characteristics of the Study Sample

Characteristic	Males, *N* = 1,493	Females, *N* = 3,583	*p*
HSCL-25, *M* (*SD*)	1.67 (0.53)	1.96 (0.60)	< .001
HSCL-10, *M* (*SD*)	1.70 (0.59)	1.98 (0.67)	< .001
HSCL-5, *M* (*SD*)	1.87 (0.74)	2.23 (0.80)	< .001
Major depressive episode (MDE)	9.8%	16.9%	< .001
Generalized anxiety disorder (GAD)	8.0%	15.7%	< .001
MDE or GAD	13.1%	23.8%	< .001

Table 4

Optimal Cut-Off Values for Females for HSCL-25, HSCL-10, and HSCL-5 for Major Depressive Episode (MDE) and Generalized Anxiety Disorder (GAD) Assessed by CIDI

HSCL version and CIDI diagnosis	Cut-off value	Youden Index	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
HSCL-25								
MDE or GAD	≥ 2.20	0.72	0.85	0.89	0.83	0.62	0.96	0.94
MDE	≥ 2.28	0.74	0.84	0.92	0.82	0.51	0.98	0.94
GAD	≥ 2.08	0.64	0.76	0.91	0.73	0.38	0.98	0.90
HSCL-10								
MDE or GAD	≥ 2.30	0.70	0.83	0.88	0.81	0.60	0.96	0.93
MDE	≥ 2.44	0.73	0.86	0.88	0.85	0.55	0.97	0.93
GAD	≥ 2.30	0.63	0.77	0.88	0.75	0.39	0.97	0.89
HSCL-5								
MDE or GAD	≥ 2.75	0.68	0.86	0.81	0.87	0.67	0.94	0.91
MDE	≥ 2.75	0.69	0.83	0.86	0.83	0.50	0.97	0.91
GAD	≥ 2.80	0.64	0.82	0.83	0.81	0.45	0.96	0.89

Note. CIDI = Composite International Diagnostic Interview; HSCL = Hopkins Symptom Checklist; MDE = Major Depressive Episode; GAD = Generalized Anxiety Disorder; PPV = Positive Predictive Value; NPV = Negative Predictive Value; AUC = Area Under the ROC Curve.

For the HSCL-25, the optimal cut-off values for identifying cases of MDE or GAD were 1.96 for males and 2.20 for females. There was a good balance between sensitivity (0.92 for males and 0.89 for females) and specificity (0.83 for both males and females), and the Youden Index was acceptable for both males (0.74) and females (0.72). The PPV and NPV for males were 0.45 and 0.99, while the corresponding numbers were 0.62 and 0.96 for females (see Table 4 and Table 5).

For the HSCL-10, the best cut-off values to identify cases of MDE or GAD were 2.10 for males and 2.30 for females. The balance between sensitivity (0.91 for males and 0.88 for females) and specificity (0.81 for both sexes) was good. The Youden Index was satisfactory for both males (0.72) and females (0.69). The PPV and NPV for males registered at 0.42 and 0.98, respectively, and for females, these values were 0.58 and 0.96.

For the HSCL-5, the optimal cut-off values to identify cases of MDE or GAD were 2.25 for males and 2.75 for females, and similar to the two longer HSCL iterations, the balance between sensitivity and specificity was notably good (see Tables 4 and 5 for details). The Youden Index showed satisfactory results for both males and females at 0.68, and the PPV and NPV for males stood at 0.41 and 0.97, respectively, whereas for females, these figures were 0.65 and 0.94.

Table 5

Optimal Cut-Off Values for Males for HSCL-25, HSCL-10, and HSCL-5 for Major Depressive Episode (MDE) and Generalized Anxiety Disorder (GAD) Assessed by CIDI

HSCL version and CIDI diagnosis	Cut-off value	Youden Index	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
HSCL-25								
MDE or GAD	≥ 1.96	0.74	0.84	0.92	0.83	0.45	0.99	0.94
MDE	≥ 2.00	0.76	0.84	0.94	0.82	0.37	0.99	0.95
GAD	≥ 1.96	0.69	0.80	0.91	0.79	0.27	0.99	0.92
HSCL-10								
MDE or GAD	≥ 2.10	0.72	0.82	0.91	0.81	0.42	0.98	0.94
MDE	≥ 2.30	0.74	0.87	0.87	0.87	0.41	0.98	0.95
GAD	≥ 2.30	0.68	0.85	0.83	0.85	0.32	0.98	0.91
HSCL-5								
MDE or GAD	≥ 2.25	0.68	0.83	0.85	0.82	0.42	0.97	0.92
MDE	≥ 2.25	0.68	0.81	0.88	0.80	0.32	0.98	0.92
GAD	≥ 2.20	0.64	0.74	0.92	0.72	0.22	0.99	0.91

Note. CIDI = Composite International Diagnostic Interview; HSCL = Hopkins Symptom Checklist; MDE = Major Depressive Episode; GAD = Generalized Anxiety Disorder; PPV = Positive Predictive Value; NPV = Negative Predictive Value; AUC = Area Under the ROC Curve.

The corresponding values for only MDE and only GAD were relatively similar to those of MDE or GAD. The same optimal cut-off values for all HSCL iterations were replicated in bootstrapped analyses with 1,000 runs.

As also detailed in Table 6, when using the optimal cut-offs, all three HSCL iterations were associated with a marked overestimation of the prevalences, according to the self-administered CIDI. For example, the 30-day prevalence rates of MDE or GAD were 13.1% for males and 23.8% for females. However, using the optimal HSCL-25 cut-offs, these rates increased to 26.9% for males and 33.8% for females, respectively.

Discussion

The present large-scale study of students in higher education employed a recently adapted self-administered psychiatric diagnostic survey (CIDI 5.0) to investigate the efficacy of three iterations of the widely used HSCL scale in detecting cases of GAD or MDE, with an emphasis on potential sex-specific cut-off values. Our results show that all three versions of the HSCL discern relatively well between students afflicted with and without generalized anxiety disorder (GAD) or major depressive episode (MDE), and that different cut-offs for males and females should be used to ensure a good balance between sensitivity, specificity and overall accuracy.

Table 6

Prevalence of Mental Disorder According to CIDI and the Optimal HSCL Cutoffs

	Males		Females	
HSCL version and CIDI diagnosis	**CIDI**	**HSCL[a]**	**CIDI**	**HSCL[a]**
HSCL-25				
MDE or GAD	13.1%	26.9%	23.8%	33.8%
MDE	9.8%	25.1%	16.9%	30.2%
GAD	8.0%	26.9%	15.7%	39.3%
HSCL-10				
MDE or GAD	13.1%	28.3%	23.8%	35.2%
MDE	9.8%	20.6%	16.9%	27.2%
GAD	8.0%	20.6%	15.7%	35.2%
HSCL-5				
MDE or GAD	13.1%	26.6%	23.8%	28.7%
MDE	9.8%	26.6%	16.9%	28.7%
GAD	8.0%	33.3%	15.7%	28.6%

Note. CIDI = Composite International Diagnostic Interview; HSCL = Hopkins Symptom Checklist; MDE = Major Depressive Episode; GAD = Generalized Anxiety Disorder.
[a]Estimated prevalences based on the optimal HSCL cut-off.

A significant finding in this study is the introduction of new, sex-specific cut-off values for all HSCL iterations. These values diverge from those in earlier validation studies, which might have lacked the statistical power to evaluate distinct cut-offs for males and females. While the conventional cut-off value of 1.75 for the HSCL-25 has been consistently used for both sexes for years, the present study indicates that adopting slightly elevated cut-offs enhances the balance between sensitivity and specificity for both males and females across all HSCL iterations. Prior research contrasting the full HSCL-25 with structured diagnostic interviews have demonstrated a sensitivity ranging from 70-88% and a specificity between 77-85% for mood disorders or depression, and 43-50% sensitivity with 83% specificity for anxiety disorders (Rodríguez-Barragán et al., 2021; Sandanger et al., 1998; Veijola et al., 2003). In the current study, both the sensitivity and specificity were generally a little higher both for MDE (sensitivity 92-94%, and specificity 82%), and GAD (sensitivity 91%, and specificity 73-79%). It is important to note that the different studies employ different anxiety and depression diagnoses.

What is noteworthy is that both the shorter HSCL-10 and HSCL-5 displayed similarly high levels of sensitivity and specificity. Although very few studies have investigated the case-detection ability of these shorter HSCL iterations, a recent study by Rodríguez-Barragán et al. found that the HSCL-5 yielded a sensitivity of 78% and a specificity of 73% for depression. Based on our dataset using the newly suggested cut-offs, we observed an even higher sensitivity and specificity for both MDE and GAD, suggesting that both

the HSCL-10 the HSCL-5 may be equally good alternatives to be used in epidemiological research for the purpose of detecting probable cases of depression and anxiety. However, it should be noted that the relative differences between HSCL and CIDI prevalences were more pronounced in men. This was driven by differences in base prevalence as measured by CIDI; specifically, a relatively lower CIDI prevalence estimate is, all other things being equal (ceteris paribus), associated with a higher HSCL prevalence estimate. Another explanation for the disparities between HSCL and CIDI prevalence is that HSCL does not assess how symptoms affect daily functioning, while this is a key requirement for diagnosing a mental disorder in the CIDI instrument. This difference might partly explain why specificity was somewhat lower for GAD than for MDE. Anxiety disorders often have a more intricate relationship with functional impairment compared to depression (McKnight et al., 2016). Future studies could explore improving the HSCL by adding a measure of daily functioning. This addition could address the tool's current limitation of potentially overestimating mental health disorders due to the lack of assessment of the real-world functional impact. A revised HSCL with functional impairment questions could offer a more thorough evaluation. Alternatively, it could be used initially for screening, followed by more detailed assessments of daily functioning in those with elevated scores, ensuring a balanced approach that combines ease of use with comprehensive symptom analysis. Integrating functional impairment assessment would thus significantly enhance HSCL's diagnostic accuracy.

Surprisingly, few recent studies, both generally and specifically on university samples, have used structured diagnostic interviews, with the latest being over a decade old (Kou et al., 2012; Verger et al., 2010). The Dutch NEMESIS-3 study (ten Have et al., 2023), which monitors mental disorders in the Dutch general population, offers valuable insights. Using CIDI 3.0 (face-to-face interviews), the 12-month prevalence of any mental disorder was found to be 40% among young adults aged 18-24 years and 35% among those aged 25-35 years. Lifetime prevalence estimates were 50% and 59%, respectively. A sub-study from the 2020 HUNT study also provides relevant data (Knudsen et al., 2021). In this study, 2,154 participants from the general population were interviewed using CIDI 5.0. The 30-day prevalence of mental disorders among those aged 20-29 years was estimated to be 25.5% just before the COVID-19 pandemic. Although the prevalence estimates from NEMESIS-3 and HUNT are lower than those in our study of college and university students, they highlight the high prevalence of mental disorders among young adults.

Some methodological considerations warrant attention. This study's reliance on the standardized and validated CIDI psychiatric survey represents a significant strength. However, the shift from traditional face-to-face interviews to a self-administered electronic format in CIDI 5.0 introduces challenges, such as the need for further validation against conventional methods. Past research (Knudsen et al., 2021) indicates no significant prevalence differences between face-to-face and telephone interviews, although

recent comparisons between face-to-face and web-based self-reporting of psychological functioning revealed that respondents in face-to-face settings reported slightly fewer symptoms of depression (Cohen's $d = 0.25$) (Kocjan et al., 2023). This raises questions about the accuracy of different administration modes, particularly as previous findings suggest that young, well-educated respondents might underreport mental health issues. Assessing the reliability of the self-administered CIDI, especially in comparison to face-to-face interviews, is crucial not only for validation but also for understanding how different modes might impact mental health assessments. However, discrepancies between HSCL and CIDI prevalence estimates, particularly among male respondents, require careful interpretation. These variations underscore the need for contextual adjustments and a deeper understanding of how assessment tools could produce divergent outcomes in mental health research. The proposed cut-offs, tailored for epidemiological research in a student population with unique demographic traits, might not generalize well to other groups due to varying baseline prevalences and cultural perceptions of mental health. Therefore, while HSCL shows promise as a screening tool in a student population, its broader applicability needs further validation. Moreover, the limited differences observed in mental health problems between CIDI respondents and non-respondents, along with consistent response rates across Norwegian regions, suggest a reasonable level geographic representativeness. However, our understanding of non-responders, limited to basic demographic details, restricts our ability to fully assess representativeness against the broader Norwegian student population. Recent findings from Denmark (Lyngsøe et al., 2023), showing minor participation variations across sociodemographic groups provide some reassurance regarding the generalizability of our results.

The best approaches to determining cut-off values based on scales, like the various HSCL variants, depend on the intended purpose. If the goal is to identify individuals potentially in need of psychiatric treatment, prioritizing high sensitivity is crucial. However, when estimating the prevalence of depression or anxiety disorders within a population, striking a balance between specificity and sensitivity relative to the true prevalence of the disorder might yield more accurate results. Recommending specific cut-off values for future studies lies beyond the scope of this study. Such norms can be effectively determined only after conducting several similar studies, ideally across diverse cultures, countries, and population segments. The benefits of using different cut-off points for men and women also warrant more comprehensive evaluation, compared against the challenges posed by adding such complexity.

Funding: Norwegian Ministry of Education and Research.

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: The authors have declared that no competing interests exist.

Ethics Statement: The study was approved by the Regional Committee for Medical and Health Research Ethics in Western Norway (no. 2022/326437). All subjects gave electronic informed consent in accordance with the Declaration of Helsinki (2013).

Preregistration: The study was not preregistered. The CIDI follow-up survey is registered at ClinicalTrials.gov (identifier: NCT05731102).

Reporting Guidelines: We adhered to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines in the design, conduct, and reporting of this study.

Statement of the Conclusions: The study revealed high criterion validity for all HSCL versions in detecting major depressive episode and generalized anxiety disorder, suggesting their efficacy as screening tools rather than for precise prevalence estimation.

Data Availability: Data are available upon reasonable request. All SHoT data set is administrated by the NIPH. Approval from a Norwegian regional committee for medical and health research ethics (https://helseforskning.etikkom.no) is a prerequisite.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Auerbach, R. P., Alonso, J., Axinn, W. G., Cuijpers, P., Ebert, D. D., Green, J. G., Hwang, I., Kessler, R. C., Liu, H., Mortier, P., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Aguilar-Gaxiola, S., Al-Hamzawi, A., Andrade, L. H., Benjet, C., Caldas-de-Almeida, J. M., Demyttenaere, K., . . . Bruffaerts, R. (2016). Mental disorders among college students in the World Health Organization World Mental Health Surveys. *Psychological Medicine, 46*(14), 2955–2970. https://doi.org/10.1017/S0033291716001665

Chi, T., Cheng, L. Y., & Zhang, Z. J. (2023). Global prevalence and trend of anxiety among graduate students: A systematic review and meta-analysis. *Brain and Behavior, 13*(4), Article e2909. https://doi.org/10.1002/brb3.2909

Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science, 19*(1), 1–15. https://doi.org/10.1002/bs.3830190102

Glaesmer, H., Braehler, E., Grande, G., Hinz, A., Petermann, F., & Romppel, M. (2014). The German Version of the Hopkins Symptoms Checklist-25 (HSCL-25) – Factorial structure, psychometric properties, and population-based norms. *Comprehensive Psychiatry, 55*(2), 396–403. https://doi.org/10.1016/j.comppsych.2013.08.020

Haro, J. M., Arbabzadeh-Bouchez, S., Brugha, T. S., de Girolamo, G., Guyer, M. E., Jin, R., Lepine, J. P., Mazzi, F., Reneses, B., Vilagut, G., Sampson, N. A., & Kessler, R. C. (2006). Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys. *International Journal of Methods in Psychiatric Research, 15*(4), 167–180. https://doi.org/10.1002/mpr.196

Ibrahim, A. K., Kelly, S. J., Adams, C. E., & Glazebrook, C. (2013). A systematic review of studies of depression prevalence in university students. *Journal of Psychiatric Research, 47*(3), 391–400. https://doi.org/10.1016/j.jpsychires.2012.11.015

Jordanova, V., Wickramesinghe, C., Gerada, C., & Prince, M. (2004). Validation of two survey diagnostic interviews among primary care attendees: A comparison of CIS-R and CIDI with SCAN ICD-10 diagnostic categories. *Psychological Medicine, 34*(6), 1013–1024. https://doi.org/10.1017/S0033291703001727

Kessler, R. C., & Üstün, T. B. (2004). The World Mental Health (WMH) Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research, 13*(2), 93–121. https://doi.org/10.1002/mpr.168

Kirkøen, B., Ørstavik, R., Reneflot, A., Skogen, J. C., Sivertsen, B., & Knudsen, A. K. S. (2024). *The ability of the Hopkins Symptoms Checklist-5 to identify anxiety and depression in the general population* [Manuscript in preparation].

Knapstad, M., Sivertsen, B., Knudsen, A. K., Smith, O. R. F., Aaro, L. E., Lonning, K. J., & Skogen, J. C. (2021). Trends in self-reported psychological distress among college and university students from 2010 to 2018. *Psychological Medicine, 51*(3), 470–478. https://doi.org/10.1017/S0033291719003350

Knudsen, A. K. S., Stene-Larsen, K., Gustavson, K., Hotopf, M., Kessler, R. C., Krokstad, S., Skogen, J. C., Øverland, S., & Reneflot, A. (2021). Prevalence of mental disorders, suicidal ideation and suicides in the general population before and during the COVID-19 pandemic in Norway: A population-based repeated cross-sectional analysis. *The Lancet Regional Health – Europe, 4*, Article 100071. https://doi.org/10.1016/j.lanepe.2021.100071

Kocjan, G. Z., Lavtar, D., & Sočan, G. (2023). The effects of survey mode on self-reported psychological functioning: Measurement invariance and latent mean comparison across face-to-face and web modes. *Behavior Research Methods, 55*(3), 1226–1243. https://doi.org/10.3758/s13428-022-01867-8

Kou, C., Meng, X. F., Xie, B., Chen, Y. F., Yu, Q., Shi, J. P., Yu, Y. Q., D'Arcy, C., & Huang, Y. Q. (2012). The prevalence and correlates of neurotic disorders among undergraduates at a mainland Chinese university. *Social Psychiatry and Psychiatric Epidemiology, 47*(12), 2011–2018. https://doi.org/10.1007/s00127-012-0500-2

Lyngsøe, S., Lophaven, S., Jepsen, R., Holmager, T., Janssens, A., & Lynge, E. (2023). Non-participation in a health examination survey in a rural-provincial area of Denmark – Results from the Lolland-Falster Health Study (LOFUS). *Scandinavian Journal of Public Health*. Advance online publication. https://doi.org/10.1177/14034948231206879

McCloud, T., Kamenov, S., Callender, C., Lewis, G., & Lewis, G. (2023). The association between higher education attendance and common mental health problems among young people in England: Evidence from two population-based cohorts. *The Lancet Public Health, 8*(10), e811–e819. https://doi.org/10.1016/S2468-2667(23)00188-3

McKnight, P. E., Monfort, S. S., Kashdan, T. B., Blalock, D. V., & Calton, J. M. (2016). Anxiety symptoms and functional impairment: A systematic review of the correlation between the two measures. *Clinical Psychology Review, 45*, 115–130. https://doi.org/10.1016/j.cpr.2015.10.005

Nordgaard, J., Sass, L. A., & Parnas, J. (2013). The psychiatric interview: Validity, structure, and subjectivity. *European Archives of Psychiatry and Clinical Neuroscience, 263*(4), 353–364. https://doi.org/10.1007/s00406-012-0366-z

Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., & Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research, 18*(3), 169–184. https://doi.org/10.1002/mpr.289

Rodríguez-Barragán, M., Fernández-San-Martín, M. I., Clavería, A., Le Reste, J. Y., Nabbe, P., Motrico, E., Gómez-Gómez, I., & Peguero-Rodríguez, E. (2023). Measuring depression in Primary Health Care in Spain: Psychometric properties and diagnostic accuracy of HSCL-5 and HSCL-10. *Frontiers in Medicine, 9*, Article 1014340. https://doi.org/10.3389/fmed.2022.1014340

Rodríguez-Barragán, M., Fernández-San-Martín, M. I., Clavería-Fontán, A., Aldecoa-Landesa, S., Casajuana-Closas, M., Llobera, J., Oliván-Blázquez, B., & Peguero-Rodríguez, E. (2021). Validation and Psychometric Properties of the Spanish Version of the Hopkins Symptom Checklist-25 Scale for Depression Detection in Primary Care. *International Journal of Environmental Research and Public Health, 18*(15), Article 7843. https://doi.org/10.3390/ijerph18157843

Sandanger, I., Moum, T., Ingebrigtsen, G., Dalgard, O. S., Sørensen, T., & Bruusgaard, D. (1998). Concordance between symptom screening and diagnostic procedure: The Hopkins Symptom Checklist-25 and the Composite International Diagnostic Interview I. *Social Psychiatry and Psychiatric Epidemiology, 33*(7), 345–354. https://doi.org/10.1007/s001270050064

Sheldon, E., Simmonds-Buckley, M., Bone, C., Mascarenhas, T., Chan, N., Wincott, M., Gleeson, H., Sow, K., Hind, D., & Barkham, M. (2021). Prevalence and risk factors for mental health problems in university undergraduate students: A systematic review with meta-analysis. *Journal of Affective Disorders, 287*, 282–292. https://doi.org/10.1016/j.jad.2021.03.054

Sivertsen, B., Knudsen, A. K. S., Kirkøen, B., Skogen, J. C., Lagerstrøm, B. O., Lønning, K.-J., Kessler, R. C., & Reneflot, A. (2023). Prevalence of mental disorders among Norwegian college and university students: A population-based cross-sectional analysis. *The Lancet Regional Health – Europe, 34*, Article 100732. https://doi.org/10.1016/j.lanepe.2023.100732

Sivertsen, B., Råkil, H., Munkvik, E., & Lonning, K. J. (2019). Cohort profile: The SHoT-study, a national health and well-being survey of Norwegian university students. *BMJ Open, 9*(1), Article e025200. https://doi.org/10.1136/bmjopen-2018-025200

Skogen, J. C., Overland, S., Smith, O. R. F., & Aaro, L. E. (2017). The factor structure of the Hopkins Symptoms Checklist (HSCL-25) in a student population: A cautionary tale. *Scandinavian Journal of Public Health, 45*(4), 357–365. https://doi.org/10.1177/1403494817700287

ten Have, M., Tuithof, M., van Dorsselaer, S., Schouten, F., Luik, A. I., & de Graaf, R. (2023). Prevalence and trends of common mental disorders from 2007-2009 to 2019-2022: Results from the Netherlands Mental Health Survey and Incidence Studies (NEMESIS), including comparison of prevalence rates before vs. during the COVID-19 pandemic. *World Psychiatry, 22*(2), 275–285. https://doi.org/10.1002/wps.21087

Veijola, J., Jokelainen, J., Läksy, K., Kantojärvi, L., Kokkonen, P., Järvelin, M. R., & Joukamaa, M. (2003). The Hopkins Symptom Checklist-25 in screening DSM-III-R axis-I disorders. *Nordic Journal of Psychiatry, 57*(2), 119–123. https://doi.org/10.1080/08039480310000941

Verger, P., Guagliardo, V., Gilbert, F., Rouillon, F., & Kovess-Masfety, V. (2010). Psychiatric disorders in students in six French universities: 12-month prevalence, comorbidity, impairment and help-seeking. *Social Psychiatry and Psychiatric Epidemiology, 45*(2), 189–199. https://doi.org/10.1007/s00127-009-0055-z

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

Examination of Gender Differences: Causal Attributions of Treatment-Seeking Individuals With Overweight and Obesity

Carmen Henning[1], Caroline Seiferth[1,2], Tanja Färber[1], Magdalena Pape[2], Stephan Herpertz[3], Sabine Steins-Loeber[2], Jörg Wolstein[1]

[1] *Department of Psychopathology, University of Bamberg, Bamberg, Germany.* **[2]** *Department of Clinical Psychology and Psychotherapy, University of Bamberg, Bamberg, Germany.* **[3]** *Department of Psychosomatic Medicine and Psychotherapy, LWL-University Hospital, Ruhr University Bochum, Bochum, Germany.*

Clinical Psychology in Europe, 2024, Vol. 6(4), Article e12089, https://doi.org/10.32872/cpe.12089

Received: 2023-05-31 • **Accepted:** 2024-08-14 • **Published (VoR):** 2024-12-20

Handling Editor: Cornelia Weise, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Corresponding Author: Carmen Henning, University of Bamberg, Markusstr. 8a, 96047 Bamberg, Germany. Phone +49 951 863 2589. E-mail: carmen.henning@uni-bamberg.de

Supplementary Materials: Materials, Preregistration [see Index of Supplementary Materials]

Abstract

Background: Addressing patients' perceptions of the causes of their overweight and obesity may be a promising approach to enhance treatment motivation and success. Previous research suggests that there are gender differences in these aspects. The objective of this study was to investigate gender differences in causal attributions among individuals with overweight and obesity who participated in a cognitive-behavioral mobile health (mHealth) intervention.

Method: Causal attributions were assessed using the revised Illness Perceptions Questionnaire, which included a rated and open answering section. An ANCOVA was conducted for each causal factor (behavioral, psychological, risk, external) as a dependent variable to determine gender differences, which were analysed with chi-squared tests for open-ended responses.

Results: The most frequently mentioned and highly rated cause was behavior for both genders (59.8% of 639 responses). The results indicated that women rated psychological causes, particularly stress-related causes, significantly higher, $F(1,211) = 14.88$, $p < .001$, $\eta^2 = .07$, and were more likely to cite emotional eating than men, $\chi^2(1, N = 639) = 15.06$, $p < .001$. Men rated alcohol stronger as cause than women, $t(125.05) = 3.79$, $p < .001$.

Conclusion: The findings of this study contribute to the understanding of the gender differences in causal attributions among individuals with overweight or obesity. Implementing stress

management interventions with a focus on emotion regulation is pivotal, especially for females. Interventions should focus on sensitizing males to the association between emotions and eating behavior. The causal attributions should be assessed with different survey methods in order to match the patient's view of their condition.

Keywords

overweight, obesity, gender, causal attributions, physical activity, health behavior, mHealth

Highlights

- Causal attributions of overweight and obesity differ between males and females.
- Female participants attributed their overweight/obesity primarily to psychological causes.
- Awareness of the association between emotion and overweight or obesity for males is necessary.
- For both genders, behavioral aspects and enhancing stress management should be focused on.

Nutrition and exercise programs for individuals with overweight or obesity (OO) are widely available, but the third pillar of evidence-based treatment, cognitive-behavioral interventions, is difficult to obtain at low-threshold. Health insurance companies cover the costs of mobile Health (mHealth) interventions in some countries (Roth et al., 2023), which can adequately bridge the long waiting times for specialized in-person treatment. Therefore, identifying the underlying mechanism of individuals with OO to engage with mHealth interventions that address cognitive-behavioral aspects of weight loss and weight-gain prevention is important.

The perception of causes among individuals with overweight (Body Mass Index, BMI = 25 – 29.99 kg/m^2) or obesity (BMI ≥ 30 kg/m^2) differs based on socio-cultural factors and self-perceived consequences. The media reinforces the ideal of thin women and the necessity of dieting for females (Pedersen, 2010), which can lead to an internalization of a thin beauty ideal and social comparisons (López-Guimerà et al., 2010). Women who have internalized the thin beauty ideal or show a high exposure to such media tend to report greater body dissatisfaction, unhealthy eating (López-Guimerà et al., 2010), and unrealistic weight goals (Dutton et al., 2010). In contrast, men seem to be less concerned about their OO and less aware of the consequences than females (Breland et al., 2023; Mozumdar & Liguori, 2011; Tronieri et al., 2017). Studies indicate that some males with OO do not perceive themselves as OO, whereas females with normal weight perceive themselves as overweight or obese (Chang & Christakis, 2003). Several gender differences in OO have been reported: overall, 53.5% of the German population is affected by overweight, including obesity, with a clear gender difference of 60.5% men and 46.6% women (Schienkiewitz et al., 2022). The prevalence of obesity is positively correlated

with age and negatively correlated with socio-economic status (Schienkiewitz et al., 2022). Women are more likely to be affected by food craving (Hallam et al., 2016) and emotional eating behavior, i.e., overeating when experiencing (negative) emotions, than men, whereby this overeating reinforces negative emotions and can create a vicious circle (Breland et al., 2023). Research indicates that there is a higher prevalence of weight loss intentions among females than males (Houle-Johnson & Kakinami, 2018). Both genders are motivated to lose weight to improve overall health, but women also tend to report more internal motivators, such as increased personal esteem (Crane et al., 2017), while men tend to be more motivated by external factors, such as improved job performance (Sabinsky et al., 2007). In general, males are under-represented in obesity research, which often leads to difficulties in the transfer of research findings (Bramlage et al., 2004; Cooper et al., 2021; Pantalone et al., 2017).

Furthermore, research has demonstrated that women tend to associate obesity with more negative emotions and worse illness perceptions than men (Henning et al., 2022), and that this mental image is negatively associated with dieting attempts and weight cycling (Prill et al., 2021). Gender differences have also been found in the assumptions about the causes of an illness, the so-called causal attributions of one's own OO. These causal attributions have direct effects on therapeutic outcomes, coping, and goal-related behavior (Mathieu et al., 2018; Zhang et al., 2018). These causal attributions can be categorized into different factors, e.g., psychological or genetic. The structure varies depending on the disease and its aetiology, whether it is multifactorial or can be attributed to a specific trigger (e.g., hereditary in the case of trisomy 21). For OO, which is a multifactorial disease, no unique structure has been identified (Daigle et al., 2019). Recent literature offers contradictory or non-comparable findings about the causal attributions of individuals with obesity, and most of the studies report no gender specific results. A cohort study with 75 individuals with OO suggested that unfavourable health behavior (e.g., excessive eating) was the most often causal attribution (58.7%) of own obesity, but individuals also considered psychological causes (e.g., worries) (Mathieu et al., 2018). Strong behavioral attributions (e.g., sedentary behavior) were also found in an investigation of individuals seeking surgical or behavioral/pharmacological weight loss treatment (Pearl et al., 2018). Agüera and colleagues (2021) categorized causal attributions, particularly for individuals with eating disorders, into four distinct categories: eating disorder-specific, psychological, risk, and external causes. The psychological factor included self-reported own behavior, but not eating behavior. This was categorized within the domain eating disorder specific causal factor, which makes it difficult to compare the results with other studies. Studies show that most of the individuals with OO named psychological causes, such as emotions, boredom, and low self-worth, followed by lifestyle aspects such as working environment (Agüera et al., 2021; Brogan & Hevey, 2009). Other causal attributions contained childhood experiences, social environment, medical reasons, eating behavior, and media influence. Brogan and Hevey (2009) conducted a network analysis,

which showed that trauma, family problems, and an "addictive personality" were distal causes for overeating and comfort eating. Passive behavior, reduced physical activity levels, overeating, and comfort eating were proximal causes for obesity (Brogan & Hevey, 2009). To date, the majority of studies have not analysed results by gender. Consequently, the investigation of patterns of gender disparities with regard to causal attributions is underrepresented.

Several studies have examined the link between BMI and causal attributions of OO, but the results have been inconsistent. Lewis and colleagues (2010) suggested that the attribution of personal responsibility as a cause for obesity leads to powerlessness of the individuals with obesity grade III (BMI ≥ 40kg/m^2) and to empowerment of individuals with lower BMI, whereas another study found an association with age but not the BMI level (Mathieu et al., 2018). Individuals with OO showed stronger attributions to heritability with their weight than normal weight individuals, which has been suggested to be associated with lower physical activity, decreased self-efficacy, and a low perception of personal control (Hilbert et al., 2009; Wang & Coups, 2010). However, their assumptions that their obesity was caused by overeating could have led to greater reported levels of physical activity (Wang & Coups, 2010).

The associations between causal attributions of OO and treatment outcomes or health behavior have been investigated by some studies. Individuals with OO showed more negative health outcomes as well as emotional and disinhibited eating behavior when they assumed psychosocial causes of their obesity (Mathieu et al., 2018). Psychosocial attributions were associated with pathologic eating patterns, which was more often prevalent in females (Mathieu et al., 2018). Research showed that interventions that match individuals' causal assumptions of their illness can be a strategy to individualize treatment in OO and lead to better weight loss results (Bauer et al., 2020; Broadbent et al., 2009; Karekla et al., 2019).

Causal attributions are modifiable, disease and gender specific, and could lead to a change of health behavior (Bonsaksen et al., 2015; Surgenor et al., 2020; Zhang et al., 2018). A gender-sensitive investigation about causal attributions of individuals with OO, who are motivated to lose weight and interested in using mHealth for weight loss is lacking. The results could give an insight in underlying mechanisms and help enhance mHealth interventions for men and women. The aim of the present study was to examine these gender differences in this group. Given the contradictory or non-existent findings in the literature, we did not have directional hypotheses about gender-specific differences.

Materials and Method

Design

This cross-sectional study was part of the I-GENDO project, which was approved by the ethics committee of the University of Bamberg, Germany and the Institutional Review Board of the Ruhr-University Bochum (no. 18-6415) (Pape et al., 2022). The study was conducted in accordance with the Declaration of Helsinki. The participants provided their informed consent to participate in this study. Data collection took place via an online questionnaire between December 2019 and August 2020 within the pre-screening for the I-GENDO project. The aim of the project was the development and evaluation of a gender-sensitive mHealth intervention with psychological contents for weight loss and self-tailoring elements (Pape et al., 2022). After a telephone interview, individuals with suicidality or binge eating disorder were excluded. To avoid a systematic selection effect of a pseudo-random sample, we targeted especially males via press releases. Consequently, the sample is disproportionately stratified concerning gender, given that the proportion of males is still less than in the population.

Sample

The study included 675 interested participants who were informed about the content of the project and screened for eligibility. The inclusion criteria of the participants were having overweight or obesity grade I and II (BMI = 25.00 – 39.9 kg/m^2), being motivated to lose weight and interested in using an mHealth application, at least 18 years old, not pregnant, and having no binge eating disorder or bulimia nervosa according to DSM-5-criteria (American Psychiatric Association, 2020) (see Henning et al., 2024S, Additional File 1 for recruitment process). Individuals with a BMI greater than 40 kg/m^2 often have comorbidities and drug therapy or bariatric surgery is advised (Deutsche Adipositas-Gesellschaft e.V., 2014). Consequently, they were excluded from the present study.

The final sample compromised 213 participants (female: 143; male: 70) between 19 and 71 years old. Power analyses were conducted using G*Power version 3.1.9.7 (Faul et al., 2007) and resulted in a sample size of 210 participants required to achieve 80% power for detecting a medium effect. More than half of the participants were married or in a partnership (55.4%), almost a third were single (31.4%) and 13.2% were divorced or widowed. Males and females did not differ in BMI (*Min:* 25.59kg/m^2; *Max:* 39.88kg/m^2), age or education (see Table 1).

Table 1

Sample Characteristics

Variable	Total (n = 213)	Females (n = 143)	Males (n = 70)	Group differences
BMI M (SD) (kg/m^2)	33.35 (3.79)	33.51 (3.71)	33.01 (3.95)	$t(211) = 0.301$; $p = .360$
Age M (SD) (years)	46.45 (12.13)	44.94 (12.58)	49.51 (10.59)	$t(160.05) = -2.78$; $p = .006$
Level of Education (%)				$\chi^2(2,213) = 2.14$; $p = .343$
Low	13.62	13.29	14.28	
Middle	24.41	27.97	17.14	
High	59.62	57.34	64.29	

Note. Significance level $p < .001$.

Instruments

We assessed the demographic variables such as age and gender at the beginning of the questionnaire. The causal attributions were assessed with the IPQ-R (Moss-Morris et al., 2002). First, 19 potential causes of OO (e.g., "stress or worries") were presented and participants were asked to rate the extent of personal agreement with each cause on a 5-point Likert scale (1 = *strongly disagree*; 5 = *strongly agree*). Subsequently, participants were requested to name three causes that are most relevant to them personally in an open answering form.

Statistical Analysis

Rating of the Causes — All analyses were conducted using IBM SPSS (Version 26). A four-factor model for causal attributions was set based on Moss-Morris and colleagues (2002), which was adapted for OO in accordance with the recommendations for this questionnaire (see Figure 1): psychological (6 items, Cronbach's α = .746), behavioral (2 items, α = .750), risk (6 items, α = .413), and external (5 items, α = .646) factor. The significance level was set at $p < .05$ and was maintained through a Bonferroni correction for multiple testing ($p < .001$).

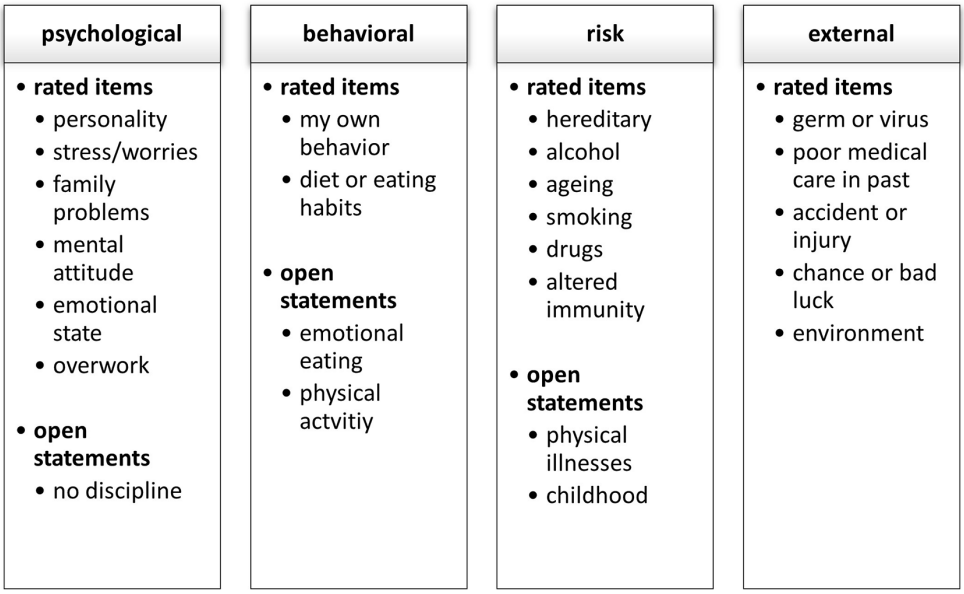
For each of the four causal factors a one-way ANCOVA was computed to analyse gender differences because the assumption of homogeneity of the regression slopes of gender for a MANCOVA was not met. We controlled for BMI and age in the first step and added gender as an independent variable in the second step. Additionally, we conducted two-sided t-tests for each item of the rated section.

The assumptions for ANCOVAs were checked: Homogeneity of regression slopes was not violated for three of the four dependent variables: behavioral, psychological, and risk factors ($p < .0125$). This assumption was not met for the external factor, as indicated by the significant interaction term for gender and age ($p = .001$). Consequently, we omitted age as a covariate in the ANCOVA for the external factor. The residuals were normally

distributed for the psychological and risk factors as determined by the Shapiro-Wilk test ($p > .05$). However, the Shapiro-Wilk test was significant for the behavioral and external factors. The Kolmogorov Smirnov test was not significant ($p > .0125$) and because of the sample size, we omitted bootstrapping in the analysis. The assumptions of homogeneity of variances were not violated (Levene's test: $ps = .339 - .804$). The leverage values (< .200) and values for Cook's distance (< 1) indicated no outliers to be removed.

Figure 1

The 4-Factor Model of Causal Attributions



Open Statements — The open statements ($n = 639$) of the second questionnaire section were scalable, structured through deductive categorization according to Mayring (2015, p. 68) by two independent raters. The categorization was based on the 4-factor-model of the rated items with further additions (see Figure 1). For the 639 open statements, the degree of agreement by kappa was .946, which is an almost perfect interrater reliability (Landis & Koch, 1977). The analysis of the differences in frequencies for males and females were computed with chi-square tests. When expected cell frequencies were below five, we used the exact calculation option of SPSS.

Results

A significant difference was found in the first part of the questionnaire for the psychological factor, with women rating the items as more likely to cause their obesity than men. In addition to psychological causes, men exhibited significantly stronger beliefs that alcohol was a possible cause. There was no significant gender difference at the factor level in the open response format. However, women were significantly more likely to report emotional eating as a cause of their obesity. Table 2 presents the descriptive statistics of the four causal attribution factors (psychological, behavioral, external, and risk) for the total sample and for females and males separately as well as the results of the ANCOVAs of the rated items section. All factors were weakly significantly correlated with each other (r = .177 – .228, ps < .001), but not with BMI or age.

Table 2

Descriptive Statistics of Causal Attributions and Results of ANCOVAs for Gender

	Means (Standard Deviations)							
Factor	**Total**	**Females**	**Males**	**Females[a]**	**Males[a]**	***F*(1,211)**	***p***	**η^2**
PSY	3.23 (0.86)	3.40 (0.80)	2.88 (0.88)	3.39 (0.07)	2.91 (0.10)	14.883	< .001	.066
BEH	4.54 (0.52)	4.56 (0.47)	4.50 (0.60)	4.56 (0.04)	4.51 (0.06)	0.440	.508	.002
RIS	2.33 (0.53)	2.33 (0.53)	2.35 (0.53)	2.34 (0.05)	2.33 (0.06)	0.002	.966	0
EXT	1.74 (0.55)	1.70 (0.54)	1.81 (0.57)	1.71 (0.05)	1.80 (0.07)	1.492	.223	.007

Note. PSY = psychological; BEH = behavioral; RIS = risk; EXT = external factor.
[a]Adjusted for BMI at all factors and for age in the psychological, behavioral, and risk factor.

Psychological Causes

After controlling for age and BMI, a significant main effect of gender was found for the psychological causes. Females showed a stronger inclination towards psychological causes, with 'stress or worries' being the highest rated item on the scale (M = 4.15; SD = 0.80) compared to males (M = 3.71; SD = 0.95) (see Table 2). The approval rate for all items on the psychological scale (see Henning et al., 2024S, Additional File 2) was higher for females than for males. There was a significant effect of gender for the items 'stress/worries', $t(118.14) = -3.34$; p < .001, 'family problems', $t(211) = 3.20$; p < .001, and 'emotional state', $t(211) = -4.79$; p < .001 (see Henning et al., 2024S, Additional File 2).

Psychological causes were the second most frequently mentioned in the open-response format, with 25.4% for females and 23.8% for males. However, no significant gender difference was found (see Table 3). Although 'family problems' were rated significantly higher by females than males, men mentioned this cause more often than

females in the open response section (refer to Henning et al., 2024S, Additional File 2 and Table 3).

Table 3

Frequencies, Percentages, and Results of the $\chi^2(1; N = 639)$ Tests of the Open Statement Section

	Frequency			Percentages					
Factor / Subcategory	**Total**	**Females**	**Males**	**Total**	**Females**	**Males**	χ^2	p^a	*V*
Emotional state	42	34	8	6.6	7.9	3.8	3.880	.049	.078
Family problems	9	3	6	1.4	0.7	2.9	4.728	.030	.086
Stress/worries	61	41	20	9.5	9.6	9.5	0	.989	.001
Work stress	11	7	4	1.7	1.6	1.9	0.062	.803	.010
No discipline	36	24	12	5.6	5.6	5.7	0.004	.951	.002
Psychological total	159	109	50	24.9	25.4	23.8	0.193	.661	.017
Diet or eating habits	212	131	81	33.2	30.5	38.6	4.106	.043	.080
Emotional eating	45	42	3	7.0	9.8	1.4	15.058	< .001***	.154
Physical activity	117	68	49	18.3	15.9	23.3	5.277	.022	.091
My own behavior	3	2	1	0.5	0.5	0.5	0	.986	.001
Habits (e.g., sleeping)	5	4	1	0.8	0.9	0.5	0.378	.539	.024
Behavioral total	382	247	135	59.8	57.6	64.3	2.640	.104	.064
Alcohol	16	5	11	2.5	1.2	5.2	9.579	.002	.122
(physical) Illness	43	37	6	6.7	8.6	2.9	7.472	.006	.108
Hereditary/past	13	10	3	2.0	2.3	1.4	.576	.448	.030
Pregnancy in past	5	5	0	0.8	1.2	0	2.457	.116	.062
Risk total	77	57	20	12.1	13.3	9.5	1.884	.170	.054
Environment	6	4	2	0.9	0.9	1.0	0.001	.980	.001
External total	6	4	2	0.9	0.9	1.0	0.001	.980	.001
Others total	15	12	3	2.3	2.8	1.4	1.152	.283	.042

Note. *V* = Cramer´s *V*, effect sizes of χ^2 tests.
[a]Significance level due to Bonferroni correction $p < .001$ (***).

Behavioral Causes

The participants rated the behavioral factor as the most important cause (see Table 2) and causes related to their behavior were most frequently mentioned in the open response section, both by females (57.6%) and males (64.3%) (see Table 3). Although the Chi-square test for the behavioral factor was nonsignificant, 'emotional eating' ('eating because I'm bored/frustrated') was reported as a cause of their OO significantly more often by women than men (see Table 3).

Risk and External Causes

Neither the risk nor the external factor were in an area of agreement (see Table 2 for descriptive results and Henning et al., 2024S, Additional File 2 for single item agreement). Males rated 'alcohol' ($M = 2.64$; $SD = 1.24$) significantly higher than females, $M = 1.98$; $SD = 1.12$; $t(125.05) = 3.79$, $p < .001$ (see Henning et al., 2024S, Additional File 2) and reported

it as a possible cause more often (see Table 3). Causes of risk were reported more frequently by females (13.3%) than males (9.5%), mainly due to physical illness (64.9% of female vs. 30.0% of male responses in this category), but again the difference was not significant (see Table 3). The number of statements categorized as external factors was less than 1% for both genders.

Discussion

This study used an exploratory design to investigate the gender-related differences in self-perceived causes among individuals with OO. Participants were permitted to rate pre-defined causes and provide open-ended responses. All participants wanted to lose weight and participated in a project involving a behavioral-cognitive mHealth intervention. Responses were analysed at both the factor level (psychological, behavioral, external, and risk factors) and the item level.

In summary, significant gender differences were observed in the agreement with the psychological causes. Women considered stress, family problems and their emotional state to be significantly more important causes of their weight than men. Behavioral causes were rated most highly by both genders, with significantly more women than men citing emotional eating as a cause in the open-ended responses. The only cause for which gender differences were observed in both survey methods was alcohol consumption. This was rated significantly more strongly and cited more frequently by men.

The highest rated item on the psychological scale was 'stress/worries' for both genders, which emphasizes the importance of adaptive enhancing coping mechanisms in individuals with OO. Stress management training should be an integral part of psychological interventions, especially for females who showed significant higher scores on stress-related causal beliefs than males. Individuals with better coping strategies and competences to handle daily stresses are more successful in maintaining weight loss (Elfhag & Rössner, 2005).

The high rating of the importance of behavioral aspects such as eating and physical activity behavior is in line with results of other studies, which showed that they are proximal causes and causal attributions of obesity (Brogan & Hevey, 2009; Haslam & James, 2005; Mathieu et al., 2018; Pearl et al., 2018). Unfavourable health behavior such as emotional eating or physical inactivity seem to be maladaptive coping strategies for stress. Investigations of cardiac patients showed that individuals with beliefs in behavioral causes were more likely to change their dietary or exercise behavior (Weinman et al., 2000). Based on our results, which focused on a psychological mHealth weight-loss intervention, and existing research, it appears that motivating patients to address their OO could be effective by emphasizing the behavioral and changeable aspects of the condition.

This is supported by Fleary and Ettienne (2014) who found an association between the causal attribution of inactivity for males and their motivation to lose weight. Research has shown that males tend to benefit more than females from exercise in terms of weight loss and prefer this method instead of dieting and restrictive eating, which is perceived as a 'female approach' of weight management (Donnelly et al., 2003; Kiefer et al., 2005). Physical activity is not necessarily a prerequisite for weight-loss or maintenance because of compensatory behaviors and less discipline in attending sport programs regularly and on a long-term basis (Foright et al., 2018). Results about the effect of psychological intervention, such as behavioral change techniques on physical activity, are inconsistent (Awoke et al., 2022; Dombrowski et al., 2012). One possible approach could be to enhance self-efficacy by action planning, providing instruction and providing rewards to increase physical activity (Williams & French, 2011). The aim of psychological interventions in OO therapy could be to strengthen perseverance and reduce reward behavior related to food intake or alcohol consumption after exercise, particularly for men. The results indicate that men are aware of the role of diet and eating behavior in causing their overeating but are not aware of emotional eating (e.g., eating because of frustration). Interventions for males should focus on the association between emotions and overeating or alcohol consumption. Alcohol consumption seems to be a pivotal causal attribution of males, which is not surprising given that males drink more alcohol than females (Nolen-Hoeksema, 2004). It is recommended that men be made aware of the association between their own maladaptive coping, emotion regulation or self-rewarding behavior, which may manifest as alcohol consumption or eating, and their OO. Psychoeducational elements regarding the influence of alcohol on weight management in men should be considered in the development of mHealth interventions. In addition, self-monitoring of alcohol consumption may be useful for men, as this behavior change technique has been shown to be effective in interventions for physical activity and healthy eating (Samdal et al., 2017). Such a diary is easy to integrate into mHealth interventions but should be optional for the user or practitioner to activate, as alcohol consumption, especially in women, was not often reported as a suspected cause of OO.

It is suggested that emotional eating behavior for females should be focused on in interventions to target gender specific causal attributions. There is some evidence that females assume emotionally driven behavior such as emotional eating as a cause of their OO. Emotional eating has also been shown to be associated with the concept of food addiction (Pape et al., 2021) and eating addiction (Hebebrand & Gearhardt, 2021), to mediate the link between obesity, change in BMI and depression (Konttinen, Männistö, et al., 2010) and to be associated with less self-efficacy for the ability to maintain physical activity (Konttinen, Silventoinen, et al., 2010). Thus, emotional eating may represent a barrier to successful treatment. Our results are in accordance with other studies, which showed that females tend to engage in emotional eating behavior (Löffler et al., 2015).

Our findings are also consistent with earlier observations, which showed that females had a much more 'emotional view' on their obesity and showed significantly stronger emotional illness representations than males (Henning et al., 2022). This means that women associate their OO with anger and guilt. Combined with the significant effect of gender at the second highest rated psychological factor, which was significantly more pronounced for women, the evidence emphasizes the importance of emotion-focused therapy in OO (mHealth) interventions, especially for females.

However, the results of our study with the two different survey methods also suggest that while men recognise the psychological component of their illness, they do not see it as being as strongly responsible for OO as women do. This can be seen from the fact that the gender differences disappear almost completely in the open responses. This phenomenon may be due to the fact that women are more aware of obesity and its consequences, e.g., health consequences, and suffer more from it than men (Audureau et al., 2016; Breland et al., 2023). This greater awareness could also lead them to participate more in weight management programmes. We therefore recommend considering this aspect in questionnaires for men that measure the strength of the perception of causes and consequences and, if necessary, that the survey be optimised by adding open questions or interviews. To encourage men to participate in weight loss interventions or research projects, it may be beneficial to reduce the emphasis on the perceived threat associated with such initiatives in recruitment activities. Instead, it may be more effective to focus on the elements of behavior that can be changed.

In contrast to other studies (Daigle et al., 2019), we did not find an association between causal attributions and BMI level. Our results are also not consistent with previous findings that individuals with a BMI of less than 40 kg/m^2 believe in causes such as social aspects or environment (Daigle et al., 2019; Lewis et al., 2010). One potential explanation is that our treatment-seeking sample was motivated to lose weight. Consequently, they may have attributed their OO more often to changeable causes such as their behavior or coping mechanisms. The practical implication of this finding is that it is important to raise awareness of the impact of the environment, in order to enhance strategies to cope with these external stimuli. However, it is also important to emphasize their role, abilities, and potentials to meet these challenges, which in turn should enhance their self-efficacy to manage weight loss and maintenance.

Finally, a number of limitations need to be considered. The use of a cross-sectional design limits any causal conclusion. It is noteworthy that all respondents self-identified as either male or female, with no individuals selecting the "other" category. However, research in the domain of non-binary environments would be invaluable in order to facilitate the transfer of results to all individuals. Apart from these limitations, the generalizability of these results is limited because the sample consisted of individuals who were motivated to attend an mHealth study, which could have led to desirability effects in answering. The rated section excluded hedonistic items (e.g., eating because it

tastes good), physical diseases, and physical activity, which might have led to a priming effect or bias in responding to the open-statement section. As with other studies, the reliability of the factors is low (Daigle et al., 2019), which is particularly evident in the "risk" factor, which in the context of OO encompasses a multitude of interrelated aspects. These include risky behaviors such as smoking, as well as external conditions, such as childhood experiences, which collectively contribute to a lack of internal consistency. We recommend an individualized view on a single item level respectively the subcategories of the open-statement section. The factorization seems to lead to a loss of information, which is needed for intervention planning. This might show the complexity and individuality of OO but could also be a chance for mHealth interventions as these can be individualized economically and easily.

The findings of this study contribute to the understanding of the gender differences in causal attributions among individuals with OO who are motivated to lose weight and interested in a psychological mHealth intervention. The practical implications are that implementing stress management interventions with a focus on emotion regulation is pivotal, especially for females. Interventions should focus on sensitizing males to the association between emotions and eating behavior. MHealth interventions that promote strategies to increase health behavior, such as physical activity and reducing alcohol consumption, may be more effective in engaging men than dieting or the proclamation of the consequences of OO. Furthermore, the causal attributions should be assessed with different survey methods in order to individualize (mHealth) interventions and to match the patient's view of their overweight and target treatment options.

Funding: This work was supported by the German Federal Ministry of Education and Research (BMBF Grant Nr. 01GL1719A and 01GL1719B).

Acknowledgments: Student assistants (Bianca Weishaupt and Katja Urwank) helped in part with data collection and acquisition of the sample.

Competing Interests: The authors have declared that no competing interests exist.

Ethics Statement: The design of the study was approved by the ethics committee of the University of Bamberg, Germany and the Institutional Review Board of the Ruhr-University Bochum (no. 18-6415). The study was conducted in accordance with the Declaration of Helsinki. All participants were informed about the content of the entire project.

Preregistration: ClinicalTrials.gov, NCT04080193, September 6, 2019; German Clinical Trials Register DRKS00016623, May 3, 2019

Author Contributions: CH was responsible for writing the paper, screening the literature, extracting, and analysing data, interpreting results, updating the reference lists and creating tables and figures. CS was a major contributor in discussing the results and collecting the data. TF and MP collected the data. All authors provided feedback on and approved the final manuscript.

Data Availability: The datasets and codes used or analysed during the current study are available from the corresponding author on request.

Supplementary Materials

The Supplementary Materials contain the following items:

- The preregistration at ClinicalTrials.gov (NCT04080193) (Henning et al., 2019S-a)
- The preregistration at the German Clinical Trials Register (DRKS00016623) (Henning et al., 2019S-b)
- Additional information (Henning et al., 2024S):
  - Additional File 1: Recruitment process
  - Additional File 2: Single items analysis

Index of Supplementary Materials

Henning, C., Seiferth, C., Färber, T., Pape, M., Herpertz, S., Steins-Loeber, S., & Wolstein, J. (2019S-a). *Gender-sensitive enhancement of common weight loss strategies for overweight and obesity (I-GENDO)* [Preregistration at ClinicalTrials.gov; ID: NCT04080193]. ClinicalTrials.gov. https://clinicaltrials.gov/study/NCT04080193

Henning, C., Seiferth, C., Färber, T., Pape, M., Herpertz, S., Steins-Loeber, S., & Wolstein, J. (2019S-b). *Gender-sensitive enhancement of common weight loss strategies for overweight and obesity: A*

personalized smartphone app [Preregistration at the German Clinical Trials Register; ID: DRKS00016623]. German Clinical Trials Register. https://drks.de/search/en/trial/DRKS00016623

Henning, C., Seiferth, C., Färber, T., Pape, M., Herpertz, S., Steins-Loeber, S., & Wolstein, J. (2024S). *Supplementary materials to "Examination of gender differences: Causal attributions of treatment-seeking individuals with overweight and obesity"* [Additional information]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.15570

References

Agüera, Z., Riesco, N., Valenciano-Mendoza, E., Granero, R., Sánchez, I., Andreu, A., Jiménez-Murcia, S., & Fernández-Aranda, F. (2021). Illness perception in patients with eating disorders: Clinical, personality, and food addiction correlates. *Eating and Weight Disorders, 26*(7), 2287–2300. https://doi.org/10.1007/s40519-020-01083-3

American Psychiatric Association. (2020). *Diagnostische Kriterien DSM-5* (2nd ed., corr.). Hogrefe.

Audureau, E., Pouchot, J., & Coste, J. (2016). Gender-related differential effects of obesity on health-related quality of life via obesity-related comorbidities: A mediation analysis of a French nationwide survey. *Circulation: Cardiovascular Quality and Outcomes, 9*(3), 246–256. https://doi.org/10.1161/CIRCOUTCOMES.115.002127

Awoke, M. A., Harrison, C. L., Martin, J., Misso, M. L., Lim, S., & Moran, L. J. (2022). Behavior change techniques in weight gain prevention interventions in adults of reproductive age: Meta-analysis and meta-regression. *Nutrients, 14*(1), Article 209. https://doi.org/10.3390/nu14010209

Bauer, K., Lau, T., Schwille-Kiuntke, J., Schild, S., Hauner, H., Stengel, A., Zipfel, S., & Mack, I. (2020). Conventional weight loss interventions across the different BMI obesity classes: A systematic review and quantitative comparative analysis. *European Eating Disorders Review, 28*(5), 492–512. https://doi.org/10.1002/erv.2741

Bonsaksen, T., Lerdal, A., & Fagermoen, M. S. (2015). Trajectories of illness perceptions in persons with chronic illness: An explorative longitudinal study. *Journal of Health Psychology, 20*(7), 942–953. https://doi.org/10.1177/1359105313504235

Bramlage, P., Wittchen, H.-U., Pittrow, D., Kirch, W., Krause, P., Lehnert, H., Unger, T., Höfler, M., Küpper, B., Dahm, S., Böhler, S., & Sharma, A. M. (2004). Recognition and management of overweight and obesity in primary care in Germany. *International Journal of Obesity and Related Metabolic Disorders, 28*(10), 1299–1308. https://doi.org/10.1038/sj.ijo.0802752

Breland, J. Y., Dawson, D. B., Puran, D., Mohankumar, R., Maguen, S., Timko, C., Frayne, S. M., & Nevedal, A. L. (2023). Common sense models of obesity: A qualitative investigation of illness representations. *International Journal of Behavioral Medicine, 30*(2), 190–198. https://doi.org/10.1007/s12529-022-10082-w

Broadbent, E., Ellis, C. J., Thomas, J., Gamble, G., & Petrie, K. J. (2009). Can an illness perception intervention reduce illness anxiety in spouses of myocardial infarction patients? A randomized controlled trial. *Journal of Psychosomatic Research, 67*(1), 11–15. https://doi.org/10.1016/j.jpsychores.2008.11.006

Brogan, A., & Hevey, D. (2009). The structure of the causal attribution belief network of patients with obesity. *British Journal of Health Psychology, 14*(1), 35–48. https://doi.org/10.1348/135910708X292788

Chang, V. W., & Christakis, N. A. (2003). Self-perception of weight appropriateness in the United States. *American Journal of Preventive Medicine, 24*(4), 332–339. https://doi.org/10.1016/S0749-3797(03)00020-5

Cooper, A. J., Gupta, S. R., Moustafa, A. F., & Chao, A. M. (2021). Sex/gender differences in obesity prevalence, comorbidities, and treatment. *Current Obesity Reports, 10*(4), 458–466. https://doi.org/10.1007/s13679-021-00453-x

Crane, M. M., Jeffery, R. W., & Sherwood, N. E. (2017). Exploring gender differences in a randomized trial of weight loss maintenance. *American Journal of Men's Health, 11*(2), 369–375. https://doi.org/10.1177/1557988316681221

Daigle, K. M., Gang, C. H., Kopping, M. F., & Gadde, K. M. (2019). Relationship between perceptions of obesity causes and weight loss expectations among adults. *Journal of Nutrition Education and Behavior, 51*(1), 86–90. https://doi.org/10.1016/j.jneb.2018.08.015

Deutsche Adipositas-Gesellschaft e.V. (Ed.). (2014). *Interdisziplinäre Leitlinie der Qualität S3 zur "Prävention und Therapie der Adipositas": Version 2.0* (2nd ed.). Deutsche Adipositas-Gesellschaft (DAG) e.V.

Dombrowski, S. U., Sniehotta, F. F., Avenell, A., Johnston, M., MacLennan, G., & Araújo-Soares, V. (2012). Identifying active ingredients in complex behavioral interventions for obese adults with obesity-related co-morbidities or additional risk factors for co-morbidities: A systematic review. *Health Psychology Review, 6*(1), 7–32. https://doi.org/10.1080/17437199.2010.513298

Donnelly, J. E., Hill, J. O., Jacobsen, D. J., Potteiger, J., Sullivan, D. K., Johnson, S. L., Heelan, K., Hise, M., Fennessey, P. V., Sonko, B., Sharp, T., Jakicic, J. M., Blair, S. N., Tran, Z. V., Mayo, M., Gibson, C., & Washburn, R. A. (2003). Effects of a 16-month randomized controlled exercise trial on body weight and composition in young, overweight men and women: The Midwest Exercise Trial. *Archives of Internal Medicine, 163*(11), 1343–1350. https://doi.org/10.1001/archinte.163.11.1343

Dutton, G. R., Perri, M. G., Dancer-Brown, M., Goble, M., & van Vessem, N. (2010). Weight loss goals of patients in a health maintenance organization. *Eating Behaviors, 11*(2), 74–78. https://doi.org/10.1016/j.eatbeh.2009.09.007

Elfhag, K., & Rössner, S. (2005). Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. *Obesity Reviews, 6*(1), 67–85. https://doi.org/10.1111/j.1467-789X.2005.00170.x

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

Fleary, S. A., & Ettienne, R. (2014). Inherited or behavior? What causal beliefs about obesity are associated with weight perceptions and decisions to lose weight in a US sample? *International Scholarly Research Notices, 2014*, Article 632940. https://doi.org/10.1155/2014/632940

Foright, R. M., Presby, D. M., Sherk, V. D., Kahn, D., Checkley, L. A., Giles, E. D., Bergouignan, A., Higgins, J. A., Jackman, M. R., Hill, J. O., & MacLean, P. S. (2018). Is regular exercise an effective strategy for weight loss maintenance? *Physiology & Behavior, 188*, 86–93. https://doi.org/10.1016/j.physbeh.2018.01.025

Hallam, J., Boswell, R., DeVito, E., & Kober, H. (2016). Gender-related differences in food craving and obesity. *The Yale Journal of Biology and Medicine, 89*(2), 161–173.

Haslam, D. W., & James, W. P. T. (2005). Obesity. *Lancet, 366*(9492), 1197–1209. https://doi.org/10.1016/S0140-6736(05)67483-1

Hebebrand, J., & Gearhardt, A. N. (2021). The concept of "food addiction" helps inform the understanding of overeating and obesity: No. *The American Journal of Clinical Nutrition, 113*(2), 268–273. https://doi.org/10.1093/ajcn/nqaa344

Henning, C., Schroeder, S., Steins-Loeber, S., & Wolstein, J. (2022). Gender and emotional representation matter: Own illness beliefs and their relationship to obesity. *Frontiers in Nutrition, 9*, Article 799831. https://doi.org/10.3389/fnut.2022.799831

Hilbert, A., Dierk, J.-M., Conradt, M., Schlumberger, P., Hinney, A., Hebebrand, J., & Rief, W. (2009). Causal attributions of obese men and women in genetic testing: Implications of genetic/biological attributions. *Psychology & Health, 24*(7), 749–761. https://doi.org/10.1080/08870440801947787

Houle-Johnson, S. A., & Kakinami, L. (2018). Do sex differences in reported weight loss intentions and behaviors persist across demographic characteristics and weight status in youth? A systematic review. *BMC Public Health, 18*(1), Article 1343. https://doi.org/10.1186/s12889-018-6179-x

Karekla, M., Karademas, E. C., & Gloster, A. T. (2019). The Common Sense Model of Self-Regulation and Acceptance and Commitment Therapy: Integrating strategies to guide interventions for chronic illness. *Health Psychology Review, 13*(4), 490–503. https://doi.org/10.1080/17437199.2018.1437550

Kiefer, I., Rathmanner, T., & Kunze, M. (2005). Eating and dieting differences in men and women. *Journal of Men's Health & Gender, 2*(2), 194–201. https://doi.org/10.1016/j.jmhg.2005.04.010

Konttinen, H., Männistö, S., Sarlio-Lähteenkorva, S., Silventoinen, K., & Haukkala, A. (2010). Emotional eating, depressive symptoms and self-reported food consumption: A population-based study. *Appetite, 54*(3), 473–479. https://doi.org/10.1016/j.appet.2010.01.014

Konttinen, H., Silventoinen, K., Sarlio-Lähteenkorva, S., Männistö, S., & Haukkala, A. (2010). Emotional eating and physical activity self-efficacy as pathways in the association between depressive symptoms and adiposity indicators. *The American Journal of Clinical Nutrition, 92*(5), 1031–1039. https://doi.org/10.3945/ajcn.2010.29732

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Lewis, S., Thomas, S. L., Blood, R. W., Hyde, J., Castle, D. J., & Komesaroff, P. A. (2010). Do health beliefs and behaviors differ according to severity of obesity? A qualitative study of Australian

adults. *International Journal of Environmental Research and Public Health, 7*(2), 443–459. https://doi.org/10.3390/ijerph7020443

Löffler, A., Luck, T., Then, F. S., Luppa, M., Sikorski, C., Kovacs, P., Tönjes, A., Böttcher, Y., Breitfeld, J., Horstmann, A., Löffler, M., Engel, C., Thiery, J., Stumvoll, M., & Riedel-Heller, S. G. (2015). Age- and gender-specific norms for the German version of the Three-Factor Eating-Questionnaire (TFEQ). *Appetite, 91*, 241–247. https://doi.org/10.1016/j.appet.2015.04.044

López-Guimerà, G., Levine, M. P., Sánchez-Carracedo, D., & Fauquet, J. (2010). Influence of mass media on body image and eating disordered attitudes and behaviors in females: A review of effects and processes. *Media Psychology, 13*(4), 387–416. https://doi.org/10.1080/15213269.2010.525737

Mathieu, L., Bitterlich, N., Meissner, F., von Wolff, M., Poethig, D., & Stute, P. (2018). Illness perception in overweight and obesity and impact on bio-functional age. *Archives of Gynecology and Obstetrics, 298*(2), 415–426. https://doi.org/10.1007/s00404-018-4827-0

Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12th ed., rev.). Beltz. http://ebooks.ciando.com/book/index.cfm/bok_id/1875625

Moss-Morris, R., Weinman, J., Petrie, K., Horne, R., Cameron, L., & Buick, D. (2002). The Revised Illness Perception Questionnaire (IPQ-R). *Psychology & Health, 17*(1), 1–16. https://doi.org/10.1080/08870440290001494

Mozumdar, A., & Liguori, G. (2011). Persistent increase of prevalence of metabolic syndrome among U.S. Adults: NHANES III to NHANES 1999-2006. *Diabetes Care, 34*(1), 216–219. https://doi.org/10.2337/dc10-0879

Nolen-Hoeksema, S. (2004). Gender differences in risk factors and consequences for alcohol use and problems. *Clinical Psychology Review, 24*(8), 981–1010. https://doi.org/10.1016/j.cpr.2004.08.003

Pantalone, K. M., Hobbs, T. M., Chagin, K. M., Kong, S. X., Wells, B. J., Kattan, M. W., Bouchard, J., Sakurada, B., Milinovich, A., Weng, W., Bauman, J., Misra-Hebert, A. D., Zimmerman, R. S., & Burguera, B. (2017). Prevalence and recognition of obesity and its associated comorbidities: Cross-sectional analysis of electronic health record data from a large US integrated health system. *BMJ Open, 7*(11), Article e017583. https://doi.org/10.1136/bmjopen-2017-017583

Pape, M., Färber, T., Seiferth, C., Roth, T., Schroeder, S., Wolstein, J., Herpertz, S., & Steins-Loeber, S. (2022). A tailored gender-sensitive mHealth weight loss intervention (I-GENDO): Development and process evaluation. *JMIR Formative Research, 6*(10), Article e38480. https://doi.org/10.2196/38480

Pape, M., Herpertz, S., Schroeder, S., Seiferth, C., Färber, T., Wolstein, J., & Steins-Loeber, S. (2021). Food addiction and its relationship to weight- and addiction-related psychological parameters in individuals with overweight and obesity. *Frontiers in Psychology, 12*, Article 736454. https://doi.org/10.3389/fpsyg.2021.736454

Pearl, R. L., Wadden, T. A., Allison, K. C., Chao, A. M., Alamuddin, N., Berkowitz, R. I., Walsh, O., & Tronieri, J. S. (2018). Causal attributions for obesity among patients seeking surgical versus

behavioral/pharmacological weight loss treatment. *Obesity Surgery, 28*(11), 3724–3728. https://doi.org/10.1007/s11695-018-3490-7

Pedersen, S. (2010). Female form in the media: Body image and obesity. In G. Tsichlia & A. Johnstone (Eds.), *Fat matters: From sociology to science* (pp. 5–11). M&K Update Ltd.

Prill, S., Henning, C., Schroeder, S., Steins-Loeber, S., & Wolstein, J. (2021). Does weight-cycling influence illness beliefs in obesity? A gender-sensitive approach. *Journal of Obesity*. Advance online publication. https://doi.org/10.1155/2021/8861386

Roth, L., Ordnung, M., Forkmann, K., Mehl, N., & Horstmann, A. (2023). A randomized-controlled trial to evaluate the app-based multimodal weight loss program zanadio for patients with obesity. *Obesity, 31*(5), 1300–1310. https://doi.org/10.1002/oby.23744

Sabinsky, M. S., Toft, U., Raben, A., & Holm, L. (2007). Overweight men's motivations and perceived barriers towards weight loss. *European Journal of Clinical Nutrition, 61*(4), 526–531. https://doi.org/10.1038/sj.ejcn.1602537

Samdal, G. B., Eide, G. E., Barth, T., Williams, G., & Meland, E. (2017). Effective behavior change techniques for physical activity and healthy eating in overweight and obese adults: Systematic review and meta-regression analyses. *The International Journal of Behavioral Nutrition and Physical Activity, 14*(1), Article 42. https://doi.org/10.1186/s12966-017-0494-y

Schienkiewitz, A., Kuhnert, R., Blume, M., & Mensink, G. B. (2022). Übergewicht und Adipositas bei Erwachsenen in Deutschland – Ergebnisse der Studie GEDA 2019/2020-EHIS. *Journal of Health Monitoring, 7*(3), 23–31. https://doi.org/10.25646/10292

Surgenor, L. J., Snell, D. L., Siegert, R. J., Kelly, S., Flint, R., & Coulter, G. (2020). Psychometric characteristics of the Revised Illness Perception Questionnaire (IPQ-R) in people undergoing weight loss surgery. *Journal of Clinical Psychology in Medical Settings, 27*(1), 79–88. https://doi.org/10.1007/s10880-019-09624-4

Tronieri, J. S., Wurst, C. M., Pearl, R. L., & Allison, K. C. (2017). Sex differences in obesity and mental health. *Current Psychiatry Reports, 19*(6), Article 29. https://doi.org/10.1007/s11920-017-0784-8

Wang, C., & Coups, E. (2010). Causal beliefs about obesity and associated health behaviors: Results from a population-based survey. *International Journal of Behavioral Nutrition and Physical Activity, 7*, Article 19. https://doi.org/10.1186/1479-5868-7-19

Weinman, J., Petrie, K. J., Sharpe, N., & Walker, S. (2000). Causal attributions in patients and spouses following first-time myocardial infarction and subsequent lifestyle changes. *British Journal of Health Psychology, 5*(3), 263–273. https://doi.org/10.1348/135910700168900

Williams, S. L., & French, D. P. (2011). What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour – and are they the same? *Health Education Research, 26*(2), 308–322. https://doi.org/10.1093/her/cyr005

Zhang, L., Schwarz, J., Kleinstäuber, M., Fritzsche, K., Hannig, W., Wei, J., Yang, J., & Zhang, L. (2018). Confirmatory factor analysis of the causal illness attribution scale in Chinese patients with multiple somatic symptoms. *Psychology, Health & Medicine, 23*(sup1), 1056–1070. https://doi.org/10.1080/13548506.2018.1455983

EACLIPT

Clinical Psychology in Europe (CPE) is the official journal of the European Association of Clinical Psychology and Psychological Treatment (EACLIPT).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.

Clinical Psychology in Europe

PsychOpen GOLD

Politics and Education

The History of Clinical Psychology in Greece: A Brief Review – Legal Deficiencies, Practical Dimensions and Challenges for the Future

Katerina Flora [1]

[1] *Department of Psychology, University of Western Macedonia, Florina, Greece.*

Clinical Psychology in Europe, 2024, Vol. 6(4), Article e12515, https://doi.org/10.32872/cpe.12515

Received: 2023-07-31 • **Accepted:** 2024-08-28 • **Published (VoR):** 2024-12-20

Handling Editor: Anton-Rupert Laireiter, University of Vienna, Vienna, Austria

Corresponding Author: Katerina Flora, Kimonos Voga 50, 54645, Thessaloniki, Greece. Phone: 0030-6973238267. E-mail: kflora@uowm.gr

Abstract

Background: The history of clinical psychology in Greece spans more than 150 years. However, this branch of psychology concerned with the assessment and treatment of mental illness and psychological problems has not yet acquired the institutional and general recognition to which it is entitled.

Aims: This article intends to highlight, chronologically, the basic elements of the history of clinical psychology in Greece, beginning with the important contribution of the work of philologist Panagiota Kazolea-Tavoularis.

Results: From the first references in the context of medical studies during the 19th century, clinical psychology gradually develops through its application in pedagogical, laboratory, and clinical contexts to become an independent discipline alongside the consolidation of general psychology. Special mention is made of the scientists who pioneered this direction.

Conclusion: The present review highlights historical milestones and concludes with the current situation, in which important steps have been taken. However, significant changes are needed at the institutional level.

Keywords

clinical psychology, history, applied psychology, Greece, review

Highlights

- During the 19th century elements of clinical psychology were formulated in the context of medical studies.
- During the 20th century, clinical psychology developed more widely in pedagogical, laboratory and clinical contexts.
- The formation of clinical psychology as an independent branch took place in parallel with the consolidation of general psychology.

The purpose of this overview is to coherently present the history of clinical psychology, the branch of psychology concerned with the assessment and treatment of mental illness and psychological problems, in Greece. The basic bibliographic source for the writing of this article is the doctoral thesis of Panagiota Kazolea-Tavoularis (2001), titled "The history of psychology in Greece (1830–1987)". This particular thesis gathers rich and original material. What is described in this article is, by a significant part, an extraction from the thesis, specifically related to clinical psychology. It is worth noting that there are no specific sources for the history of clinical psychology in Greece. Therefore, the source material mainly concerned the history of psychology in Greece in general, the history of applied psychology, and the history of psychiatry. Processing of the found material resulted in five general origins and strands from which clinical psychology in Greece is fed: medicine, especially neurology and psychiatry; philosophy and religion; educational science and pedagogy; psychoanalysis and psychotherapy; and the development of American and European (experimental and applied) psychology in general, with its various fields and disciplines.

This article follows a chronological order and concludes with a description of the current state of clinical psychology in Greece. The chronological order follows the work of Kazolea-Tavoularis' thesis, with the addition of older and modern sources added to the basic elements included in the historical review.

The Beginning – 19th Century

The development of clinical psychology in the 19th century was influenced mainly by medicine, particularly neurology and psychiatry, and its institutions (clinics, universities, and practitioners). Since the 19th century, applied psychology subjects have been taught in medical departments throughout Greece. Psychological theories concerning mental disorders and shaping the early practice of clinical psychology appeared during the 19th century and were taught at medical schools in the context of general pathology, since psychiatry as a specialty was unestablished. The influence of psychological factors on physical health was emphasized, which equated with the harmony of body and soul. Outside of a strict academic context, psychological issues were also referred to by

doctors, and an early use of the term "psychosis" was recorded with reference to the history of Western psychiatry regarding the classification and treatment of phrenosis (Kazolea-Tavoularis, 2001). As early as 1885, scientific publications such as the *Journal Medical* had been established, in which doctors at the time – such as Ioannis Foustanos, a student of Jean-Martin Charcot – referred to hypnosis, removing any metaphysical dimension from the behaviour of the hypnotized.

Apart from academic, theoretical psychology, an applied clinical psychology was practised in psychiatric institutions, and it may have been influenced, to some extent, by religious and philosophical currents. According to Dimitris Ploubidis (1995), practical psychiatry was practised in several monasteries through wishes, exorcisms, and incantations but also with more violent means such as binding with chains and confinement. These practices were succeeded from 1838 on by the establishment of the insane asylum of Corfu and the Dromokaiteio sanatorium of Attica according to the standards of Western European sanatoriums based on the theories of Jean-Etienne Dominique Esquirol (Karamanolakis, 1998). These institutions were followed at the beginning of the 20th century by the Aegineteio (Athens) 1905, the public psychiatric hospital of Athens in the 1920s, and the public psychiatric hospitals of Thessaloniki and Souda, Chania (Crete) (Chartokollis, 1991).

During that period, psychological theories and clinical practices that the Greek doctors had learned from their studies in Europe were applied in these institutions, following the French school of degeneration of Valentin Magnan or the German school of psychiatry, which, at the end of the century, ended up in the Kraepelin classification of primitive dementia. There is a discrepancy in the diagnoses, as could be seen from the archives of the Corfu psychiatric hospital (1880–1885) where diagnostic terms such as lysomania, bimorphous phrenitis, bimorphous paranoia, rationalized paranoia, isodemia, enteric frenzy, and exphylogenic frenzy were mentioned. It is worth mentioning the work of two Greeks with international recognition, Grigorios Rosolimos in Russia and Konstantinos Oikonomou in Austria, in the field of neuropsychology during the same period.

In Greece in the 19th century, scientific events unfolded with a delay, as the state structures and educational ties were absent in the newly established state. Therefore, beyond the above references, psychology remained largely philosophical and was slow to evolve into an applied field.

Clinical Psychology in the 20th Century

Overview and Lines of Development

The 20th century saw a more intensive and heterogeneous development of psychology and clinical psychology in particular. The development of clinical psychology in Greece

came in two phases, the first lasting until about 1960 and the second following thereafter, as the discipline was largely established and consolidated as an independent scientific and practical field.

The First Phase: Structure and Development

The first phase of the development of clinical psychology was strongly influenced not only by European and American psychology but also by pedagogy and psychotherapy. Institutes and professorships of psychology were established in Greece's large cities and in institutions of applied clinical psychology in the fields of health and education.

In 1964, the first chairs of psychology were established at the universities of Ioannina and Thessaloniki (Kazolea-Tavoularis, 2001). Hypnosis and psychoanalysis were being practised elsewhere in Europe, and Greece imported knowledge about hypnotism for medical or parapsychological use. This sparked a dialogue that contributed to the awakening of interest in the soul as an entity and, by extension, in psychic phenomena. References to hypnosis, dreams, and parapsychological phenomena were noted in general psychology throughout the first half of the 20th century. Some interested parties, however, leaned more towards mysticism, promoting religious beliefs and accusing scientific and research sources of being atheistic or materialistic.

Scientific Psychology in Universities and Laboratories

While psychology was taught as a general subject at the Philosophical School of Athens until 1908, as psychopathology at the Medical School of Athens, and as general psychology in teaching and secondary education, it was still part of philosophy courses at the beginning of the 20th century. The first psychology laboratory, which provided important services to student practice and research, was founded in 1926 at the University of Athens by Theophilos Boreas. The second such laboratory was established in 1935 in Thessaloniki, where the activities of Professor Georgios Sakellariou were particularly noteworthy; he founded the magazine *Prometheus* in 1951 and the Hellenic Psychological Society in 1955. At the same time, psychopedagogical research was being conducted by Alexandros Delmouzos, a professor of pedagogy in Thessaloniki and a student of Wilhelm Wundt.

The psychological laboratory of Athens was divided during the 1950s into four departments: psychological research, professional guidance, clinical psychology, and enlightenment of parents and young people; this was the first reference to clinical psychology at an academic level. The aims of the laboratory also included applied psychology. The work of S. Paraskeva-Sakka, who studied philosophy in Thessaloniki and psychology in the United States and was a colleague of Sakellariou, was noteworthy. Paraskeva-Sakka specialized in vocational guidance and psychoanalytic psychotherapy and translated and used many of the psychological tests still in use today.

Applied Psychology, the Beginnings

Psychological scientists were sought after by the Greek Army early in the 20th century to help select and train conscripts in the most efficient use of new weapons and machines. In 1917–1918, a special committee drafted tests for the selection of aviators while selection methods were being implemented for conscripts and candidates for military schools. In 1948–1949, Sakellariou trained Army officers in the psychological laboratory at which intelligence tests were administered. Since the 1950s, the Army has used intelligence scales such as the Terman–Sakellariou scale and the Nikolaos Exarchopoulos Progressive Matrix Test scale. Additionally, applied psychology was used in the Greek Army until 1997, when the institution of psychosocial care was established to address conscripts' problems.

Important contributions to the spread of psychology in Greece were made by two female pioneers, Sofia Gedeon and Aikaterini Striftou-Kriaras. They were collaborators of Nikolaos Exarchopoulos, the chair of pedagogy in the Laboratory of Experimental Pedagogy, who was succeeded by Spyridon Kalliafas. Kalliafas turned to the study of psychological issues, as can be seen by his publication of *Characters or Psychological Types* in 1935, in which he referred to older characterizations starting with Plato and Aristotle and continued with the psychoanalytic schools of Freud, Adler, and Jung. His study mentioned the relevant psychology of individual differences of Fechner, Charcot, Taine, Binet, and Stern and the contributions made by two Greeks, Nikolaos Exarchopoulos and Georgios Sakellariou. Kalliafas glorified Jung's typology and was perhaps the most basic and important exponent of his work in Greece.

Konstantinos Specieris succeeded Kalliafas in 1953. In his work, *The Psychosynthesis of Man*, Specieris attempted a philosophical approach to the psyche; however, he emphasized the therapeutic effects of the psychoanalysis of Freud, Adler, and Jung. For the evaluation of the personality, he proposed his own psychograph type with 13 questions. Specieris' work, *The Mental Life of Man* (Specieris, 1960) was a revision of his previous work based on the latest scientific findings. It referred equally critically to behaviourism, psychoanalysis, individual psychology, and the philosophy of existence. He also directly questioned the method of psychological tests and, in general, quantitative measurements that provide only partial knowledge of phenomena. He argued that the totality of mental life was greater than the sum of its parts in accordance with morphological psychology, while insightful understanding and deepening introspection were required to understand the human psyche.

Applied psychology was introduced in Greece at the beginning of the 20th century with the particularities that characterized Greek scientific and social reality at that time. Psychology was not clearly accepted as an independent scientific field, so discussion about Freud's theories of psychology and psychoanalysis took place within the circles of pedagogy. The first translations of Freud's works were published in the early 1900s, while at the same time the psychoanalytic movement was remembered mainly for its

pedagogical application. Beyond Freud, Adlerian principles of pedagogy were taught in selected schools of the country and also in special education through the work of the pioneering pedagogue Roza Imbrioti.

The medical community in Greece initially had a negative attitude towards psychoanalysis, which favoured the demedicalization of psychoanalysis in Greece during the first period of its introduction. However, Dimitrios Kouretas, a neurologist and psychiatrist, delivered the first lecture on psychoanalysis in Greece in 1927. He contributed to the creation of the first psychoanalytic nucleus in Greece in 1947 with Andreas Empirikos, a poet and exponent of surrealism in Greece, and Georgios Zavicsianos, a psychiatrist, under the supervision of Princess Marie Bonaparte, a student and translator of Freud and president of the Paris Psychoanalytic Society. The name Nikolaos Drakoulides was also mentioned; he wrote a number of studies, including one on Freud and psychoanalysis (1936) and another on the psychoanalytic interpretation of art (1948). Psychoanalysis, however, was practiced clinically and was referred to as the new psychotherapeutic method with encouraging results since the 1930s. Prominent practitioners were psychiatrist Mihail Vlastos, psychiatrist Fotis Skouras, and neurologist–psychiatrist Konstantinos D. Konstantinidis, a professor at the Medical School of Athens and director of the Public Psychiatry of Athens. Countering this group was Georgios Zouraris, a member of the Institute for Sexual Research in Berlin, who favoured psychobiology and criticized concepts of psychoanalysis such as childhood sexuality and the Oedipus complex.

Given the interaction and mutual borrowing between disciplines and the theoretical issues of psychology and psychiatry (Tzavaras, 1991), a reference to psychiatry in Greece also concerned the history of psychology in the country, insofar as the conceptions of the "soul", its functions, and its pathology were common points of concern for both. The popularity of psychological theories and therapeutic practices such as psychoanalysis and others, since it was only in 1963 that neurology and psychiatry were institutionally separated, erased the special characteristics of clinical psychology in institutions and as taught by university departments. After all, the concept of mental illness was part of a specific institutional framework, which, reflected the socio-economic conditions at the time as well as the ideological parameters in Greece, as Kazolea-Tavoularis (2001) commented.

Mental Health Services: The Treatment of the Mentally Ill

Of interest was the treatment of the mentally ill, which during the period of the Ottoman Empire – which is a long period for which there is not much evidence – was done in asylum-type institutions and sometimes in churches and monasteries. Faith was thought to heal the mentally ill in a sense, while inhumane practices such as folk psychosurgery with red-hot irons were reported (Ploubidis, 1995).

The first asylum was established in Corfu, in 1838, by the British administration and was housed in the equestrian stables. The first doctors of the asylum were British, while the first Greek doctor, Christodoulos Tsirigotis, took over as director in 1874.

Already in Constantinople-during Ottoman Empire- there were several institutions that accepted the mentally ill. In the 16th century, we have the establishment of the hospital of Galata of Gemintzidon, which accepted the mentally ill.

In 1780 the hospital of Stadrodomiu, in 1839 of Heptapyrgio, in 1855 the La Paix asylum in Constantinople was founded by Catholic nuns, in 1748 the Greek hospital with an insane asylum department.

In Constantinople, the Ottomans founded two insane asylums in 1465, the Fatih Mosque Hospital and in 1527 the second one in the Suleiman Mosque Hospital. In 1583, Sultana Valide was founded and in 1850 the TopTahi Hospital, both of which received mental patients (Madianos 1994; Ploubidis, 1995).

At the end of the 19th century and the beginning of the 20th century, some forms of psychiatric treatment were available, as it was reported that hospitals in Chios, Smyrna, and Constantinople received mentally ill patients with various diagnoses such as primitive dementia, progressive general paralysis, and mania-melancholia (Ploubidis, 1995).

Opinions about mental illness in Greece were shaped according to the studies of Greek psychiatrists abroad. Some followed the German school and others the French school. Essentially, psychiatry was organized after 1930 and the founding of the Neurological and Psychiatric Society. The prevailing opinion was that psychiatric diseases derived from organic causes, and the therapeutic practices were similar, including bed rest, electric shock, drugs, cold and hot water, occupational therapy, and hydrotherapy.

Public psychiatric hospitals, such as the University Psychiatric Clinic Aeginetio founded at the start of the 20th century, were staffed by academic psychiatrists and treated patients using the well-known psychotherapeutic methods of the time. After the 1960s, psychosocial approaches such as drama groups, psychodrama, and group psychotherapy were also mentioned. More generally, however, the psychiatric world remained oriented toward the neurobiological and organic basis of mental disorders while promoting chemical treatments and expressing doubts about the findings of the new science of psychology and psychoanalysis. Over time, however, psychoanalysis was increasingly supported by a growing number of medical representatives. However, Greek society at that time – characterized by its low level of education, conservatism, orthodox Christian religion, and poor economic situation for large numbers of the population – generally maintained a negative attitude toward psychoanalysis and its emphasis on sexuality.

In addition to academic psychology, a network of mental health services was created in the context of social welfare in the early days of the 20th century. Associations promoting mental health, institutions treating children's psychosomatic health, and, since the mid-20th century, institutions for the blind and deaf, psycho–pedagogical centres, medi-

cal–pedagogical centres, psychotherapeutic clinics, vocational guidance centres, student perception centres, and psychological test institutes were established.

An important development for community mental health was the establishment in the 1950s, following a proposal by psychoanalyst Anna Potamianou, of the Centre for Mental Hygiene, which offered a series of counselling and therapeutic prevention and intervention services. A corresponding development was the Psychological Centre of Northern Greece in Thessaloniki in the 1960s, which aimed to address the broader mental health problems in Northern Greece. Its function was closely related to the scientific activity of Eftychia Nanakou, a doctor and psychologist who treated children with mental retardation. In the 1960s, the Athenian Centre for the Study of Man was founded by psychiatrist Georgios Vassiliou and psychologist Vaso Vassiliou as the first institute of systemic and family therapy, with methods based on family and group dynamics such as those taught at Loyola University of Chicago.

After 1960 – The Consolidation of Clinical Psychology

During the period 1964–1997, the above processes were been completed, and clinical psychology was consolidated and developed into various fields. Hence, Greece saw the consolidation of the science of psychology in both academic instruction and scientific activity. At the Universities of Ioannina, Thessaloniki, and Athens, psychology courses were taught within the faculty of philosophy in the Departments of Philosophy, Pedagogy, and Psychology. Later the field of psychology became autonomous, and independent departments of psychology were created. Clinical psychology was taught alongside the other subjects of psychology during these years. Key representatives of applied psychology were Maria Nasiakou in Ioannina as a school psychologist, Mika Haritou Fatourou in Thessaloniki as a clinical psychologist, and Ioannis Paraskevopoulos and Anastasia Kalantzi-Azizi in Athens in psychometry and clinical psychology, respectively.

Applied Psychology, as we saw above, was taught as early as the 1930s in pedagogic departments, academies and teaching in the context of either general or special education. Educational and school psychology, developmental and evolutionary psychology, developmental psychopathology, assessment of intelligence and learning disabilities as well as therapeutic interventions are the main courses taught at this time, after 1960, in these schools. The 1970s saw the establishment of parenting schools through the work of psychologist Maria Khourdakis, which aimed to protect the mental health of both children and parents. These schools were based on the theories of family psychology, educational psychology, and school pedagogy. They had a significant impact on the public and 135 schools were operating nationwide by the 1980s.

Important in the consolidation and independence of the science of psychology in Greece was the establishment of the Association of Greek Psychologists in 1963 and the Panhellenic Psychological Association in 1997, which defended the rights of licensed professional psychologists. However, despite continuous efforts, psychologists in Greece

have not succeeded in establishing a single association that constitutes a legal entity under public law. Additionally important is the presence of the Hellenic Psychological Society, a scientific association founded in 1990, that includes branches of clinical psychology and health psychology, among others. Apart from this organization, there is no separate association of clinical psychology in Greece.

Moreover, the connection between clinical psychology and psychotherapy lacks a clear framework in Greece. Psychotherapy is clearly a part of clinical psychology; for example, in the postgraduate programs of clinical psychology, psychotherapy is certainly a responsibility of the clinical psychologist. By contrast, many private institutes for the training of psychotherapists are not directly related to clinical psychology, as this discipline is taught in universities. It is imperative to institutionally clarify the professions of clinical psychologist and psychotherapist, a distinction that remains unclear.

Additionally, in the context of social welfare, various institutions have been established to accommodate children with various disabilities or those who have been removed from their families due to neglect or abuse. Children's psychological support and therapeutic interventions, occupational therapy, psychosocial support, and special vocational training have been broadly applied in these structures.

At the same time, since the 1950s, various vocational orientation tests have been introduced in the psychological laboratories of Athens and Thessaloniki, while counselling for students has also been offered. From 2000 onwards, programs to remove social exclusion have been extended to various categories of people, such as prisoners, addicted people and disabled people, mainly with funding from the European Union.

Clinical Psychology, Psychiatry, and Psychiatric Reform

Clinical psychology in Greece remains a popular discipline among young psychologists. Its relationship with psychiatric medicine was partly conflictual due to the different backgrounds of the two sciences and the therapeutic practices that each proposes. However, the psychological training of doctors should be noted, with the introduction of psychology courses in medical schools in 1970 and the increasing participation of psychiatrists in psychotherapeutic training programs. At the same time, the contribution of clinical psychologists to psychiatric reform, oriented towards deinstitutionalization, community psychology, and social psychiatry, was important. The work of psychiatrists Kostas Stefanis, Panagiotis Sakellaropoulos, Petros Hartokollis, Stavroula Berati, Nikolaos Tzavaras, Charalampos Hierodiakonou, and Georgios Anastasopoulos and psychologists Anna Kokkevi, Alexandra Routsoni, R. Diakogianni, Maria Dolianiti, Ilias Fragos, and Thaleia Vergopoulou is important in this direction.

The operation of psychiatric hospitals and the structure of psychiatric reform from the 1980s onwards contributed to the promotion of clinical psychology in Greece. The same decade saw the establishment of the first addiction centres in Greece, KETHEA and 18 Ano, with the significant contributions of psychiatrists Phoebus Zafirides and

Katerina Matsa, respectively. Psychiatrists, psychologists, social workers, and psychiatric nurses worked at these centres, comprising the interdisciplinary teams.

Clinical Child and Adolescent Psychology and Psychiatry

Child psychiatry organisations have been established since 1950 to treat the mental health of children and adolescents. An important activity for the psychosomatic health of children was developed at the child psychiatric clinic of the Agia Sophia Children's Hospital, founded in 1978. This department provides services such as diagnostic consultation, short hospitalization and treatment, psychosocial support, and pre-professional training. The department trains doctors and psychologists and has prepared prevention programs. Psychiatrists and psychologists Gerasimos Stefanatos, Ioannis Tsiantis, Olga Maratou, and Anna Kokkevi pioneered this effort.

Also related to the practice of clinical psychology in Greece is the establishment of various organizations from 1960 onwards with the aim of providing psychotherapeutic services, the education and training of mental health professionals, and, sometimes, research. These groups followed various psychological approaches and practices and include the Hellenic Counselling Society; Hellenic Psychoanalytic Society; Hellenic Society for Dyslexia; Hellenic Society for Research and Behavior Therapy; Centre for Family Therapy; Alcoholics Anonymous; Centre for Individual Psychology; Centre for Psychotherapy and Gestalt Training; Hellenic Society for Community Therapists, Sociotherapy, and Psychodrama; and the Medical Sexology Institute. These societies have become widely associated with applied psychology, as many graduates of psychology departments have continued their education there to be trained in a psychotherapeutic approach, thus complementing the knowledge offered by academic study. The training in these organizations had a more practical orientation that psychology departments usually did not offer except at the postgraduate level.

Academic Departments

Perhaps the most important event for the development of psychology in Greece and, by extension, clinical psychology was the establishment in the mid-1980s of independent university departments of psychology. The first Department of Psychology was founded in 1987 at the University of Crete. The contribution of psychologist Maria Khourdaki was important as, together with professors Ioannis Nestoros, Nikos Papadopoulos, and Grigoris Potamianos, she founded the new department. Following the example of the University of Crete, other psychology departments were established in Athens and Thessaloniki. All departments arose from philosophical schools with the exception of the one at Panteion University, which was rooted in the sociology department. The newly established Department of Psychology of Panteion University was staffed by professors

of the Department of Sociology such as Stamos Papastamou, Foteini Tsalikoglou, and Aimilios Metaxopoulos.

In 1992, a Department (Program) of Psychology was established at the Kapodistrian University of Athens with professor George Paraskevopoulos and assistant professors Ilias Bezevegis, Nikolaos Giannitsas, and Anastasia Kalantzi-Azizi staffing this new department, which was part of the Department of Philosophy, Pedagogy, and Psychology. Postgraduate programs were not initially offered.

Since 1993, an autonomous Department of Psychology has been operating at the Aristotle University of Thessaloniki, which also emerged from the Department of Philosophy, Pedagogy, and Psychology. The department followed three directions – school–evolutionary, experimental–cognitive, and social–clinic – and was founded by professors Maria Maniou-Vakali, Anastasia Euclides, Dimitra Papadopoulou, Mika Charitou-Fatourou, Diomedes Markoulis, Dimitrios Natsopoulos, and Kostas Bayraktaris. Postgraduate studies have been offered almost from the beginning, while various scientific activities such as conferences, workshops, and seminars have been conducted.

The map of psychology departments in Greece was completed with the establishment of two new departments in 2019: the University of Western Macedonia, based in Florina, and the University of Ioannina, which also emerged from the older Department of Philosophy, Pedagogy, and Psychology.

Instead of an Epilogue: Clinical Psychology in Greece Today

In Greece at the moment there are six psychology departments in public universities, four of which provide at least two years of postgraduate study in clinical psychology. Additionally, at least 10 affiliates of foreign universities have been established in which undergraduate studies in psychology are taught, and at least five offer postgraduate studies in clinical psychology. Private institutes providing training in various treatment approaches are numerous, and it is difficult to calculate their exact number.

The professional training of clinical psychologists, as can be deduced from the above, is provided at the postgraduate level mainly at public universities but also at private colleges. These programs have an academic structure and also provide extensive practical training.

As mentioned above, professional clinical psychology is not recognized officially and it is not institutionally protected. Hence, there are no financial benefits for people who use the services of clinical psychologists through Greece's healthcare system. At the institutional level, there is the branch of clinical psychology in the Hellenic Psychological Society that validates this specialty.

Clinical psychologists in Greece work in both medical and psychiatric clinics, addiction treatment programs, and other therapeutic settings as well as on an outpatient

basis or in private practice. In recent years, the positions of clinical psychologists in public institutions have been scarce, leading to more clinicians to be employed in private practices. The work of clinical psychologists in medical and psychiatric clinics includes diagnosis, therapy, consultation, crisis intervention, and supervision. In private practices, it involves mainly therapy and consultation.

Another important issue is the unclear distinction between clinical psychology and psychotherapy. Officially, neither is officially recognized as a profession in Greece. However, psychotherapy is undeniably part of the work of a clinical psychologist. Informally, psychotherapists are those who have completed training in a particular therapeutic approach, and their numbers are probably greater than clinical psychologists, but official figures are lacking. The public health system employs more clinical psychologists since academic qualifications, such as a master's degree and a doctorate, are usually required.

Through the above-described developments, clinical psychology has established itself as a discipline in Greece. However, there are institutional gaps in care that must be addressed to secure and further develop the field of clinical psychology (see Kalantzi-Azizi & Karadimas, 2009). First and most importantly, the Greek legislation regarding the profession of psychologist includes no mention of registered specializations such as clinical, school, or cognitive psychology. Therefore, clinical psychology is taught at an academic level and practiced at a clinical level, but it does not have legal status as a specialization.

A second deficiency that undermines the field of clinical psychology is the fact that the licence to practice the profession of psychologist in Greece is issued directly after the completion of undergraduate studies, with only two months of practice. This is unprecedented, as a psychologist can typically undertake clinical work without having completed the necessary supervised clinical practice. The most basic requirement for a clinical psychologist should be supervised clinical practice for at least one year.

In conclusion, clinical psychology in Greece has a long history and has been served by notable scientists, initially from the field of medicine and later from psychology. However, the regulation by the state is needed to secure the field of clinical psychology in the present and to create the groundwork for its development in the future.

Funding: The author has no funding to report.

Acknowledgments: The author has no additional (i.e., non-financial) support to report.

Competing Interests: The author declares that she has no relevant or material financial interests that relate to the research described in this paper.

Reporting Guidelines: The author has followed the EQUATOR Reporting Guidelines (https://www.equator-network.org) and has used the PRISMA guidelines (http://prisma-statement.org).

References

Chartokollis, P. (1991). *Εισαγωγή στην ψυχιατρική* [Introduction to psychiatry]. Themelio Publications.

Kalantzi-Azizi, A., & Karadimas, E. (Eds.). (2009). *Η Σχέση της Κλινικής Ψυχολογίας με τους Εφαρμοσμένους Κλάδους της Ψυχολογίας και την Ψυχοθεραπεία* [The relationship of clinical psychology with other applied fields of psychology and psychotherapy]. Topos Publications.

Karamanolakis, V. (1998). Το Δρομοκαϊτειο Φρενοκομείο: 1887-1903. Όψεις της εγκατάστασης ενός ιδρυματικού θεσμού [Dromokaitio, the psychiatric hospital: 1887-1903. Views of facility of one institutional institution]. *Mnimon, 20*, 45–66. https://doi.org/10.12681/mnimon.666

Kazolea-Tavoularis, P. (2001). *Η ιστορία της ψυχολογίας στην Ελλάδα (1830-1987)* [The history of Psychology in Greece (1830-1987)]. [Doctoral dissertation]. Panteion University of Social and Political Sciences. https://thesis.ekt.gr/thesisBookReader/id/15021#page/1/mode/2up

Madianos, M. (1994). *Η ψυχοκοινωνικη αποκατασταση απο το ασυλο στην κοινοτητα. Παραρτημα: το νομικο πλαισιο της αποκατασταση*ς [Psychosocial rehabilitation from the asylum to the community. Annex: the legal framework of rehabilitation]. Ellinika Grammata.

Ploubidis, D. (1995). *Η ιστορία της ψυχιατρικής στην Ελλάδα. Θεσμοί, Ιδρύματα και Κοινωνικό Πλαίσιο* [The history of Psychiatry in Greece: Institutions and Social Context]. Exantas Publication.

Specieris, K. (1960). *Η ψυχική ζωή του ανθρώπου* [The mental life of man]. Typois Mina Myrtidi.

Tzavaras, T. (1991). Οι σχέσεις ψυχολογίας και ψυχιατρικής από ιστορική-επιστημολογική άποψη [The relations between psychology and psychiatry from a historical-epistemological point of view]. *Psichologika Themata, 4*(2), 113–121.