ORIGINAL ARTICLE

JARID bild WILEY

Psychometric analysis of the Group Climate Inventory—Revised in adults with mild intellectual disability or borderline intellectual functioning in a secure residential facility

A. Turhan[1,2] | J. J. Roest[3] | M. J. Delforterie[1,2] | G. H. P. Van der Helm[3,4] | E. G. Neimeijer[2] | R. Didden[1,2]

[1]Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

[2]Trajectum, Zwolle, The Netherlands

[3]Research Group Residential Youth Care, Leiden University of Applied Sciences, Leiden, The Netherlands

[4]Faculty of Social and Behavioural Sciences, University of Amsterdam, Amsterdam, The Netherlands

Correspondence
A. Turhan, Trajectum, Dokter Stolteweg 17, 8025 AV Zwolle, The Netherlands.
Email: aturhan@trajectum.info

Abstract

Background: In secure residential facilities, group climate perceptions of clients with mild intellectual disability or borderline intellectual functioning are systematically assessed for quality improvement. A valid and reliable measure may ensure that this process is consistent. The Group Climate Inventory—Revised (GCI-R) is a new measure to assess group climate perceptions.

Method: Confirmatory factor analysis was conducted in 148 adult clients (79% male) with mild intellectual disability or borderline intellectual functioning in a secure facility to examine internal structure validity and internal consistency reliability of the GCI-R.

Results: The results indicate support for the five-factor structure of the GCI-R ('Support', 'Growth', 'Repression', 'Peer interactions', and 'Physical environment'). The internal consistency reliability of its scales ranged from acceptable to good (α: .72–.87; ω: .76–.86).

Conclusion: The GCI-R demonstrates evidence of psychometric adequacy when applied to adult clients with mild intellectual disability or borderline intellectual functioning in secure residential facilities.

KEYWORDS
borderline intellectual functioning, confirmatory factor analysis, Group Climate Inventory, mild intellectual disability, residential treatment, secure facility

1 | INTRODUCTION

Providing treatment within an open group climate is considered essential to the successful recovery and rehabilitation of clients in secure residential facilities (Day et al., 2012; Harding, 2014; Leipoldt et al., 2019; Puzzo et al., 2023). Group climate in secure residential facilities is characterised by transactional processes between the clients in a group or between clients and staff members (Sameroff, 2009), and it has been defined as 'the quality of the social- and physical environment in terms of the provision of sufficient and

necessary conditions for physical and mental health, well-being, contact and personal growth of the residents, with respect for their human dignity and human rights as well as (if not restricted by judicial measures) their personal autonomy, aimed at recovery and successful participation in society' (Stams & Van der Helm, 2017, p. 4). There is relative consensus about which elements constitute group climate in secure residential facilities. A meta-analysis by Eltink (2020) identified seven dimensions of group climate, including (1) supportive and responsive staff behaviour, (2) opportunities for growth and learning, (3) a structured environment with clear rules, (4) safety of clients from physical and psychological harm, (5) justice and fairness, (6) social interactions between clients, and (7) repression by staff. Group climate may vary from an open to a closed group climate.

An open group climate is characterised by mutual trust and respect. In wards with an open group climate, a positive group atmosphere is common, clients are provided with support and growth opportunities aimed at strengthening their autonomy in a clean, safe, and structured environment, and little or no repression is exerted by staff members. In contrast, a closed group climate is characterised by mutual hostility and low levels of safety, support, and growth opportunities. In wards with a closed group climate, treatment is often provided in a less structured environment. Moreover, wards with a closed group climate often have high power distance, unfair rules as perceived by clients, and high levels of repression and staff turnover (Roest, 2022; Schaftenaar, 2018).

An open group climate is associated with beneficial outcomes in clients without intellectual disability in secure residential facilities. For example, it is associated with greater treatment motivation (Van der Helm et al., 2018), better quality of life (O'Flynn et al., 2018), and less aggressive behaviour (Robinson et al., 2018). Research on group climate in clients with mild intellectual disability or borderline intellectual functioning (IQ 50–85) has been limited thus far, although similar benefits of an open group climate may apply to this population. For example, available research in this population suggests similar associations between an open group climate and less aggressive behaviour (Neimeijer et al., 2021). It is therefore unsurprising that in recent years more attention has been paid to measuring group climate in clients with intellectual disabilities (Bell, 2022; Neimeijer et al., 2019; Van der Helm et al., 2021).

Several self-report measures are used to assess perceptions of group climate in clients and staff (for a comprehensive overview, see Tonkin, 2016), which vary in terms of scale and item content, operationalisation, and which dimensions of group climate are measured. Group climate measures provide an overview rather than an in-depth assessment of group climate (De Vries et al., 2018). Regardless of which measure is used, an in-depth assessment of group climate requires follow-up conversations between clients and staff on a ward about its aggregated results (De Vries et al., 2018; Van Ginneken & Nieuwbeerta, 2020). Routinely assessing the group climate perceptions of clients and staff, discussing the results, and assigning improvement actions generally results in a more open group climate within secure residential facilities (Levrouw et al., 2018). The measures most frequently reported in the reviews by Dickens et al. (2022)

and Tonkin (2016) are the Ward Atmosphere Scale (WAS: Moos, 1968) and the EssenCES (Schalast et al., 2008).

The WAS consists of 100 items across 10 scales, grouped into the factors 'Relationship', 'Personal growth', and 'System maintenance'. Relationship assesses the quality of relationships on a ward, personal growth evaluates the extent to which personal development is stimulated in clients and system maintenance maps how well a ward is organised. While the WAS is widely used in Britain and the United States, concerns have been raised about its reliability and proposed factor structure. As WAS-items were formulated in the late 1960s, some items may no longer be appropriate in contemporary secure mental health settings (Røssberg & Friis, 2003; Schalast et al., 2008). In addition, the WAS may be unsuitable for routine use as it takes a relatively long time to complete (Schalast et al., 2008; Tonkin, 2016). Furthermore, the utility of WAS in people with intellectual disabilities is limited (Bakken et al., 2012; McGee & Woods, 1978).

The EssenCES is a brief measure of 15 items that cover the factors 'Therapeutic hold', 'Patients' cohesion and mutual support', and 'Experienced safety'. Therapeutic hold assesses the quality of therapeutic relationships, patients' cohesion and mutual support evaluates the extent to which clients support each other's rehabilitation, and experienced safety maps safety from aggression and violence. Originally developed in Germany, the EssenCES has since been translated into several languages. There is increasing empirical evidence demonstrating good psychometric properties of the EssenCES in a wide variety of populations in secure settings (Tomlin & Tonkin, 2023). However, the current version of the EssenCES is not suitable for clients with intellectual disabilities due to the difficult wording of its items (Quinn et al., 2012). Bell (2022) recently proposed a modified EssenCES for use in forensic services for people with intellectual or neurodevelopmental disorders. Future research has yet to determine its psychometric properties within this setting. Other researchers have modified the EssenCES for use with clients with intellectual disabilities in a low-risk forensic setting and reported preliminary support for its psychometric properties in this population (Barker et al., 2021).

In view of the WAS and EssenCES, the Group Climate Inventory (GCI: Neimeijer et al., 2019) represents a good alternate measure for routine use that is considered to provide a comprehensive measurement of group climate using a relatively small number of items (Eltink, 2020; Van der Helm et al., 2021). In addition, the measure was designed from the outset to be suitable for use with clients with mild intellectual disability or borderline intellectual functioning, as reflected in its relatively simply worded items. The GCI, derived from the Prison Group Climate Instrument (Van der Helm et al., 2009, 2011), is a 29-item measure used in secure residential facilities for youth or adults in the Netherlands, Germany, and Australia (Heynen et al., 2014, 2017; Van Miert et al., 2021). Using four scales that represent the factors 'Support', 'Growth', 'Repression', and 'Atmosphere', the GCI maps whether there is a more open or more closed perceived group climate. The Support factor refers to the extent to which staff members respond to the needs of clients and whether they invest in establishing good relationships with them.

Growth involves facilitation of learning and preparation for a meaningful life within and beyond the facility. Repression covers the extent and fairness of staff members' controlling behaviour, the amount and enforcement of rules, how much privacy is afforded to clients and the degree of boredom that exists amongst them. Finally, Atmosphere concerns the extent to which both the social and physical environment promote feelings of safety and trust in clients. Evidence has been provided for the validity and reliability of the GCI in adults with mild intellectual disability or borderline intellectual functioning in low to high security settings (Neimeijer et al., 2019), and in different age groups of clients without intellectual disability from different institutional settings (Heynen et al., 2014; Strijbosch et al., 2014, 2018; Van der Helm et al., 2011).

During the period 2016–2019, the GCI was further developed into the Group Climate Inventory—Revised (GCI-R: Van der Helm et al., 2021), based on feedback from clients and staff during its use in practice. In addition, the continued development of the GCI was driven by content-related considerations (De Valk, 2019; Sonderman et al., 2020) (see Section 2.2 for the adaptation process of the GCI-R). Ongoing research by Van der Helm and colleagues aims to explore the psychometric properties of the GCI-R in youth without intellectual disability. However, the GCI-R has not yet been validated for adults with mild intellectual disability or borderline intellectual functioning. Considering that Neimeijer et al. (2019) reported good psychometric results for this target group with its predecessor, the question remains as to whether the GCI-R produces similar psychometric results for this target group.

The GCI-R is a newly deployed measure amongst clients with mild intellectual disability or borderline intellectual functioning in secure residential facilities, with which group climate perceptions of clients are systematically assessed and monitored for quality improvement. It is an ongoing evaluation and targeted improvement process in which a valid and reliable measure may ensure that this process is consistent. Therefore, primary aim of the present study is to explore the internal structure validity and internal consistency reliability (Messick, 1989) of the GCI-R within this target group. Consonant with findings by Neimeijer et al. (2019) on the original GCI, we hypothesise that we find support for its five-factor structure, and internal consistency reliability on each scale.

2 | METHODS

2.1 | Participants

Clients were eligible to participate in the study if they were ≥18 years old; were diagnosed with mild intellectual disability or borderline intellectual functioning; and resided at Trajectum at the time of data collection.

Of the 178 clients from whom data were drawn, 30 were excluded from the analyses due to missing data of more than 20% on items within any scale of the GCI-R (see Section 2.2). This resulted in a sample of 148 participants, of whom 97 (66%) were residing

in 27 in-patient treatment wards with six to 18 beds each ($Mo = 8$ beds), and 51 (34%) in 16 specialised residence wards with four to 25 beds each ($Mo = 7$ beds).

All 148 participants were diagnosed with mild intellectual disability or borderline intellectual functioning based on data from clinical observations, questionnaires that map adaptive skills (e.g., SRZ-P: Kraijer et al., 2004; ADAPT: Jonker et al., 2021), and intelligence tests. WAIS-IV (Wechsler, 2008) was used in 45% of clients, WAIS-III (Wechsler, 1997) in 28%, WISC-III (Prifitera et al., 1998) in 7% and another intelligence test in 20%. Data on other diagnoses were available for 142 participants. One comorbid mental disorder was diagnosed in 23 and more than one in 119 participants (see Table 1 for participant demographics).

Clients with an IQ between 70 and 85 constitute a substantial part of the client population in secure facilities. In addition, they are internationally recognised to share characteristics and needs with clients with mild intellectual disability, as a result of which clients with borderline intellectual functioning are often grouped together with clients with mild intellectual disability (see e.g., Berger et al., 2023; Sauter et al., 2023; Smit et al., 2023; Van Cappellen et al., 2023). Therefore, we grouped participants with both intellectual levels into a single sample.

The current research was conducted in Trajectum, a Dutch secure residential facility that provides treatment to adult clients with mild intellectual disability or borderline intellectual functioning with

TABLE 1 Participant demographics ($N = 148$).

Variable	
Age	
Mean (SD)	39.5 (12.6)
Range	18–69
Sex (N; %)	
Male	117 (79%)
Female	31 (21%)
IQ, mean (SD)	
Total IQ	68.3 (10.7)
Verbal IQ	71.8 (12.8)
Performance IQ	71.9 (13.2)
Ward type (N; %)	
Mixed gender	120 (81%)
Male-only	28 (19%)
Diagnoses (N; %)	
Substance use disorder	72 (49%)
Psychotic disorder	63 (43%)
Neurodevelopmental disorder	57 (39%)
Personality disorder	54 (36%)
Conduct disorder	19 (13%)
Paraphilia	18 (12%)
Mood disorder	16 (11%)
Another type of disorder	21 (14%)

Abbreviations: IQ, intelligence quotient; SD, standard deviation.

comorbid behavioural and mental health problems, who have been admitted by court order. As part of a sentence to minimise recidivism, the Dutch court can order individuals to be treated in a secure residential facility for mental health problems believed to have influenced their crimes. In Trajectum, clients receive combinations of psychotherapy and cognitive behavioural therapy aimed at reducing risk factors for addiction, aggression, sexuality problems, and trauma-related problems. A significant proportion of the clients also receive support with adaptive living skills, such as practicing good hygiene.

An essential part of treatment, in addition to providing individual and group therapies, is to create a learning environment in which clients can transfer therapeutic benefit to one another through interactions and group dynamics (see e.g., De Boer-Van Schaik & Derks, 2010; Neimeijer et al., 2019).

Trajectum holds *In-patient treatment* [Klinisch behandelen] and *Specialised residence* [Specialistisch verblijf] wards, which vary in low, medium, or high levels of security, and intensity of care provided to clients. After completing treatment in an in-patient treatment ward, clients progress to a specialised residence ward, where they practice the skills they have learned in (simulated) real-life areas, such as housing and work, in close proximity to staff members.

2.2 | Instrument

2.2.1 | Group Climate Inventory—Revised

The Group Climate Inventory—Revised (GCI-R; Van der Helm et al., 2021) consists of 32 self-report items belonging to one of its five scales of Support, Growth, Peer interactions, Physical environment, and Repression. Using a five-point Likert-type scale (1 = 'strongly disagree' to 5 = 'strongly agree'), clients had to indicate how much they agreed with each item. Following Bell et al. (2018) and Neimeijer et al. (2019), the GCI-R was adapted for clients with mild intellectual disability or borderline intellectual functioning by visually supporting each answering category by smiley icons and including an 'I do not know' answering category. The descriptions of the GCI-R scales are shown in Table 2.

Each scale separately yields an average score between 1 and 5. Higher scores on each scale are more favourable, with the exception that a lower score on the Repression scale is more favourable. The GCI-R also generates a total perceived group climate score between 1 and 5, with all its items combined (after reverse-coding the items of the Repression scale). A higher total group climate score indicates a more open group climate, while a lower score indicates a more closed group climate. It takes approximately 15 minutes to complete the GCI-R.

2.2.2 | Psychometric properties for the original Group Climate Inventory

Validation studies of the GCI in adult samples (De Vries et al., 2018; Heynen et al., 2014; Neimeijer et al., 2019; Van der Helm et al., 2011)

TABLE 2 Descriptions of the scales of the Group Climate Inventory—Revised.

Scale	Number of items	Content (example item)
Support	7	Measures the support provided by staff members and their responsiveness to the specific needs of clients (e.g., '*The staff members listen to me when I want to discuss something*')
Growth	5	Measures the extent to which clients give meaning to their stay in the facility, working towards being able to participate in society (e.g., '*I learn the right things here*')
Repression	8	Evaluates the amount, fairness and enforcement of rules, and the level of privacy and boredom present amongst clients (e.g., '*I am being watched all the time*')
Peer interactions	7	Addresses the atmosphere clients experience when interacting with each other (e.g., '*We accept each other as we are at the ward*')
Physical environment	5	Addresses the atmosphere clients experience in relation to aspects of the physical environment (e.g., '*I have a nice room*')

using confirmatory factor analysis reported generally satisfactory model fit (comparative fit index: 0.82–0.93, Tucker-Lewis index: 0.81–0.92, root mean square error of approximation: 0.05–0.08), with good internal consistency reliability for Support (α: .86–.92), acceptable to good for Growth (α: .79–.88), and questionable to acceptable for Repression (α: .64–.76) and Atmosphere (α: .67–.81).

2.2.3 | Development of the Group Climate Inventory—Revised

The GCI went through an iterative development and validation process, which resulted in the GCI-R. Van der Helm and colleagues convened a series of focus groups with several clients and practitioners in three different facilities who provided feedback on the content of pre-existing scales and items (Neimeijer et al., 2019; Van der Helm et al., 2009, 2011), as well as topics from other measures under development at the time that map institutional repression (De Valk, 2019) and peer interactions (Sonderman et al., 2020). Using data from the focus groups, an item pool was created of 101 items that were included in a pilot version of the GCI-R. This pilot version was administered to 190 youth from 41 different wards in four secure youth care facilities and three youth prisons. Based on a series of confirmatory factor analyses, Van der Helm and colleagues examined item

content, factor loadings and item-total correlations. Next, they examined the model fit of a multifactor model consisting of five factors 'Support', 'Growth', 'Repression', 'Peer interactions', and 'Physical environment'. Of the 101 items, 45 were retained. Between 2018 and 2019, this 45-item version was administered to 207 youth from 42 different wards in two secure and one open youth care facilities and three youth prisons. Van der Helm and colleagues arrived at a final 32-item version from subsequent deliberations. Throughout the development process of the GCI-R, experts in the field of intellectual and developmental disabilities helped craft each item to closely match the language deficits of people with mild intellectual disability or borderline intellectual functioning and autism spectrum disorder.

Based on empirical research and feedback from clients and practitioners, the item content for each scale was revised and the Atmosphere and Repression scales were reconceptualised. In the Atmosphere scale of the GCI, interactional and environmental aspects were combined. However, when clients had to rate the Atmosphere scale, they tended to consider only the environmental aspects, but not the interactional aspects (Van der Helm et al., 2021). As the interactions amongst clients are considered an important part of the group climate quality in relation to the Atmosphere scale (Sonderman et al., 2020), the Atmosphere scale has been replaced by two separate scales in the GCI-R, namely Peer interactions and Physical environment scales. In addition, the Repression scale only included items of overt repression that reflected the authoritative threat of sanctions to enforce compliance by clients. In the operationalisation of the Repression scale in GCI-R, attention was also devoted to a more subtle form of repression as described by De Valk (2019), namely persuasive repression. As a result, Repression has been operationalised as a more complete representation of repression that includes both overt and more subtle repression (e.g., 'The staff members meddle with me all the time').

2.3 | Procedure

Within Trajectum, group climate data are systematically collected to provide performance-related feedback to wards, primarily aimed at evaluating and improving the treatment of the clients. The current study uses data collected between September 2020 and October 2021 using the GCI-R (see Section 2.2) from voluntarily participating clients across the organisation's wards. A total of 31.2% of the clients participated in the data collection.

The GCI-R was administered to clients by trained research assistants, who read aloud the items and answering categories, showed clients the smiley icons corresponding to each answering category, and then recorded the oral ratings clients gave of the items. Research assistants were independent of participating wards and treatment of clients as Kooijmans et al. (2022) recommend minimising potential biases in the administration of self-report measures. As they were not allowed to enter the wards due to the COVID-19 pandemic, the measure was administered by the research assistants via a secure video calling application used by therapists in the organisation to treat clients during that time. The data were stored on a secure Trajectum network and later retrieved for use in the current study.

Clients were asked to participate in the data collection and to consent to their data being used for any research purpose, in addition to providing performance-related feedback to wards. Oral informed consent was obtained from all clients. Their therapists had conversations with them beforehand to explain what the requested informed consent entailed. Written informed consent was provided by clients, or legal representatives of clients who were deemed legally incapable by the court to provide written informed consent themselves.

Ethical approval was granted by the Ethics Committee of the Faculty of Social Sciences of the Radboud University (ECSW-2021-084).

2.4 | Data analysis

We performed confirmatory factor analysis (CFA) for three models using *lavaan* (Rosseel, 2012, 2022) in R (R Core Team, 2022) to validate the hypothesised five-factor structure of the GCI-R (Hayduk et al., 2007; McIntosh, 2007). The current study only included clients who completed at least 80% of all items within each scale of the GCI-R to minimise possible biases due to missing data (Van der Helm et al., 2021). 'I do not know' answers were treated as missing. The overall percentage of missing data in 148 clients across the 32 GCI-R items was 1.2% ($N = 58$ out of 4736 possible values). Little's MCAR test indicated that data were not missing completely at random ($\chi^2 = 1061.088$; $p < .001$). However, based on the small number of missing values and further inspection of the data, missing at random
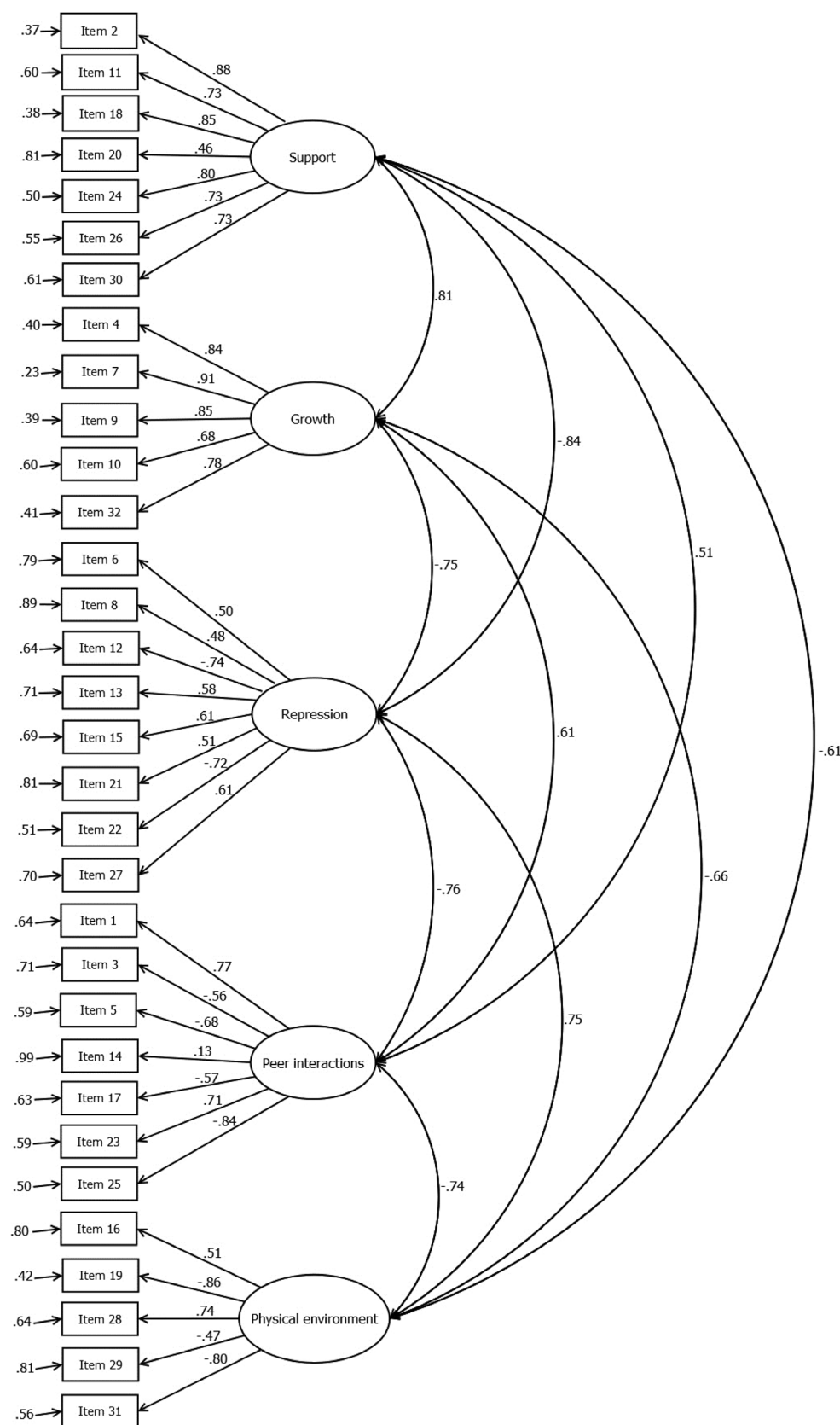
TABLE 3 Results from the confirmatory factor analyses with the Group Climate Inventory—Revised.

Model	χ^2 (df)	RMSEA (90% CI)	CFI	TLI	SRMR	p
5-factor model—32 items	670.956 (454)	0.057 (0.048–0.066)	0.987	0.985	0.084	<.001
5-factor model—31 items (w/o 14)	644.600 (424)	0.059 (0.050–0.069)	0.986	0.985	0.084	<.001
4-factor model—32 items	728.385 (458)	0.063 (0.055–0.072)	0.983	0.982	0.087	<.001

Note: Item 14 = 'I can discuss personal matters with a client from my ward'. The five-factor model consists of the factors 'Support', 'Growth', 'Repression', 'Peer interactions', and 'Physical environment'. In the four-factor model, the Peer interactions and Physical environment factors were combined into one factor (as in the original Group Climate Inventory), to compare model fit between models.
Abbreviations: CFI, comparative fit index; CI, confidence interval; df, degrees of freedom; RMSEA, root mean square error of approximation; SRMR, standardised root mean residual; TLI, Tucker-Lewis index; χ^2, Chi-square test of model fit.

FIGURE 1 Path diagram of the 32-item model tested in the confirmatory factor analysis. Items (i.e., variables) are represented by rectangles, and scales (i.e., factors) by ellipses. The arrows pointing from the ellipses to the rectangles represent standardised factor loadings, and arrows pointing to the rectangles from the left represent standardised residual variances. The arrows between the ellipses represent standardised estimates of the covariances.

(MAR) was assumed. Pairwise deletion was used to handle missing data, which is considered a reasonable approach when missing data are MAR, as it makes efficient use of available data, and would not introduce systematic bias into our study as there appears to be no relationship between the missing data and the variables of interest (Kline, 2016; Muthén et al., 1987).

To account for non-normally distributed ordinal variables, the mean and variance-adjusted weighted least squares (WLSMV) estimation procedure was chosen (Li, 2015). Modification indices, giving the expected drop in Chi-square if the parameter in question is freely estimated, were used to improve model fit. Model 1, defined a priori, included all 32 items of the GCI-R (five-factor model: 'Support', 'Growth', 'Repression', 'Peer interactions', and 'Physical environment'). In model 2, items with insufficient loading (standardised loading ≤ 0.4; Hair et al., 2010) were removed to achieve optimal model fit. Next, a four-factor model (model 3) was specified in which the factors Peer interactions and Physical environment were combined into one factor, as in the original GCI, to examine whether the five-factor model demonstrated better model fit compared to the four-factor model.

Multiple indices were used to examine model fit, including the minimum fit function χ^2 statistic, the Tucker-Lewis index (*TLI*), the comparative fit index (*CFI*), the standardised root mean square residual (*SRMR*), and the root mean square error of approximation (*RMSEA*). Small and nonsignificant χ^2 values, CFI and TLI values ≥ 0.95, RMSEA ≤ 0.06 and SRMR ≤ 0.08 indicate a good model fit, and CFI and TLI between 0.90 and 0.95, RMSEA ≤ 0.08 and SRMR ≤ 0.10 indicate an acceptable model fit (Hu & Bentler, 1999; MacCallum et al., 1996). Unlike the other fit indices used, χ^2 is sensitive to sample size, as a result of which it is not considered a realistic test to determine model fit (Byrne, 2010; Schermelleh-Engel et al., 2003).

The internal consistency reliability of the scales was examined by calculating ordinal α (Zumbo et al., 2007), construct reliability (also referred to as McDonald's Omega [ω; McDonald, 1999]), and Cronbach's α. Values larger than .80 are indicative of good construct reliability (Cheung et al., 2023; Cicchetti, 1994).

Intraclass correlation coefficients (ICC)(1), ICC(2), and within-group rater agreement for multi-item measures (rWG(*J*); James et al., 1984) were calculated for the GCI-R scale scores using the *bruceR* package (Bao, 2023). ICC(1) refers to the proportion of variance explained by groups (i.e., heterogeneity between groups), ICC(2) refers to the extent to which an individual score can be considered a reliable assessment of a group-level construct, and rWG (*J*) is a measure of within-group agreement across all items of a scale. ICC(2) and rWG(*J*) ≥ 0.70 are generally accepted as good (Bliese, 2000).

3 | RESULTS

3.1 | Evidence of validity based on the internal structure of each scale

A five-factor model (model 1) with all 32 items of the GCI-R showed good fit for RMSEA, CFI and TLI, acceptable fit for SRMR and inadequate fit for χ^2. In model 2, item 14 of the Peer interactions scale was removed as this item had insufficient factor loading ('I can discuss personal matters with a client from my ward'; standardised loading = 0.126) (Hair et al., 2010). However, model 1 was retained as model 2 did not improve model fit. In addition, model fit was compared between the four-factor model (model 3) and five-factor model (model 1). Results indicated that the five-factor model provided a significantly better fit to the data ($\Delta\chi^2$ (df) = 22.95 (4), *p* < .001). The results of the confirmatory factor analyses are shown in Table 3. A path diagram of estimates of the standardised factor loadings of model 1 with all 32 items is shown in Figure 1.

3.2 | Evidence of reliability

Table 4 shows the estimates of the resulting acceptable to good internal consistency reliabilities of the scales.

TABLE 4 Cronbach's alpha, ordinal alpha, McDonald's omega, and average variance extracted values for each scale of the Group Climate Inventory—Revised resulting from the confirmatory factor analysis with 32 items.

Scale	α	Ordinal α	ω	AVE
Support	.857	.893	.853	0.563
Growth	.868	.898	.856	0.665
Repression	.761	.813	.763	0.363
Peer interactions	.730	.794	.756	0.417
Physical environment	.715	.768	.756	0.481

Abbreviations: AVE, average variance extracted; α, Cronbach's alpha; ω, McDonald's omega.

TABLE 5 Descriptive statistics of the scales of the Group Climate Inventory—Revised (*N* = 148).

Scale	ICC(1)	ICC(2)	rWG(*J*)	Min.	Max.	M (SD)	Skewness (SE)	Kurtosis (SE)
Support	0.03	0.08	0.90	1.86	5.00	3.82 (0.56)	−1.017 (0.199)	2.194 (0.396)
Growth	0.00	0.00	0.85	1.60	5.00	3.67 (0.73)	−1.062 (0.199)	.898 (0.396)
Repression	0.06	0.18	0.85	1.38	4.25	2.46 (0.55)	.632 (0.203)	.301 (0.403)
Peer interactions	0.14	0.34	0.89	2.14	4.71	3.49 (0.54)	−.269 (0.200)	−.435 (0.397)
Physical environment	0.41	0.67	0.89	2.20	5.00	3.79 (0.64)	−.699 (0.199)	−.181 (0.396)

Note: On all scales, a higher score indicates a more favourable outcome, except for Repression, on which a lower score indicates a more favourable outcome.

Abbreviations: ICC, intraclass correlation coefficient; rWG(*J*), within-group rater agreement for multi-item measures; SD, standard deviation; SE, standard error.

3.3 | Descriptive data of the scores

The minimum and maximum scores, means and standard deviations, skewness, and kurtosis values for all scales and items are presented in Tables 5 and 6, respectively. The skewness and kurtosis values were all within an acceptable range for normality (Hair et al., 2010; Kline, 2016). ICCs and rWG(J) values for the scales are also reported (see Table 5). These findings suggest that the percentage of the variance in scores attributable to the group (i.e., ward) in which clients reside varies between scales. ICC(2), referring to the reliability of

TABLE 6 Descriptive statistics of the items of the Group Climate Inventory—Revised (N = 148).

Item no.	Scale/item	Min.	Max.	Mean (SD)	Skewness (SE)	Kurtosis (SE)
	Support					
2	_The staff members help me to achieve my (treatment) goals._	1	5	3.83 (0.82)	−1.137 (0.201)	1.770 (0.399)
11	_The staff members spend enough time with me._	1	5	3.57 (0.87)	−1.015 (0.199)	.515 (0.396)
18	_The staff members listen to me when I want to discuss something._	1	5	3.91 (0.75)	−1.020 (0.199)	1.954 (0.396)
20	_The staff members compliment me when I do something right._	2	5	4.03 (0.64)	−.818 (0.199)	2.192 (0.396)
24	_The staff members help me when I ask them for help._	1	5	3.86 (0.75)	−1.244 (0.200)	2.272 (0.397)
26	_The staff members keep their appointments with me._	1	5	3.70 (0.78)	−1.096 (0.199)	1.600 (0.396)
30	_I can discuss personal matters with the staff members._	1	5	3.86 (0.76)	−1.274 (0.201)	2.729 (0.399)
	Growth					
4	_I receive the treatment that I need._	1	5	3.60 (0.92)	−.876 (0.201)	.244 (0.399)
7	_What I learn at the ward is helping me._	1	5	3.65 (0.87)	−.925 (0.200)	.478 (0.397)
9	_I learn the right things here._	1	5	3.69 (0.83)	−1.150 (0.200)	.748 (0.397)
10	_I find my treatment useful._	1	5	3.57 (0.93)	−.960 (0.199)	.308 (0.396)
32	_What I learn here will help me when I am outside._	1	5	3.89 (0.87)	−1.132 (0.203)	1.422 (0.403)
	Repression					
6	_The staff members meddle with me all the time._	1	5	2.43 (0.88)	.966 (0.200)	.217 (0.397)
8	_The entire ward gets into trouble with the staff if one person does something wrong._	1	4	2.09 (0.74)	1.127 (0.202)	1.765 (0.401)
12	_The rules are fair on the ward._	1	5	3.46 (0.88)	−.725 (0.201)	−.560 (0.400)
13	_I am being watched all the time._	1	5	2.48 (0.92)	.948 (0.200)	−.047 (0.397)
15	_The staff members always get their way._	1	5	2.83 (0.98)	.349 (0.200)	−1.152 (0.397)
21	_I am sent to my room by the staff for no reason._	1	5	2.04 (0.75)	1.320 (0.201)	2.817 (0.399)
22	_I feel that my complaints are taken seriously by the staff members._	1	5	3.51 (0.94)	−.818 (0.202)	−.165 (0.401)
27	_The staff members are always making up new rules._	1	5	2.67 (0.95)	.294 (0.202)	−1.001 (0.401)
	Peer interactions					
1	_At the ward there is a nice atmosphere._	1	5	3.51 (0.88)	−.512 (0.199)	−.081 (0.396)
3	_The clients from my ward threaten each other._	1	5	2.44 (0.91)	.753 (0.200)	−.311 (0.397)
5	_The clients from my ward are teaching each other bad things._	1	5	2.52 (0.90)	.588 (0.203)	−.536 (0.403)
14	_I can discuss personal matters with a client from my ward._	1	5	3.14 (0.94)	−.278 (0.201)	−1.369 (0.399)
17	_The clients provoke each other at the ward._	1	5	2.67 (0.90)	.524 (0.201)	−.717 (0.399)
23	_We accept each other as we are at the ward._	2	5	3.70 (0.72)	−.717 (0.201)	.477 (0.400)
25	_At the ward, some clients are excluded from the group._	1	5	2.29 (0.88)	1.085 (0.204)	.850 (0.406)
	Physical environment					
16	_Many things are broken at the ward._	1	5	2.27 (0.82)	1.169 (0.200)	1.377 (0.397)
19	_It looks cosy at the ward._	1	5	3.56 (0.93)	−.840 (0.200)	−.155 (0.397)
28	_It is dirty at the ward._	1	5	2.20 (1.06)	.571 (0.199)	−.735 (0.396)
29	_I have a nice room._	1	5	4.11 (0.93)	−1.210 (0.199)	1.071 (0.396)
31	_There is a nice garden or courtyard at the ward._	1	5	3.73 (0.87)	−1.044 (0.201)	0.589 (0.399)

Abbreviations: SD, standard deviation; SE, standard error.

scores as a group construct, also varied across scales. ICC(2) was moderate to high for Physical environment and low for the other scales. rWG(J) values indicated good within-group agreement in scores on items of all GCI-R scales.

Pearson correlation coefficients indicating significant moderate to strong correlations between the scales of the GCI-R in the expected direction are shown in Table 7.

4 | DISCUSSION

This study aimed to examine the internal structure validity of the Group Climate Inventory—Revised (GCI-R) and the internal consistency reliability of its scales using confirmatory factor analysis in a sample of adults with mild intellectual disability or borderline intellectual functioning in a Dutch secure residential facility. The results provide support for the presence of the five-factor structure of the GCI-R in this target group ('Support', 'Growth', 'Repression', 'Peer interactions', and 'Physical environment'). Next, acceptable to good internal consistency reliabilities of the scales of the GCI-R were found, which are largely consistent with previous research on the GCI (De Vries et al., 2018; Heynen et al., 2014; Neimeijer et al., 2019; Strijbosch et al., 2014, 2018; Van der Helm et al., 2011).

Over the past decade, increasing attention has been paid to assessing group climate in secure residential facilities (Doyle et al., 2017; Leipoldt et al., 2019; Tonkin, 2016). To this end, measures have been developed and validated that map the group climate from the perspective of clients, giving clients a voice, for example, about how they perceive support, repression, learning opportunities and other elements that constitute group climate. In addition, these measures may provide performance-related feedback to treatment units, allowing treatment units to continue or adjust treatment. Research has highlighted the importance of using client-reported data in secure settings to achieve an open group climate (Levrouw et al., 2018), which is imperative for good treatment. For example, Billen (2022) showed that an open group climate resulted in good self-regulation skills in clients. Good self-regulation skills in clients led to fewer incidents, allowing staff members to better maintain a balance between flexibility and control that is necessary to maintain an open group climate.

Within the context of measuring group climate, our findings support the use of the GCI-R with clients with mild intellectual disability or borderline intellectual functioning in secure residential facilities. In addition, our findings indicated high within-group agreement on the items across all scales, suggesting that aggregated scores at the ward level may provide reliable data to use within wards for discussing group climate. Measuring group climate using the GCI-R and discussing the results with clients and staff can function as an intervention to create greater awareness and improvement plans in daily practice. Considering that perceptions of group climate may differ between clients and staff, conversations between clients and staff about GCI-R results should be encouraged as meaning can be given to results, and goals can be set with which group climate perceptions may be improved. The process of measuring, providing feedback, discussing results, and making improvement plans can be described as a Plan-Do-Check-Act cycle (Deming, 1986) that encourages staff members to reflect on their professional behaviour and adapt it where needed (Neimeijer, 2021).

Our study is a valuable addition to intellectual disability research and practice, as it examined the psychometric qualities of the GCI-R in adults with mild intellectual disability or borderline intellectual functioning and found encouraging results for this target group. However, there are some limitations that should be mentioned. First, group climate was measured in clients using a self-report measure, which may have influenced the results of the current study due to problems inherent in self-reporting, such as social desirability or recall bias. Second, the use of a convenience sample from one facility to conduct this study may have compromised the validity of our results, as the sample used may not be representative of the general population of adult clients with mild intellectual disability or borderline intellectual functioning in secure residential facilities. Third, as we used data in this study that had already been collected within the facility for quality improvement, data were not available from other possible measures to investigate the convergent validity of the GCI-R, such as the EssenCES (Schalast et al., 2008). In addition, we had no possibility to collect data within the facility to test convergent validity. However, a study by De Vries et al. (2018) reported significant moderate (r: .36–.49) and strong correlations (r: .58–.83) between the original GCI and EssenCES scales, suggesting that both measures assess related climate dimensions. Similarly, criterion validity and test–retest

TABLE 7 Pearson correlation coefficients between the scales of the Group Climate Inventory—Revised ($N = 148$).

	Support	Growth	Repression	Peer interactions	Physical environment
Support	-				
Growth	.696*	-			
Repression	−.684*	−.564*	-		
Peer interactions	.390*	.450*	−.557*	-	
Physical environment	.452*	.490*	−.574*	.577*	-

Note: On all scales, a higher score indicates a more favourable outcome, except for Repression, on which a lower score indicates a more favourable outcome.

*$p < .001$ (2-tailed).

reliability were not tested. Fourth, following the ICCs that indicated mixed results, performing multilevel factor analysis would have provided detailed information at both a client and ward level (Bosma et al., 2020; Van Ginneken & Nieuwbeerta, 2020). However, the sample size in the current study was too small to perform multilevel analysis. Due to limited sample size, we also could not verify measurement invariance of the GCI-R across participants with mild intellectual disability and participants with borderline intellectual functioning.

4.1 | Conclusion

The psychometric analysis suggests that the Group Climate Inventory—Revised demonstrates evidence of psychometric adequacy when applied to adults with mild intellectual disability or borderline intellectual functioning in secure residential facilities. However, due to the limitations of the study, our results should be interpreted with caution. Future studies should improve on the current study by conducting research using a larger sample drawn from multiple facilities. This may strengthen the generalisability and validity of our findings by providing multilevel psychometric results for GCI-R and enable investigation of measurement invariance across clients with mild intellectual disability and clients with borderline intellectual functioning. Furthermore, future research is needed into the convergent validity, criterion validity and test–retest reliability of the GCI-R for this target group.

CONFLICT OF INTEREST STATEMENT
The authors declare that there are no conflicts of interest.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

PATIENT CONSENT
Yes, explanation is provided in the methods section.

ORCID
A. Turhan [ID] https://orcid.org/0000-0001-8923-3350
M. J. Delforterie [ID] https://orcid.org/0000-0001-7987-4664
E. G. Neimeijer [ID] https://orcid.org/0000-0002-6119-7287

REFERENCES
Bakken, T. L., Røssberg, J. I., & Friis, S. (2012). The ward atmosphere scale for psychiatric inpatients with intellectual disability: A pilot study. *Advances in Mental Health and Intellectual Disabilities*, 6, 265–272. https://doi.org/10.1108/20441281211261159

Bao, H. W. S. (2023). *bruceR: Broadly useful convenient and efficient R functions*. R package version 0.8.x.

Barker, L., McKeown, A., Small, M., & Meggs, J. (2021). Validating the Essen Climate Evaluation Schema modified for people with learning disabilities in a low-risk secure forensic setting. *Criminal Behaviour and Mental Health*, 31, 143–150. https://doi.org/10.1002/cbm.2175

Bell, N. (2022). *Adapting the Essen Climate Evaluation Schema (EssenCES) for use in forensic Intellectual Developmental Disorder (IDD) services*. Doctoral Dissertation. Birmingham City University.

Bell, N., Tonkin, M., Chester, V., & Craig, L. (2018). Adapting measures of social climate for use with individuals with intellectual developmental disability in forensic settings. *Psychology, Crime & Law*, 24, 362–378. https://doi.org/10.1080/1068316X.2017.1298761

Berger, I., Bruineberg, A., Van Ewijk, M., De Jong, L., Van der Hout, M., Van Weeghel, J., & Van der Meer, L. (2023). Developing a recovery-oriented intervention for people with severe mental illness and an intellectual disability: Design-oriented action research. *Frontiers in Psychiatry*, 14, 1184798. https://doi.org/10.3389/fpsyt.2023.1184798

Billen, E. (2022). *Improving the understanding of self-regulation: Perspectives for forensic research and practice*. Doctoral dissertation. Tilburg University.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 349–381). Jossey-Bass, Inc.

Bosma, A. Q., Van Ginneken, E., Palmen, H., Pasma, A. J., Beijersbergen, K. A., & Nieuwbeerta, P. (2020). A new instrument to measure prison climate: The psychometric quality of the prison climate questionnaire. *The Prison Journal*, 100, 355–380. https://doi.org/10.1177/0032885520916

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Routledge.

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2023). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*. https://doi.org/10.1007/s10490-023-09871-y

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

Day, A., Casey, S., Vess, J., & Huisy, G. (2012). Assessing the therapeutic climate of prisons. *Criminal Justice and Behaviour*, 39, 156–168. https://doi.org/10.1177/0093854811430476

De Boer-Van Schaik, J., & Derks, F. (2010). The Van der Hoeven clinic: A flexible and innovative forensic psychiatric hospital based on therapeutic community principles. In R. Shuker & E. Sullivan (Eds.), *Grendon and the emergence of forensic therapeutic communities: Developments in research and practice* (pp. 45–60). John Wiley & Sons. https://doi.org/10.1002/9780470661444.ch3

De Valk, S. M. (2019). *Under pressure: Repression in residential youth care*. University of Amsterdam.

De Vries, M. G., Brazil, I. A., Van der Helm, G. H. P., Verkes, R.-J., & Bulten, B. H. (2018). Ward climate in a high-secure forensic psychiatric setting: Comparing two instruments. *The International Journal of Forensic Mental Health*, 17, 247–255. https://doi.org/10.1080/14999013.2018.1478915

Deming, W. E. (1986). *Out of the crisis*. MIT Press.

Dickens, G. L., Johnson, A., Everett, B., & Tonkin, M. (2022). Interventions to improve social climate in acute mental health inpatient settings: Systematic review of content and outcomes. *SAGE Open Nursing*, 8, 1–21. https://doi.org/10.1177/23779608221124291

Doyle, P., Quayle, E., & Newman, E. (2017). Social climate in forensic mental health settings: A systematic review of qualitative studies. *Aggression and Violent Behaviour*, *36*, 118–136. https://doi.org/10.1016/j.avb.2017.06.008

Eltink, E. M. A. (2020). *Back to basic: Relations between residential group climate and juvenile antisocial behaviour*. Doctoral dissertation. University of Amsterdam.

Hair, J., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson Educational International.

Harding, R. (2014). Rehabilitation and prison social climate: Do 'what works' rehabilitation programs work better in prisons that have a positive social climate? *Australian & New Zealand Journal of Criminology*, *47*, 163–175. https://doi.org/10.1177/0004865813518543

Hayduk, L. A., Cummings, G. C., Boadu, K., Pazderka-Robinson, H., & Boulianne, S. (2007). Testing! Testing! One, two, three—Testing the theory in structural equation models! *Personality and Individual Differences*, *42*, 841–850.

Heynen, E., Van Der Helm, G. H. P., Cima, M., Stams, G. J. J. M., & Korebrits, A. (2017). The relation between living group climate, aggression, and callous-unemotional traits in delinquent boys in detention. *International Journal of Offender Therapy and Comparative Criminology*, *61*, 1701–1718. https://doi.org/10.1177/0306624X16630543

Heynen, E. J. E., Van der Helm, G. H. P., Stams, G. J. J. M., & Korebrits, A. M. (2014). Measuring group climate in a German youth prison: A German validation of the prison group climate instrument. *Journal of Forensic Psychology Practice*, *14*, 45–54. https://doi.org/10.1080/15228932.2013.868176

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. https://doi.org/10.1080/10705519909540118

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, *69*, 85–98. https://doi.org/10.1037/0021-9010.69.1.85

Jonker, F. J. R., Didden, R., Goedhard, L. E., Korzilius, H. P. L. M., & Nijman, H. L. I. (2021). The Adaptive Ability Performance Test (ADAPT): A new instrument for measuring adaptive skills in people with intellectual disabilities and borderline intellectual functioning. *Journal of Applied Research in Intellectual Disabilities*, *34*, 1156–1165. https://doi.org/10.1111/jar.12876

Kline, R. B. (2016). *Methodology in the social sciences. Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Kooijmans, R., Langdon, P. E., & Moonen, X. (2022). Assisting children and youth with completing self-report instruments introduces bias: A mixed-method study that includes children and young people's views. *Methods in Psychology*, *7*, 100102. https://doi.org/10.1016/j.metip.2022.100102

Kraijer, D. W., Kema, G. N., & De Bildt, A. A. (2004). *Handleiding SRZ en SRZ-i Sociale Redzaamheidsschalen (6^e uitgave) [manual of SRZ social self-reliance scales]* (6th ed.). Harcourt Test Publishers.

Leipoldt, J. D., Harder, A. T., Kayed, N., Grietens, H., & Rimehaug, T. (2019). Determinants and outcomes of social climate in therapeutic residential youth care: A systematic review. *Children and Youth Services Review*, *99*, 429–440. https://doi.org/10.1016/j.childyouth.2019.02.010

Levrouw, D., Roose, R., Van der Helm, G. H. P., Strijbosch, E. L. L., & Vandevelde, S. (2018). Developing a positive living group climate in residential youth care: A single case study. *Child & Family Social Work*, *23*, 709–716. https://doi.org/10.1111/cfs.12467

Li, C. (2015). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behaviour Research Methods*, *48*, 936–949. https://doi.org/10.3758/s13428-015-0619-7

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130–149. https://doi.org/10.1037/1082-989X.1.2.130

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates Publishers.

McGee, M. G., & Woods, D. J. (1978). Use of Moos' Ward Atmosphere Scale in a residential setting for mentally retarded adolescents. *Psychological Reports*, *43*, 580–582. https://doi.org/10.2466/pr0.1978.43.2.580

McIntosh, C. N. (2007). Rethinking fit assessment in structural equation modelling: A commentary and elaboration on Barrett (2007). *Personality and Individual Differences*, *42*, 859–867.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). Macmillan Publishing Co, Inc; American Council on Education.

Moos, R. H. (1968). The assessment of the social climates of correctional institutions. *Journal of Research in Crime & Delinquency*, *5*, 174–188. https://doi.org/10.1177/002242786800500207

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, *52*, 431–462. https://doi.org/10.1007/BF02294365

Neimeijer, E. G. (2021). *Close(d) care. Group climate in a secure forensic setting for individuals with mild intellectual disability*. Doctoral Dissertation. Radboud University.

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

Neimeijer, E. G., Roest, J. J., Van der Helm, G. H. P., & Didden, R. (2019). Psychometric properties of the group climate instrument (GCI) in individuals with mild intellectual disability or borderline intellectual functioning. *Journal of Intellectual Disability Research*, *63*, 215–224. https://doi.org/10.1111/jir.12567

O'Flynn, P., O'Regan, R., O' Reilly, K., & Kennedy, H. (2018). Predictors of quality of life among inpatients in forensic mental health: Implications for occupational therapists. *BMC Psychiatry*, *18*, 16. https://doi.org/10.1186/s12888-018-1605-2

Prifitera, A., Weiss, L. G., & Saklofske, D. H. (1998). The WISC—III in context. In A. Prifitera & D. H. Saklofske (Eds.), *WISC—III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 1–38). Academic Press. https://doi.org/10.1016/B978-012564930-8/50002-4

Puzzo, I., Aldridge-Waddon, L., Morley, E., Vacher, J., Mitchell, J., Murphy, D., & Farr, C. (2023). Developing ward social climate and sense of community within a high security forensic psychiatric service: Evaluating a sense of community and social climate intervention. *The Journal of Forensic Psychiatry & Psychology*, *34*, 20–36. https://doi.org/10.1080/14789949.2023.2168205

Quinn, M., Thomas, C., & Chester, V. (2012). The Essen Climate Evaluation Schema measure of social climate in a secure service for people with intellectual disabilities. *Advances in Mental Health and Intellectual Disabilities*, *6*, 171–178. https://doi.org/10.1108/20441281211236562

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Robinson, J. E., Craig, L. A., & Tonkin, M. (2018). Perceptions of social climate and aggressive behaviour in forensic services: A systematic review. *Trauma, Violence, and Abuse*, *19*, 391–405. https://doi.org/10.1177/1524838016663936

Roest, J. J. (2022). *The therapeutic alliance in child and adolescent psychotherapy and residential youth care*. Doctoral dissertation. University of Amsterdam.

Røssberg, J. I., & Friis, S. (2003). A suggested revision of the Ward Atmosphere Scale. *Acta Psychiatrica Scandinavica*, *108*, 374–380.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Rosseel, Y. (2022). *Package 'lavaan' version 0.6-12*.

Sameroff, A. J. (2009). The transactional model. In A. J. Sameroff (Ed.), *The transactional model of development* (pp. 3–21). American Psychological Association.

Sauter, F. M., Van den Bogaard, M., Van Vliet, C., & Liber, J. M. (2023). An AAIDD-informed framework for cognitive behavioural case formulation and cognitive behaviour therapy for young people with mild intellectual disabilities or borderline intellectual functioning. *Clinical Psychology: Science and Practice*, 30, 299–311. https://doi.org/10.1037/cps0000151

Schaftenaar, P. (2018). *Aiming at contact. Relational caring and the everyday interaction as effective principles in clinical forensic care*. Doctoral dissertation. University of Humanistic Studies.

Schalast, N., Redies, M., Collins, M., Stacey, J., & Howells, K. (2008). EssenCES, a short questionnaire for assessing the social climate of forensic psychiatric wards. *Criminal Behaviour and Mental Health*, 18, 49–58. https://doi.org/10.1002/cbm.677

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8, 23–74.

Smit, M. J., Emck, C., Scheffers, M., Van Busschbach, J. T., & Beek, P. J. (2023). The impact of sexual abuse on body experience in adults with mild intellectualdisability or borderline intellectual functioning. *Journal of Intellectual & Developmental Disability*, 48, 324–333. https://doi.org/10.3109/13668250.2022.2158725

Sonderman, J., Bekken, F. F., Van der Helm, G. H. P., Roest, J. J., Kuiper, C. H. Z., & Van de Mheen, D. (2020). Peer interactions in residential youth care: A validation study of the peer interactions in residential youth care (PIRY) questionnaire. *Residential Treatment for Children and Youth*, 3, 239–260. https://doi.org/10.1080/0886571X.2020.1787924

Stams, G. J. J. M., & Van der Helm, G. H. P. (2017). What works in residential programs for aggressive and violent youth? Treating youth at risk for aggressive and violent behaviour in (secure) residential care. In P. Sturmey (Ed.), *The Wiley handbook of violence and aggression* (p. 4). John Wiley & Sons.

Strijbosch, E. L. L., Van der Helm, G. H. P., Stams, G. J. J. M., & Wissink, I. B. (2018). Young children (4-8 years) in group care: Development and validation of a group climate instrument. *Child Indicators Research*, 11, 1769–1780. https://doi.org/10.1007/s12187-017-9496-1

Strijbosch, E. L. L., Van der Helm, G. H. P., Van Brandenburg, M. E. T., Mecking, M., Wissink, I. B., & Stams, G. J. J. M. (2014). Children in residential care: Development and validation of a group climate instrument. *Research on Social Work Practice*, 24, 462–469. https://doi.org/10.1177/1049731513510045

Tomlin, J., & Tonkin, M. (2023). The EssenCES measure of ward atmosphere: Mokken scaling, confirmatory factor analysis, and investigating patient-level characteristics. *International Journal of Forensic Mental Health*, 22, 199–209. https://doi.org/10.1080/14999013.2022.2134946

Tonkin, M. (2016). A review of questionnaire measures for assessing the social climate in prisons and forensic psychiatric hospitals. *International Journal of Offender Therapy and Comparative Criminology*, 60, 1376–1405. https://doi.org/10.1177/0306624X15578834

Van Cappellen, S. M., Kühl, E., Schuiringa, H. D., Matthys, W., & Van Nieuwenhuijzen, M. (2023). Social information processing, normative beliefs about aggression and parenting in children with mild intellectual disabilities and aggressive behaviour. *Research in Developmental Disabilities*, 136, 104468. https://doi.org/10.1016/j.ridd.2023.104468

Van der Helm, G. H. P., Klapwijk, M., Stams, G. J. J. M., & Van der Laan, P. (2009). "What works" for juvenile prisoners: The role of group climate in a youth prison. *Journal of Children's Services*, 4, 36–48. https://doi.org/10.1108/17466660200900011

Van der Helm, G. H. P., Kuiper, C. H. Z., & Stams, G. J. J. M. (2018). Group climate and treatment motivation in secure residential and forensic youth care from the perspective of self determination theory. *Children and Youth Services Review*, 93, 339–344. https://doi.org/10.1016/j.childyouth.2018.07.028

Van der Helm, G. H. P., Roest, J., Dekker, A. L., Van Miert, V. S. L., & Klapwijk, G. (2021). *Handleiding herziene versie leefklimaatvragenlijst Group Climate Inventory—Revisited (GCI-R) [manual Group Climate Inventory—Revised]*. University of Applied Sciences Leiden.

Van der Helm, G. H. P., Stams, G. J. J. M., & Van der Laan, P. H. (2011). Measuring group climate in prison. *The Prison Journal*, 91, 158–176. https://doi.org/10.1177/0032885511403595

Van Ginneken, E. F. J. C., & Nieuwbeerta, P. (2020). Climate consensus: A multilevel study testing assumptions about prison climate. *Journal of Criminal Justice*, 69, 101693. https://doi.org/10.1016/j.jcrimjus.2020.101693

Van Miert, V. S. L., Dekker, A. L., Van der Helm, G. H. P., & Robinson, M. (2021). Young people's perception of group climate in juvenile justice centres in an Australian state, a pilot study. *International Journal of Offender Therapy and Comparative Criminology*, 66, 856–875. https://doi.org/10.1177/0306624X211022664

Wechsler, D. (1997). *Wechsler adult intelligence scale—third edition (WAIS-III) [database record]*. APA PsycTests. https://doi.org/10.1037/t49755-000

Wechsler, D. (2008). *Wechsler adult intelligence scale—fourth edition (WAIS-IV) [database record]*. APA PsycTests. https://doi.org/10.1037/t15169-000

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*, 6, 21–29. https://doi.org/10.22237/jmasm/1177992180