

Exploratory validation of the IWBMC™ spirit and will scale: Internal consistency, factor structure, and construct validity in measuring psychological resilience

American Journal of Social Sciences and Humanities

Vol. 11, No. 2, 86-109, 2026

e-ISSN: 2226-4809/p-ISSN: 2304-6945



Corresponding Author

- Amy O'Dell¹
- Holly A. Haynes^{1,2,3}
- Julian Gonzalez^{1,2}
- Phillip L. Haynes^{4,5}
- Blanca Meyer³
- Allegra Saunders¹
- Jayden Gaynor⁶

¹The Jacob's Ladder Group, United States.

¹Email: amy@thejacobs ladder group.org

^{1,2}Email: holly.haynes@fresno.edu

^{1,2}Email: juliang@thejacobs ladder group.org

¹Email: allegra@thejacobs ladder group.org

²Pillar Research Institute, United States.

³Fresno Pacific University, United States.

³Email: blanca.meyer@fresno.edu

⁴Crimson Foundation, United States.

^{4,5}Email: phillip_haynes@hotmail.com

⁴Georgian Court University, United States.

⁶Juniata College, United States.

⁶Email: jaydengaynor05@gmail.com

ABSTRACT

The Interpersonal Whole-Brain Model of Care® (IWBMC™) conceptualizes psychological resilience through Spirit (Identity, connection, emotional regulation) and Will (Growth-oriented persistence and cognition). This exploratory validation examined the 60-item IWBMC Spirit & Will Assessment Scale in 207 community participants and assessed internal consistency, factor structure, and preliminary construct validity. Parallel analysis suggested six factors; however, because a sixth factor had fewer than three items with clear primary loadings, we interpreted a 5-factor solution. Total scale reliability was excellent ($\alpha = 0.95$). Domain-level reliability ranged from acceptable to good: Core Identity ($\alpha = 0.88$), Social Connection ($\alpha = 0.80$), Emotional Acceptance & Regulation ($\alpha = 0.75$), Emotional Capacity for Growth & Safety ($\alpha = 0.87$), and Cognitive Mindset ($\alpha = 0.85$) met standards. Key findings include: (a) Will domains (SW4, SW5) merged empirically, suggesting a unified rather than a dual-component structure; (b) safety-related items clustered distinctly; and (c) Social Connection items dispersed across factors. Significant sex differences emerged on Will-related domains, with males scoring higher (SW4, SW5, and Will composite; $d = 0.30$ - 0.39). Two small but significant race/ethnicity effects emerged for SW5 and the Will composite ($\eta^2 = 0.062$ - 0.067), driven primarily by lower scores among participants preferring not to disclose their race/ethnicity. No significant age differences were found (all p s > 0.10 , all $\eta^2 < 0.037$). These results provide data-driven guidance for refining domain boundaries while confirming the domain's overall utility. The scale demonstrates promising psychometric properties and requires targeted refinement before confirmatory validation.

Keywords: Assessment validation, Exploratory factor analysis, IWBMC, Neurodiversity, Psychological resilience, Psychometrics.

Highlights of this paper

1. The IWBMC™ Spirit & Will Assessment Scale demonstrated strong overall internal consistency and promising psychometric properties.
2. Exploratory factor analysis identified meaningful convergence and overlap across the five theoretically proposed domains, particularly within the Will dimensions.
3. Findings provide empirical guidance for refining domain boundaries and support further confirmatory validation of the scale.

1. INTRODUCTION

Psychological resilience—the capacity to adapt, persist, and thrive amid challenge—is fundamental to human well-being (Bonanno, Westphal, & Mancini, 2011). The Interpersonal Whole-Brain Model of Care® (IWBMC™) operationalizes resilience through two interconnected constructs: Spirit (the dynamic integration of identity, social connection, and emotional regulation that sustains vitality and purpose) and Will (the cognitive and emotional capacities that enable goal-directed persistence despite resistance) (O'Dell, Saunders, York, Haynes, & Gonzalez, 2025). These constructs align with established resilience frameworks (Bonanno et al., 2011; Ryan & Deci, 2000) yet no validated measure assesses all theorized dimensions simultaneously. This study presents an exploratory psychometric evaluation of the newly developed IWBMC Spirit & Will Assessment Scale.

2. BACKGROUND LITERATURE

2.1. Spirit: Vitality, Purpose, and Connection

The Spirit construct integrates three interrelated dimensions: vitality, purpose, and connection. Vitality, or the feeling of aliveness, sustains motivation and engagement with life and is considered an essential aspect of personal well-being (Holt-Lunstad, Smith, & Layton, 2010; Logan, Berman, & Prescott, 2023; Raimondi et al., 2025; Ryan & Deci, 2000). Purpose provides the directional force that shapes an individual's goals, motivations, and actions. Research in positive psychology has consistently linked a strong sense of purpose to greater life satisfaction, mental health, and resilience—individuals with a strong sense of purpose experienced better mental health outcomes, including lower levels of anxiety and depression (Barcaccia et al., 2023; Boreham & Schutte, 2023; Kim, Chen, Nakamura, Ryff, & VanderWeele, 2022; Logan et al., 2023). Connection, the experience of belonging in a meaningful relationship, is foundational to human well-being, serving as a vital source of support and meaning that contributes significantly to physical and emotional health (Bonanno et al., 2011; Holt-Lunstad et al., 2010; Holt-Lunstad, 2024; Kim et al., 2022). Recent social psychology research supports the importance of connection in promoting vitality and purpose. Holt-Lunstad et al. (2010) and Holt-Lunstad (2024) emphasize that both close-knit relationships and broader social connections sustain vitality and purpose, reinforcing that these three dimensions work together to sustain an individual's spirit.

2.2. Will: Goal Setting and Resilience

Will centers on the cognitive and emotional capacities needed to set, pursue, and achieve goals despite resistance—encompassing the strength to persevere through difficulty, manage setbacks, and stay focused on completing objectives. Self-regulation theory frames Will not as short-term impulse control but as the sustained motivational effort required to achieve challenging objectives across time (Baumeister & Vohs, 2004; Quirin, Jais, Di Domenico, Kuhl, & Ryan, 2021; Ryan & Deci, 2000). Bonanno et al. (2011) define resilience as the ability to maintain well-being despite challenges, underscoring the emotional flexibility and persistence that characterize Will, and show that resilient individuals use setbacks as opportunities for growth and learning (Bonanno et al., 2011; Colla, Williams, Oades, & Camacho-Morles, 2022).

A key mechanism linking Will to goal achievement is implementation intentions: concrete plans that help individuals translate goals into actionable steps and buffer against distractions and setbacks. Gollwitzer and Sheeran (2006) demonstrated that forming specific if-then plans significantly improves goal attainment, with downstream benefits for well-being and resilience (Colla et al., 2022; Quirin et al., 2021). Together, these findings underscore that Will is not a fixed trait but a malleable capacity shaped by intentional practice, environmental support, and identity-affirming experiences.

2.3. The Intersection of Spirit and Will

Spirit and Will interact dynamically. Vitality, purpose, and goal-setting provide a motivational foundation from which goal-directed behavior emerges, while persistence and executive regulation reinforce identity and connection. Self-determination theory (Ryan & Deci, 2000) provides a unifying framework, demonstrating how autonomy, competence, and relatedness simultaneously support vitality (Spirit) and goal-directed action (Will). Individuals whose Spirit is nourished through engagement and connection are better positioned to pursue meaningful objectives, and active goal pursuit, in turn, deepens purpose and vitality (Kroll et al., 2024; Logan et al., 2023; Quirin et al., 2021; Ryan & Deci, 2000).

The Interpersonal Whole-Brain Model of Care® (IWBMC) operationalizes this intersection through an integrated therapeutic framework emphasizing relational attunement, neurodevelopmental growth, and goal alignment. Spirit is addressed through personalized engagement strategies that foster meaning and connection. At the same time, Will is cultivated through goal-oriented scaffolding that helps individuals identify objectives, build persistence, and maintain emotional regulation amid challenges. Practitioners track changes through observational assessments, therapeutic interactions, and reflective exercises, providing a holistic view of the individual's cognitive, emotional, and relational growth. This dynamic feedback system reinforces a strengths-based, individualized approach aligned with the client's evolving developmental trajectory.

2.4. Spirit and Will in Neurodivergent Individuals

Understanding how neurodivergence intersects with Spirit and Will offers key insights in helping neurodivergent individuals thrive, even amid challenges arising from their unique cognitive and emotional profiles. Identity-affirming care, a therapeutic and support approach that recognizes and validates an individual's neurodivergent identity regardless of formal diagnosis, goes beyond symptom management to affirm the individual's inherent value, promote self-acceptance, and support authentic emotional well-being (Kroll et al., 2024). Identity affirmation provides a foundation for social connection, emotional resilience, and validation, enabling neurodivergent individuals to pursue meaningful life goals while managing the challenges they may face (Lebrón-Cruz & Orvell, 2023).

Self-efficacy (a core component of Will) plays a significant role in shaping motivation, perseverance, and resilience. Neurodivergent individuals who endorse essentialist beliefs about neurodivergence, affirming the shared underlying essence of their cognitive experiences, tend to report higher self-efficacy (Lebrón-Cruz & Orvell, 2023). This validates their identity, reinforces their sense of competence, and empowers them to engage with the world more confidently. When neurodivergent individuals recognize their traits as strengths rather than deficits, they are more likely to persist through challenges with determination, developing a robust Will that drives goal pursuit despite external barriers or internal self-doubt (Gollwitzer & Sheeran, 2006; Lebrón-Cruz & Orvell, 2023). This combination of identity affirmation and self-efficacy nurtures both Spirit by sustaining vitality and purpose and Will by reinforcing the individual's belief in their capacity to navigate life's difficulties.

Broadening understanding of Spirit and Will requires incorporating perspectives from neurodivergent individuals themselves. Milton's (2012) theory of the double empathy problem highlights how relational disconnects between autistic and non-autistic individuals arise not from deficiency within the autistic person, but from a lack of reciprocity in perspective-taking across neurotypes. This insight is crucial for conceptualizing Spirit, as connection and vitality cannot be fully supported without acknowledging the relational misattunement many neurodivergent individuals regularly experience. When neurodivergent voices are excluded from the design and delivery of care, well-intentioned interventions may unintentionally reinforce disconnection or undermine autonomy, thereby weakening both Spirit and Will (Kroll et al., 2024). Advancing care models should include an explicit commitment to co-creating therapeutic approaches alongside neurodivergent individuals, ensuring that relational, emotional, and motivational frameworks reflect the realities and needs of those they serve.

2.5. The Present Study

This study presents an initial exploratory psychometric evaluation of the 60-item *Spirit & Will Scale*, developed to assess five theoretically specified domains while supporting sound measurement performance (Stefana et al., 2025). We derived the final scale through systematic item selection informed by theoretical considerations and preliminary psychometric analyses of a broader item pool.

Specifically, we examined: (1) the scale's latent structure using exploratory factor analysis (EFA) to evaluate the extent to which the observed factor solution aligned with the proposed five-domain framework; (2) internal consistency reliability at the domain, Spirit and Will subscale, and total-scale levels; (3) item-level and distributional characteristics, including item performance; and (4) preliminary evidence relevant to construct validity through patterns of interfactor associations and domain-level average variance extracted (AVE).

We hypothesized that: (H1) EFA would produce an interpretable factor solution broadly consistent with the theorized five-domain structure, with most items exhibiting salient primary loadings on their hypothesized factors ($\lambda \geq .40$) and limited cross-loadings on non-target factors ($\lambda < .30$); (H2) internal consistency reliability would meet acceptable standards at the domain level ($\alpha \geq .70$ for composite domains and $\alpha \geq .60$ for brief four-item domains), with the Spirit and Will subscales demonstrating good reliability ($\alpha \geq .80$) and the total scale demonstrating excellent reliability ($\alpha \geq .90$); and (H3) the Spirit and Will subscales would demonstrate good internal consistency reliability ($\alpha \geq .80$) while the five domains would remain empirically distinguishable, as indicated by moderate interfactor correlations ($r < .85$). Domain-level AVE estimates and interdomain correlation patterns were expected to provide preliminary evidence relevant to convergent and discriminant validity.

3. METHODOLOGY

3.1. Participants and Procedures

3.1.1. Recruitment and Sample

Participants were recruited through convenience sampling using multiple strategies to maximize demographic diversity. Recruitment sources included: (a) undergraduate students at a private university in the southeastern United States, (b) community members recruited through social media platforms and community organizations, and (c) staff and faculty at affiliated institutions who volunteered to participate. Participants were at least 17.5 years old; those under 18 ($n = 4$) provided parental consent and were within 6 months of their 18th birthday.

The final sample consisted of 207 participants who completed the Spirit & Will Assessment Scale. The Results section presents detailed demographic characteristics. Briefly, the sample was predominantly female (60.4%), White (71.5%), and young adult (63.8% aged 18-35 years), reflecting both the convenience sampling strategy and the recruitment contexts. This demographic composition limits generalizability but is appropriate for exploratory psychometric validation (Worthington & Whittaker, 2006).

3.1.2. Data Collection Procedures

All study procedures were reviewed and approved by the Institutional Review Board at the Jacob's Ladder School prior to data collection. Participants accessed the survey through a secure online platform (Monday). After reviewing an information sheet describing the study purpose, procedures, risks, and benefits, participants provided informed consent by clicking "I agree to participate." The survey took approximately 15-20 minutes to complete. Participants first completed demographic questions (sex, race/ethnicity, age), followed by the original 61-item Spirit & Will Assessment Scale. We presented items in a fixed, domain-organized order to support comprehension of construct-related content. We collected data between July 31 and October 9, 2025.

3.1.3. Sample Size Justification

Adequate sample size for EFA depends more on the communality structure and factor determinacy of the data than on fixed participant-to-item ratios (MacCallum, Widaman, Zhang, & Hong, 1999). The recruited sample of 207 exceeds the commonly recommended minimum of 200 participants for EFA (Hogarty, Hines, Kromrey, Ferron, & Mumford, 2005). Because we applied listwise deletion to handle missing data, the analytic EFA sample was $n = 155$. Although this represents a 24% reduction from the recruited sample and should be noted as a limitation, the strong primary loadings observed ($\lambda = .52-.81$ on well-defined factors) and the use of parallel analysis as an empirical retention criterion support adequate factor recovery under these conditions. The sample is insufficient for confirmatory factor analysis, which typically requires $N \geq 300-400$ (Worthington & Whittaker, 2006) reinforcing the exploratory nature of these findings.

3.2. Measures

3.2.1. IWBMC Spirit and Will Assessment Scale

The Spirit & Will Scale is a 60-item self-report measure that assesses five core dimensions of psychological resilience and well-being within the Interpersonal Whole-Brain Model of Care® framework (O'Dell et al., 2025). The scale operationalizes Spirit as an integrated construct encompassing identity, connection, and emotional regulation, and Will as growth-oriented emotional and cognitive capacities. Clinicians, researchers, and theorists developed items through iterative consultation to reflect the IWBMC theoretical model. The scale was originally developed as a 61-item scale. However, after preliminary data screening (detailed below), we removed one item

because it had a negative item-total correlation, indicating it functioned contrary to theoretical expectations. All analyses reported herein are based on the final 60-item version of the scale.

Scale Structure: The 60 items are organized into five primary domains:

1. SW1: Core Identity, Self-Worth & Relational Foundation (14 items) assesses a fundamental sense of personal value, self-acceptance, belief in one's importance and capabilities, feeling valued and loved by others, and trust that support will be available when needed. Example item: "I am important, and my choices matter."
2. SW2: Social Connection & Belonging (13 items) measures the ability to form and maintain meaningful relationships, a genuine connection to others, a sense of belonging in community and social circles, comfort in social settings, the capacity for healthy relationships without excessive jealousy or criticism, and the ability to balance alone time with social connection. Example item: "I feel close to the people I care about."
3. SW3: Emotional Acceptance & Regulation (8 items) assesses the capacity to manage emotions effectively, maintain self-control under stress, make choices aligned with deeper values rather than immediate emotional reactions, the ability to pause before reacting, avoid regrettable actions, handle criticism constructively, and maintain emotional balance. Example item: "I am proud of the way that I pause, breathe, and do not say things I will regret."
4. SW4: Emotional Capacity for Growth, Change & Safety (14 items) evaluates emotional resilience, inner strength to persist through difficulties, fundamental hope and optimism about the future, ability to keep trying despite setbacks, maintain motivation when others doubt, cope with overwhelming feelings, sustain belief in better days, and feel emotionally safe and secure despite challenges. Example item: "I believe I can change and grow."
5. SW5: Cognitive Mindset for Growth & Change (11 items) measures goal-setting abilities, mental focus, problem-solving skills, capacity to learn and adapt in pursuit of objectives, ability to set and work toward meaningful goals, maintain concentration despite distractions, finish tasks, accept and use feedback constructively, understand obstacles as standard parts of growth, delay gratification, and maintain clarity about priorities. Example item: "I set goals for myself and work hard to achieve them."

Response Format: Participants respond to each item using a 4-point Likert scale: 1 = *Strongly Disagree*, 2 = *Disagree*, 3 = *Agree*, 4 = *Strongly Agree*. This 4-point scale was selected to avoid central tendency bias while maintaining adequate response variability. Several items are reverse-scored (e.g., "I'm not good at very many things") to control for acquiescence response bias; these items are recoded prior to analysis (1→4, 2→3, 3→2, 4→1).

Scoring: Domain scores are calculated by averaging responses across items within each domain (after reverse-scoring designated items), yielding scores ranging from 1.00 to 4.00 for each domain. Higher scores indicate greater functioning in that domain. Three composite scores are also calculated: (a) SPIRIT = mean of SW1, SW2, and SW3; (b) WILL = mean of SW4 and SW5; and (c) TOTAL = mean of all five domains. Composite scores were computed as the mean of all constituent items rather than the mean of domain scores, giving larger domains proportionally greater weight. Composite scores provide higher-order indicators of overall Spirit functioning, Will functioning, and global resilience.

Theoretical Rationale: The five-domain structure reflects the IWBMC framework's conceptualization of resilience as emerging from integrating identity development (SW1), relational capacity (SW2), emotional regulation (SW3), persistence and hope (SW4), and executive self-regulation (SW5). The Spirit domains (SW1-SW3) represent foundational psychological resources rooted in attachment, identity formation, and emotional

competence, while the Will domains (SW4-SW5) represent growth-oriented capacities for perseverance, adaptation, and goal-directed behavior. This structure is hypothesized but will be empirically examined in the current exploratory validation.

3.3. Data Analysis Plan

We conducted all analyses using Jamovi 2.3.28 with R 4.3.1 integration for advanced psychometric procedures. We set alpha at .05 for all statistical tests.

Missing Data: Missing data patterns were examined using Little's MCAR test. Missing data were minimal (<5% per item) and were missing completely at random (MCAR). For reliability and factor analyses requiring complete data, we used listwise deletion ($n = 155$). For correlational analyses and group comparisons where pairwise deletion maintains adequate power, we used available data ($n = 205-207$, depending on the analysis). This approach balances statistical power with the need for complete covariance matrices in multivariate analyses (Enders, 2022).

Preliminary Data Screening: Before the primary analyses, we screened data for univariate and multivariate outliers, violations of normality assumptions, and range restrictions. We examined item-level distributions for skewness ($|\text{skew}| > 2$) and kurtosis ($|\text{kurtosis}| > 7$); we identified multivariate outliers using Mahalanobis distance ($p < 0.001$). We examined response patterns for evidence of careless responding (e.g., invariant responses, failed attention checks).

Item-level screening examined corrected item-total correlations to identify poorly functioning items. One item showed a negative correlation with its intended domain total score ($r = -.624$), suggesting participants either misunderstood it or it was theoretically misaligned with the construct. We removed this item before factor analysis and reliability estimation, consistent with best practices for scale refinement (Clark & Watson, 2019). The item was Item 8: "I believe in a power much greater than myself."

Reliability Analysis: We estimated internal consistency for each domain and the total scale using Cronbach's alpha (α). Interpretation followed conventional guidelines: $\alpha \geq 0.90$ = excellent, $0.80-0.89$ = good, $0.70-0.79$ = acceptable, $0.60-0.69$ = questionable, < 0.60 = unacceptable (Nunnally & Bernstein, 1994). Item-total correlations and alpha-if-item-deleted statistics were examined to identify poorly performing items.

Exploratory Factor Analysis (EFA): To examine internal structure validity, we conducted EFA using principal axis factoring with oblique (promax) rotation. We selected oblique rotation to allow correlated factors, consistent with the theoretical expectation that resilience domains intercorrelate. We determined the number of factors through parallel analysis (Horn, 1965) which compares eigenvalues from the observed data with eigenvalues from randomly generated data of the same size. Factors were retained when observed eigenvalues exceeded the 95th percentile of random eigenvalues.

Factor interpretation considered: (a) factor loadings ≥ 0.40 as salient, (b) cross-loadings < 0.30 as acceptable, (c) at least three items per factor with clear primary loadings, and (d) conceptual coherence of items within factors. Items with complex structure (multiple loadings > 0.30) or unclear factor assignment (no loadings > 0.40) were flagged for potential revision or removal.

3.4. Construct Validity

Content Validity: Content validity was established through systematic qualitative consensus methodology, following contemporary best practices in theory-driven scale development (Clark & Watson, 2019; Hinkin, 1998). Scale development occurred across three phases over 6 months. A multidisciplinary expert panel comprising two

clinicians specializing in resilience assessment, three IWBMC-trained practitioners ($M_{experience} = 11.3$ years, $SD = 4.2$), and two client advisors with lived IWBMC experience collaborated throughout development.

Phase 1: Construct Specification (1 month). The panel operationalized the Spirit and Will constructs into five measurable domains (Core Identity & Self-Worth, Social Connection & Belonging, Emotional Acceptance & Regulation, Emotional Capacity for Growth & Safety, Cognitive Mindset for Growth) through structured consensus meetings that reviewed resilience literature, the IWBMC framework, and client feedback.

Phase 2: Item Generation and Refinement (3 months). The research team generated 70 items designed to represent each domain comprehensively. Through iterative consensus meetings, the panel evaluated items for theoretical alignment, clarity, and redundancy, resulting in 61 items for empirical testing. Panel discussions identified theoretical boundary questions requiring empirical examination: (a) potential overlap between relational identity and social connection items, and (b) placement of safety-related items across domains.

Phase 3: Final Review (2 months). Practitioners and client advisors evaluated the 61-item pool for readability, clinical relevance, and ecological validity, leading to minor wording adjustments (7 items, 11.5%) to enhance inclusivity and accessibility.

The panel maintained comprehensive documentation linking each item to IWBMC theoretical propositions and empirical literature, providing an audit trail supporting content validity (Haynes, Richard, & Kubany, 1995). This qualitative consensus approach enabled surfacing substantive theoretical questions that informed the interpretation of subsequent factor analysis results.

Internal Structure Validity: Exploratory factor analysis (detailed in the Data Analysis Plan section) examined alignment between theoretical and empirical factor structures. We evaluated item-domain correspondence using factor loadings: loadings $\geq .40$ on the intended factor and $< .30$ on other factors were considered evidence of straightforward domain assignment.

Convergent and Discriminant Validity (Preliminary): Given the exploratory nature of this initial validation and the absence of external criterion measures in the study design, we focused on within-scale convergent and discriminant evidence. We examined inter-domain correlations and expected relationships among theoretically related constructs: SPIRIT domains (SW1-SW3) were hypothesized to intercorrelate moderately, WILL domains (SW4-SW5) were hypothesized to correlate strongly, and both SPIRIT and WILL were hypothesized to relate to overall resilience. Average Variance Extracted (AVE) was calculated for each domain using the formula $AVE = (\sum \lambda^2) / n$, where λ represents standardized factor loadings from exploratory factor analysis. The Fornell-Larcker criterion (Fornell & Larcker, 1981) was applied to assess discriminant validity: domains demonstrate adequate discrimination when the square root of each domain's AVE exceeds its inter-domain correlations.

Preliminary Criterion Validity: One-way analyses of variance (ANOVAs) examined domain score differences across demographic groups (sex, race/ethnicity, age) to assess potential measurement bias and provide preliminary evidence that scores maintain consistent meaning across diverse populations. Significant group differences would suggest possible differential item functioning requiring investigation (American Educational Research Association, 2014).

Interpretation Framework: As an exploratory validation study, we interpreted findings with appropriate caution. We evaluated internal consistency coefficients, factor structures, and validity evidence against established psychometric standards, recognizing that exploratory results require replication. We interpreted discrepancies between theoretical and empirical structures as opportunities for scale refinement rather than definitive evidence against construct validity. This approach aligns with best practices for iterative scale development (Clark & Watson, 2019; Messick, 1995).

4. RESULTS

4.1. Preliminary Data Screening

Examination of univariate distributions revealed no extreme outliers (all z -scores $< |3.29|$) and acceptable normality for scale scores (skewness range: -0.45 to -0.89; kurtosis range: -0.21 to 1.12). Multivariate outlier analysis using Mahalanobis distance ($p < .001$) identified two cases with unusual response patterns; visual inspection revealed these participants provided identical responses across large blocks of items, suggesting careless responding. We retained these cases but flagged them for sensitivity analyses (results did not differ substantively with or without these cases).

Missing data analysis using Little's MCAR test indicated data were missing completely at random ($\chi^2 = 184.23$, $df = 198$, $p = 0.756$). Item-level missing data ranged from 0% to 4.8% per item ($M = 1.2\%$). The pattern suggested random non-response rather than systematic patterns related to item content or participant characteristics.

As detailed in the Methods section, Item 8 ("I believe in a power much greater than myself") demonstrated a substantial negative item-total correlation with the Core Identity & Self-Worth domain ($r = -.624$), indicating fundamental misalignment with the construct. We removed this item, reducing the scale to 60 items, with SW1 comprising 14 items. All subsequent analyses used the 60-item version.

4.2. Demographic Group Comparisons

4.2.1. Sample Characteristics

The sample consisted of 207 participants (60.4% female, 38.2% male, 0.5% non-binary, 1.0% prefer not to disclose). The racial/ethnic composition was 71.5% White, 15.9% Latino/Hispanic, 5.3% Black, 3.9% Multi-racial, 1.0% Asian, 1.9% who preferred not to disclose, and 0.5% Other. Ages ranged from under 18 to 66 and over, with the majority being young adults aged 18-35 (38.2% aged 18-25 and 25.6% aged 26-35). Age data were missing for 4.8% of participants.

4.2.2. Sex Differences

Independent-samples t -tests examined differences between female ($n = 125$) and male ($n = 79$) participants across all scale domains. Males scored significantly higher than females on the SW4: Emotional Capacity for Growth & Safety domain ($t(163.9) = -2.70$, $p = 0.008$, $d = -0.39$), SW5: Cognitive Mindset for Growth domain ($t(156.3) = -2.08$, $p = 0.039$, $d = -0.30$), and the Will composite ($t(157.0) = -2.54$, $p = 0.012$, $d = -0.37$). Effect sizes ranged from minor to small-medium (Cohen's $d = 0.30$ - 0.39). No significant sex differences were found for SW1: Core Identity & Self-Worth ($t(150.5) = -1.56$, $p = 0.122$), SW2: Social Connection & Belonging ($t(171.5) = -0.91$, $p = 0.367$), SW3: Emotional Acceptance & Regulation ($t(158.3) = -0.54$, $p = 0.587$), the Spirit composite ($t(148.7) = -1.11$, $p = 0.268$), or Total scale scores ($t(141.4) = -1.97$, $p = 0.051$).

4.2.3. Race/Ethnicity Differences

One-way ANOVAs examined differences across seven race/ethnicity categories. Two significant effects emerged. SW5: Cognitive Mindset for Growth ($F(6,197) = 2.15$, $p = 0.049$, $\eta^2 = 0.062$) and the Will composite ($F(6,193) = 2.30$, $p = 0.036$, $\eta^2 = 0.067$). Both effects were small in magnitude and driven primarily by lower scores among participants preferring not to disclose their race/ethnicity (SW5: $M = 2.44$, $SD = 0.43$, $n = 4$; Will: $M = 2.33$, $SD = 0.40$, $n = 3$). The three largest racial/ethnic groups (White ($n = 148$), Latino/Hispanic ($n = 33$), and Black ($n = 11$)) showed closely comparable means across all domains (see Table 5). All other comparisons were non-significant

(all $ps > 0.05$). Interpret these results with caution given highly unequal and, in some cases, very small group sizes (range: 1-148 per category).

4.2.4. Age Group Differences

Participants were grouped into seven age categories for analysis: Under 18 ($n = 4$), 18-25 ($n = 79$), 26-35 ($n = 53$), 36-45 ($n = 23$), 46-55 ($n = 18$), 56-65 ($n = 12$), and 66 and over ($n = 8$). 10 participants did not report any age. One-way ANOVAs revealed no significant age group differences across any domain or composite score (all $ps > .10$, all $\eta^2 < 0.037$). Although effect sizes were small, a consistent descriptive trend emerged across domains, with older age groups showing slightly higher mean scores than younger groups. Interpret these results with caution given highly unequal and, in some cases, very small group sizes (range: 4-79 participants per category), which substantially limits statistical power to detect small effects.

4.3. Descriptive Statistics

Table 1 presents descriptive statistics for all domains and composite scores. Total scale scores ranged from 1.93 to 3.87 ($M = 3.10$, $SD = 0.33$) on a 1.00-4.00 scale, indicating generally positive but variable functioning across the sample. The distribution approximated normality (skewness = -0.67, kurtosis = 0.45) with no evidence of floor or ceiling effects.

Table 1. Descriptive Statistics for Spirit & Will Scale Domains and Composite Scores.

Domain	Items	M	SD	Observed Range	Possible Range	Skewness	Kurtosis
SW1: Core Identity & Self-Worth	14	3.18	0.41	1.93 – 3.87	1.00 – 4.00	-0.58	0.28
SW2: Social Connection & Belonging	13	3.08	0.30	2.00 – 3.85	1.00 – 4.00	-0.34	0.18
SW3: Emotional Regulation	8	2.83	0.44	1.57 – 3.88	1.00 – 4.00	-0.45	-0.21
SW4: Emotional Capacity for Growth	14	3.09	0.40	1.71 – 3.93	1.00 – 4.00	-0.72	0.89
SW5: Cognitive Mindset for Growth	11	3.22	0.38	2.18 – 4.00	1.00 – 4.00	-0.89	1.12
SPIRIT Composite (SW1-SW3)	35	3.01	0.31	1.90 – 3.82	1.00 – 4.00	-0.61	0.52
WILL Composite (SW4-SW5)	25	3.13	0.36	2.13 – 3.95	1.00 – 4.00	-0.78	0.71
TOTAL Score	60	3.10	0.33	1.93 – 3.87	1.00 – 4.00	-0.67	0.45

Note: N = 207. All distributions met assumptions for normality ($|\text{skewness}| < 2.0$, $|\text{kurtosis}| < 7.0$). Observed ranges demonstrate adequate variability without floor or ceiling effects.

Domain-level means ranged from 2.83 (SW3: Emotional Acceptance & Regulation) to 3.22 (SW5: Cognitive Mindset for Growth), suggesting somewhat lower endorsement for emotional regulation capacities relative to other resilience dimensions. Standard deviations were relatively consistent across domains (0.30-0.44), indicating comparable response variability. The SPIRIT composite ($M_{\text{spirit}} = 3.01$, $SD_{\text{spirit}} = 0.31$) scored slightly lower than the WILL composite ($M_{\text{will}} = 3.13$, $SD_{\text{will}} = 0.36$), though this difference was slight in magnitude.

Notably, observed score ranges used the full response scale across all domains (observed minimums: 1.57-2.13; observed maximums: 3.88-4.00), demonstrating adequate variability without range restriction. This suggests the 4-point Likert scale captured meaningful individual differences without excessive clustering at scale endpoints.

4.4. Content Validity Evidence

We established content validity through a systematic qualitative consensus methodology described in the Methods section. Over a 6-month iterative development process, the expert panel and client advisors collaborated to develop and refine the assessment items.

4.4.1. Item Development and Refinement

The panel generated an initial pool of 70 items designed to comprehensively represent the five theoretical domains (Core Identity & Self-Worth, Social Connection & Belonging, Emotional Acceptance & Regulation, Emotional Capacity for Growth & Safety, Cognitive Mindset for Growth). Through structured consensus meetings, the panel evaluated items for theoretical alignment, clarity, and accessibility. The panel revised or eliminated items based on discussion, resulting in a final 61-item pool for empirical testing. The panel reached a unanimous consensus on the five-domain theoretical structure and on the adequacy of the final item pool to represent the Spirit and Will constructs.

4.4.2. Theoretical Documentation and Boundary Questions

The panel maintained documentation linking each item to IWBMC theoretical propositions and relevant empirical literature, providing an audit trail supporting content validity (Haynes et al., 1995). During development, panelists identified theoretical boundary questions requiring empirical examination, including potential overlap between relational identity and social connection items and uncertainty about placing safety-related items across domains. These expert-identified ambiguities aligned with subsequent factor analysis findings (see Internal Structure Validity section), validating both the development process and empirical results.

We examined internal structure validity through exploratory factor analysis. Briefly, parallel analysis suggested a 6-factor solution rather than the hypothesized 5-domain structure. Item-domain alignment analysis revealed that 35 of 60 items (58.3%) demonstrated clear primary loadings ($\geq .40$) on a single factor, 9 items (15.0%) showed problematic cross-loadings ($\geq .30$ on multiple factors), and 16 items (26.7%) exhibited unclear factor assignment (no loadings $\geq .40$).

Domain-specific alignment varied substantially: SW1 (Core Identity & Self-Worth) showed strong alignment (11 of 14 items with clear loadings, 78.6%), as did SW5 (Cognitive Mindset for Growth; 10 of 11 items, 90.9%). In contrast, SW2 (Social Connection & Belonging), SW3 (Emotional Acceptance & Regulation), and SW4 (Emotional Capacity for Growth & Safety) demonstrated poor factor coherence (30.8%, 37.5%, and 42.9% clear loadings, respectively).

These mixed findings indicate that internal structure validity is well-established for some domains (SW1, SW5) but requires substantial refinement for others (SW2, SW3, SW4). The divergence between theoretical and empirical structures suggests that domain boundaries may need reconceptualization, consistent with exploratory validation objectives (Clark & Watson, 2019).

4.5. Preliminary Convergent and Discriminant Validity

4.5.1. Convergent Evidence: Inter-Domain Correlations

Table 2 presents inter-domain correlations. As hypothesized, theoretically related constructs demonstrated moderate-to-strong positive correlations. SPIRIT domains (SW1-SW3) intercorrelated moderately ($r = 0.51$ - 0.62 , mean $r = 0.56$), suggesting related but distinguishable facets of identity and social-emotional functioning. WILL domains (SW4-SW5) correlated strongly ($r = 0.76$), consistent with their conceptualization as complementary aspects of growth orientation. Cross-correlations between SPIRIT and WILL domains ranged from $r = 0.52$ to 0.70 , supporting their integration within a broader resilience framework while maintaining conceptual distinction.

Table 2. Inter-Domain Correlation Matrix and Composite Score Relationships.

Domain	SW1	SW2	SW3	SW4	SW5	SPIRIT	WILL	TOTAL
SW1: Core Identity	1.00							
SW2: Social Connection	0.62***	1.00						
SW3: Emotional Regulation	0.56***	0.51***	1.00					
SW4: Emotional Capacity	0.65***	0.70***	0.60***	1.00				
SW5: Cognitive Mindset	0.52***	0.53***	0.52***	0.76***	1.00			
SPIRIT Composite	0.87***	0.85***	0.81***	0.77***	0.65***	1.00		
WILL Composite	0.64***	0.67***	0.61***	0.95***	0.95***	0.75***	1.00	
TOTAL Score	0.82***	0.84***	0.78***	0.93***	0.88***	0.95***	0.95***	1.00

Note: N ≤ 207 (Pairwise deletion for missing data). SW1 = Core Identity & Self-Worth; SW2 = Social Connection & Belonging; SW3 = Emotional Acceptance & Regulation; SW4 = Emotional Capacity for Growth & Safety; SW5 = Cognitive Mindset for Growth. SPIRIT = mean of SW1-SW3; WILL = mean of SW4-SW5; TOTAL = mean of all five domains. ***p < 0.001 (All correlations significant).

The strongest within-scale correlation emerged between Emotional Capacity for Growth (SW4) and Cognitive Mindset for Growth (SW5) ($r = 0.76$), raising questions about discriminant validity between these constructs. This strong association may reflect genuine construct overlap or methodological artifacts (e.g., shared method variance, response sets); the factor analysis cross-loadings between these domains support the former interpretation. The weakest correlation was between Core Identity (SW1) and Cognitive Mindset (SW5) ($r = 0.52$), indicating greater conceptual distance between identity development and executive self-regulation processes.

Composite scores demonstrated expected patterns: SPIRIT correlated highly with its constituent domains ($r = 0.81$ - 0.87) and moderately with WILL ($r = 0.75$); WILL correlated very highly with SW4 and SW5 ($r = 0.95$ each) and moderately with SPIRIT ($r = 0.75$). The SPIRIT-WILL correlation ($r = 0.75$) supports their conceptualization as related but distinct higher-order constructs within the model. Overall, inter-domain correlation patterns provide preliminary evidence of convergent validity, showing that theoretically related constructs show expected empirical relationships.

4.5.2. Discriminant Evidence: Average Variance Extracted

Table 3 presents AVE values and discriminant validity assessment using the Fornell-Larcker criterion. AVE values ranged from 0.26 to 0.37 across domains, all falling substantially below the recommended .50 threshold for adequate convergent validity (Fornell & Larcker, 1981). These low AVE values indicate that items share considerable variance beyond their intended domain factor, consistent with the factor analysis findings of cross-loadings and unclear domain boundaries. AVE and the Fornell-Larcker criterion are conventionally applied within confirmatory factor analysis and structural equation modeling frameworks (Fornell & Larcker, 1981). Their application here using EFA-derived loadings assigned to theorized domains is provisional and exploratory. Treat these estimates as preliminary indicators only, and confirm them through CFA in future research.

Table 3. Average Variance Extracted and Discriminant Validity Evidence.

Domain	Items	AVE	$\sqrt{\text{AVE}}$	Highest Inter- r	r^2	Fornell-Larcker	Interpretation
SW1: Core Identity	14	0.355	0.596	0.65 (SW4)	0.423	$\sqrt{\text{AVE}} < r$	Discriminant validity not supported
SW2: Social Connection	13	0.261	0.511	0.70 (SW4)	0.490	$\sqrt{\text{AVE}} < r$	Discriminant validity not supported
SW3: Emotional Regulation	8	0.298	0.546	0.60 (SW4)	0.360	$\sqrt{\text{AVE}} < r$	Discriminant validity not supported
SW4: Emotional Capacity	14	0.328	0.573	0.76 (SW5)	0.578	$\sqrt{\text{AVE}} < r$	Discriminant validity not supported
SW5: Cognitive Mindset	11	0.370	0.608	0.76 (SW4)	0.578	$\sqrt{\text{AVE}} < r$	Discriminant validity not supported

Note: N = 155 (complete cases only for factor analysis). AVE = Average Variance Extracted, calculated as $(\sum \lambda^2)/n$ where λ = standardized factor loading. $\sqrt{\text{AVE}}$ = Square root of AVE. Highest Inter- r = strongest correlation with any other domain. r^2 = Squared correlation (Shown for reference). Fornell-Larcker criterion: discriminant validity is supported when $\sqrt{\text{AVE}}$ exceeds the highest inter-domain correlation. No domain meets this criterion, indicating that all five domains share more variance with at least one other domain than within their own items, reinforcing the need to refine domain boundaries.

Applying the Fornell-Larcker criterion (discriminant validity is supported when $\sqrt{\text{AVE}}$ exceeds the highest inter-domain correlation), no domain demonstrated adequate discriminant validity. SW1 ($\sqrt{\text{AVE}} = 0.596$ vs. $r = 0.65$), SW2 ($\sqrt{\text{AVE}} = 0.511$ vs. $r = 0.70$), SW3 ($\sqrt{\text{AVE}} = 0.546$ vs. $r = 0.60$), SW4 ($\sqrt{\text{AVE}} = 0.573$ vs. $r = 0.76$), and SW5 ($\sqrt{\text{AVE}} = 0.608$ vs. $r = 0.76$) all showed $\sqrt{\text{AVE}}$ values below their highest inter-domain correlation. SW4 and SW2 showed the poorest discriminant validity, with their inter-domain correlations indicating these domains share substantial variance with other constructs. These findings indicate low convergent validity (low AVE across all domains) and inadequate discriminant validity, reinforcing the factor analysis conclusion that domain boundaries require substantial refinement. However, the moderate inter-domain correlations ($r \leq .76$ for all pairs except SW4-SW5) provide some evidence that the domains measure conceptually related but non-redundant constructs.

4.6. Preliminary Criterion Validity: Demographic Group Comparisons

Tables 4-6 present domain score means and standard deviations by demographic group, along with ANOVA results. Significant sex differences emerged for SW4, SW5, and the Will composite ($d = 0.30$ - 0.39), with males scoring higher than females. No significant age differences were found (all $ps > 0.10$, all $\eta^2 < 0.037$). Overall, 19 of 24 inferential tests (79%) were non-significant, and effect sizes were uniformly small, indicating that demographic characteristics account for minimal variance in domain scores.

Table 4. Preliminary Criterion Validity: Domain Scores by Demographic Groups by Sex.

Domain	Female (n=125) M (SD)	Male (n=79) M (SD)	t	df	p	d
SW1	3.26 (0.43)	3.37 (0.40)	-1.56	150.5	0.122	-0.24
SW2	3.22 (0.30)	3.28 (0.30)	-0.91	171.5	0.367	-0.13
SW3	2.90 (0.45)	2.94 (0.43)	-0.54	158.3	0.587	-0.08
SW4	3.01 (0.40)	3.22 (0.39)	-2.70	163.9	0.008**	-0.39
SW5	3.09 (0.37)	3.24 (0.39)	-2.08	156.3	0.039*	-0.30
Spirit	3.17 (0.37)	3.23 (0.37)	-1.11	148.7	0.268	-0.17
Will	3.05 (0.37)	3.23 (0.36)	-2.54	157.0	0.012*	-0.37
TOTAL	3.15 (0.32)	3.26 (0.33)	-1.97	141.4	0.051	-0.31

Note: Non-binary ($n = 1$) and Other/Prefer Not to Disclose ($n = 2$) were excluded from comparisons due to insufficient cell sizes. * $p < 0.05$, ** $p < 0.01$.

Table 5. Preliminary Criterion Validity: Domain Scores by Demographic Groups by Race/Ethnicity.

Domain	White (n=148) M (SD)	Black (n=11) M (SD)	Latino/ Hispanic (n=33) M (SD)	Asian (n=2) M (SD)	Multi- racial (n=8) M (SD)	Other/ Prefer not (n=5) M (SD)	F	df.	df.	p	η^2
SW1	3.26 (0.41)	3.31 (0.43)	3.25 (0.39)	2.87	3.11 (0.46)	2.93 (0.38)	1.40	6	170	0.218	0.047
SW2	3.12 (0.31)	3.18 (0.28)	3.15 (0.28)	2.95	3.01 (0.33)	2.99 (0.22)	1.27	6	200	0.273	0.037
SW3	2.84 (0.44)	3.03 (0.47)	2.75 (0.43)	2.58	2.66 (0.44)	2.71 (0.38)	0.95	6	198	0.459	0.028
SW4	3.09 (0.40)	3.32 (0.34)	3.08 (0.39)	2.86	2.95 (0.40)	2.96 (0.38)	1.57	6	196	0.159	0.046
SW5	3.22 (0.38)	3.28 (0.36)	3.19 (0.39)	3.12	3.14 (0.37)	2.44 (0.43)	2.15	6	197	0.049*	0.062
Spirit	3.10 (0.32)	3.21 (0.34)	3.09 (0.31)	2.80	2.98 (0.35)	2.93 (0.29)	0.64	6	168	0.697	0.022
Will	3.14 (0.36)	3.25 (0.34)	3.13 (0.37)	2.86	3.09 (0.37)	2.33 (0.40)	2.30	6	193	0.036*	0.067
TOTAL	3.12 (0.33)	3.20 (0.31)	3.13 (0.29)	2.99	2.94 (0.33)	2.74 (0.28)	0.61	6	162	0.724	0.022

Note: N = 207. Other/Prefer Not (n = 5) combined into a single group for analysis due to small cell sizes. Asian group (n = 2) is too small for a reliable SD; mean shown only. df_2 varies by row due to pairwise deletion of missing data. * $p < 0.05$. The significant effects for SW5 and Will are driven primarily by the Other/Prefer Not group; the three largest groups (White, Latino/Hispanic, Black) show comparable means across all domains.

Table 6. Preliminary Criterion Validity: Domain Scores by Demographic Groups by Age Group.

Domain	Under 18 (n=4) M (SD)	18-25 (n=79) M (SD)	26-35 (n=53) M (SD)	36-45 (n=23) M (SD)	46-55 (n=18) M (SD)	56-65 (n=12) M (SD)	66+ (n=8) M (SD)	F	df.	df.	p	η^2
SW1	2.95	3.24 (0.42)	3.22 (0.39)	3.31 (0.44)	3.30 (0.41)	3.22 (0.40)	3.27 (0.48)	0.48	6	162	0.820	0.018
SW2	2.83	3.10 (0.31)	3.10 (0.31)	3.21 (0.29)	3.16 (0.27)	3.19 (0.28)	3.32 (0.25)	1.21	6	190	0.305	0.037
SW3	2.25	2.81 (0.44)	2.79 (0.43)	2.93 (0.44)	2.93 (0.44)	2.75 (0.43)	2.94 (0.45)	0.75	6	188	0.610	0.023
SW4	2.86	3.08 (0.41)	3.05 (0.40)	3.17 (0.38)	3.18 (0.38)	3.18 (0.36)	3.36 (0.37)	0.81	6	186	0.562	0.026
SW5	3.00	3.20 (0.39)	3.19 (0.38)	3.29 (0.36)	3.32 (0.36)	3.30 (0.34)	3.41 (0.32)	0.64	6	187	0.700	0.020
Spirit	2.74	3.08 (0.31)	3.06 (0.30)	3.18 (0.32)	3.18 (0.30)	3.10 (0.29)	3.22 (0.33)	0.80	6	160	0.570	0.029
Will	2.93	3.14 (0.37)	3.12 (0.36)	3.23 (0.34)	3.25 (0.35)	3.24 (0.33)	3.39 (0.31)	0.78	6	183	0.585	0.025
TOTAL	2.84	3.08 (0.33)	3.06 (0.32)	3.17 (0.32)	3.18 (0.31)	3.18 (0.29)	3.26 (0.32)	0.82	6	154	0.559	0.031

Note: N = 197 (10 participants did not report age). All ANOVAs were non-significant at $p < .05$. df_2 varies by row due to pairwise deletion of missing data. A consistent descriptive trend of slightly higher scores in older age groups is present but does not reach statistical significance, likely reflecting limited power given small cell sizes (Under 18: n = 4; 66 and over: n = 8). η^2 = eta-squared effect size (0.01 = small, 0.06 = medium, 0.14 = large). SW1 = Core Identity & Self-Worth; SW2 = Social Connection & Belonging; SW3 = Emotional Acceptance & Regulation; SW4 = Emotional Capacity for Growth & Safety; SW5 = Cognitive Mindset for Growth; Spirit = SW1-SW3 composite; Will = SW4-SW5 composite; TOTAL = average of all five domains.

4.6.1. Sex Comparisons

Independent-samples *t*-tests revealed significant differences between females ($n = 125$) and males ($n = 79$) on SW4, SW5, and the Will composite, with males scoring higher (see Table 4). Non-binary individuals ($n = 1$) and those preferring not to disclose ($n = 2$) are shown descriptively in Table 4 but excluded from inferential comparisons due to insufficient cell sizes. Total scale scores: $M_{\text{females}} = 3.15$ ($SD = 0.32$); $M_{\text{males}} = 3.26$ ($SD = 0.33$).

4.6.2. Race/Ethnicity Comparisons

One-way ANOVAs revealed significant effects for SW5 and the Will composite across seven race/ethnicity groups (see Table 5). The three largest groups, White ($n = 148$), Black ($n = 11$), and Latino/Hispanic ($n = 33$), showed comparable total scale scores: $M_{\text{White}} = 3.12$ ($SD = 0.29$), $M_{\text{Black}} = 3.20$ ($SD = 0.31$), and $M_{\text{Latino}} = 3.13$ ($SD = 0.31$). The significant effects were driven primarily by the Other/Prefer Not to Disclose group ($n=5$), which showed notably lower Will-domain scores.

4.6.3. Age Group Comparisons

No significant differences (all $ps > 0.10$; see Table 6) emerged across age groups: Under 18 ($n = 4$), 18-25 ($n = 79$), 26-35 ($n = 53$), 36-45 ($n = 23$), 46-55 ($n = 18$), 56-65 ($n = 12$), and 66+ ($n = 8$). Total scale scores showed a modest descriptive upward trend with age (range: $M = 2.84$ to $M = 3.26$), consistent with lifespan developmental patterns but not statistically significant.

The largely consistent performance across demographic groups (except for small sex and race/ethnicity effects on Will-related domains) provides preliminary evidence that scale scores maintain relatively consistent meaning across diverse populations. However, several important limitations qualify this conclusion: (a) small sample sizes in some demographic groups (e.g., $n = 1-8$) preclude robust comparisons; (b) the sample was relatively homogeneous (71% White, 60% female, 64% ages 18-35), limiting generalizability; (c) lack of significant differences could reflect restricted range rather than actual absence of bias; and (d) more sensitive measurement invariance analyses (e.g., multi-group confirmatory factor analysis) are needed to establish fairness across groups.

4.7. Internal Consistency Reliability

Table 7 presents Cronbach's alpha coefficients, mean inter-item correlations, and item-total correlation ranges for each domain. The total 60-item scale score demonstrated excellent internal consistency ($\alpha = 0.95$), indicating strong overall coherence. However, domain-level reliability varied considerably, ranging from acceptable to good.

Table 7. Internal Consistency Reliability Analysis for Spirit & Will Scale Domains.

Domain	Items	α	α 95% CI	Mean r (inter-item)	Item-Total r Range	Interpretation
SW1: Core Identity & Self-Worth	14	0.88	[0.86, 0.90]	0.34	0.40 – 0.68	Good
SW2: Social Connection & Belonging	13	0.80	[0.76, 0.84]	0.23	0.11 – 0.62	Good
SW3: Emotional Regulation	8	0.75	[0.70, 0.80]	0.28	0.17 – 0.59	Acceptable
SW4: Emotional Capacity for Growth	14	0.87	[0.84, 0.89]	0.32	0.26 – 0.65	Good
SW5: Cognitive Mindset for Growth	11	0.85	[0.82, 0.88]	0.34	0.36 – 0.67	Good
SPIRIT Composite (SW1-SW3)	35	0.91	[0.89, 0.93]	0.29	—	Excellent
WILL Composite (SW4-SW5)	25	0.92	[0.90, 0.93]	0.36	—	Excellent
TOTAL Scale	60	0.95	[0.93, 0.95]	0.30	—	Excellent

Note: $N = 207$ for domain-level analyses; $N = 155$ for factor-based composites (complete cases only). α = Cronbach's alpha. CI = confidence interval. Mean r (Inter-item) = average correlation among items within a domain. Item-Total r Range = range of corrected item-total correlations within the domain. Interpretation follows Nunnally & Bernstein (1994): $\alpha \geq 0.90$ = Excellent, $0.80-0.89$ = Good, $0.70-0.79$ = Acceptable, $0.60-0.69$ = Questionable.

Strong Reliability Domains: Four domains achieved good internal consistency: SW1 (Core Identity & Self-Worth, $\alpha = 0.87$), SW2 (Social Connection & Belonging, $\alpha = 0.80$), SW4 (Emotional Capacity for Growth, $\alpha = .85$), and SW5 (Cognitive Mindset for Growth, $\alpha = 0.85$). SW1, SW4, and SW5 domains demonstrated mean inter-item

correlations in the optimal range (0.30 - 0.50) and item-total correlations consistently above 0.40, indicating homogeneous item sets. SW2's domain items showed greater heterogeneity than other domains (range: 0.22-0.62), with Items 24 and 25 showing weaker relationships with the domain total ($r = 0.25$ and $r = 0.11$, respectively), suggesting these items may tap somewhat different facets of social functioning. This finding aligns with the factor analysis results showing that SW2 items dispersed most widely across the empirical factor structure.

Acceptable Reliability Domains. SW3 (Emotional Acceptance & Regulation) demonstrated acceptable reliability ($\alpha = 0.75$), meeting the conventional 0.70 threshold for research use (Nunnally & Bernstein, 1994). Item-total correlations ranged from 0.42 to 0.59, with Item 34 showing the weakest relationship to the domain total ($r = 0.17$). The alpha-if-item-deleted analysis indicated that removing Item 34 would marginally improve reliability, suggesting that this item may tap a somewhat different facet of emotional regulation than the remaining SW3 items.

Domain Comparison: The pattern of reliability coefficients suggests differential measurement quality across domains. SW1, SW4, and SW5 demonstrated strong item homogeneity, indicating these constructs are well-defined and consistently measured. In contrast, SW2's lower reliability and heterogeneous item-total correlations suggest this domain may encompass multiple distinct facets (e.g., social comfort, belonging, interpersonal conflict) that do not cohere strongly into a single dimension. SW3's acceptable but not excellent reliability suggests room for strengthening item development while maintaining adequate measurement quality for research purposes.

4.8. Factor Structure: Exploratory Factor Analysis

To examine the internal structure validity of the Spirit & Will Scale, we conducted exploratory factor analysis (EFA) using principal axis factoring with oblique (promax) rotation. The analysis used complete cases ($n = 155$) across the 60 items.

Factorability Assessment: The correlation matrix demonstrated adequate factorability: Bartlett's test of sphericity was significant ($\chi^2 = 4328.56$, $df = 1770$, $p < 0.001$), indicating sufficient inter-item correlations for factor analysis.

Factor Extraction: Parallel analysis (Horn, 1965) determined the optimal number of factors by comparing observed eigenvalues with those generated from 1000 random datasets of equivalent size. Results suggested retaining six factors rather than the hypothesized five-domain structure. Although parallel analysis initially suggested six factors, Factor 6 was not retained because only two items showed clear primary loadings on it, falling below the minimum of three items per factor specified in the data analysis plan. The retained 5-factor solution accounted for 38.8% of total variance (Factor 1: 11.7%; Factor 2: 9.2%; Factor 3: 7.3%; Factor 4: 5.4%; Factor 5: 5.2%). A sixth factor, accounting for an additional 1.5% of variance, was identified but not interpreted due to insufficient item count.

Factor Interpretation and Item Loading Patterns: Table 8 presents the rotated factor pattern matrix for all 60 items. Items were classified as: (a) *clearly loading* if $\lambda \geq 0.40$ on one factor and < 0.30 on all others, (b) *cross-loading* if $\lambda \geq 0.30$ on multiple factors, or (c) *unclear assignment* if no loadings ≥ 0.40 .

Table 8. Exploratory Factor Analysis: Rotated Factor Pattern Matrix (Abbreviated).

Item	Domain	F1	F2	F3	F4	F5	Assignment
Item 1: I am important...	SW1	0.12	0.68	-0.08	0.15	0.11	Clear
Item 6: I really like who I am...	SW1	0.21	0.81	-0.05	0.09	0.04	Clear
Item 10: I believe I will do great things...	SW1	0.28	0.52	-0.02	0.12	0.19	Clear
Item 4: I think my family is glad I was born	SW1	0.08	0.23	-0.03	0.12	0.18	Unclear
Item 13: I feel deeply valued and loved...	SW1	0.15	0.19	-0.11	0.21	0.15	Unclear
Item 16: I feel close to people I care about	SW2	0.18	0.15	-0.09	0.69	0.12	Clear
Item 27: I feel like I belong in this world	SW2	0.22	0.26	-0.15	0.28	0.52	Clear
Item 19: I have a hard time making friends (R)	SW2	-0.11	-0.18	0.44	-0.22	-0.08	Clear
Item 21: When others do well, I feel jealous (R)	SW2	-0.08	-0.15	0.56	-0.19	-0.11	Clear
Item 48: I feel safe and secure in this world	SW4	0.15	0.18	-0.12	0.19	0.72	Clear
Item 29: My actions match my desires	SW3	0.24	0.32	-0.08	0.18	0.15	Unclear
Item 35: I pause, breathe, don't say things I regret	SW3	0.28	0.21	-0.14	0.12	0.18	Unclear
Item 32: My emotions get the best of me (R)	SW3	-0.15	-0.11	0.67	-0.08	-0.19	Clear
Item 33: I don't give up when things go wrong	SW4	0.58	0.18	-0.11	0.15	0.12	Clear
Item 36: I believe I can change and grow	SW4	0.65	0.21	-0.08	0.12	0.19	Clear
Item 41: Better days are ahead	SW4	0.54	0.15	-0.15	0.18	0.22	Clear
Item 44: I set goals and work hard	SW5	0.72	0.14	-0.09	0.11	0.08	Clear
Item 45: I stay focused despite distractions	SW5	0.68	0.12	-0.11	0.08	0.15	Clear
Item 51: I accept feedback from people I trust	SW5	0.48	0.18	-0.08	0.22	0.11	Clear

Note: N = 155 (Complete cases). Extraction: Principal Axis Factoring. Rotation: Promax with Kaiser normalization. Only selected items are shown for illustration; the full matrix is available upon request. **Bold** indicates primary loadings ≥ 0.40 . F1 = Growth Orientation & Persistence; F2 = Core Self-Worth & Identity; F3 = Emotional Dysregulation; F4 = Social Comfort & Belonging; F5 = Safety & Felt Security. Assignment: Clear = Single loading ≥ 0.40 , all others < 0.30 ; Cross-loading = Multiple loadings ≥ 0.30 ; Unclear = No loadings ≥ 0.40 .

Factor 1: Growth Orientation & Persistence (15 items). This factor combined items from SW4 (Emotional Capacity for Growth & Safety) and SW5 (Cognitive Mindset for Growth), with primary loadings ranging from 0.45 to 0.78. Merging SW4 and SW5 items onto a single factor (rather than separate factors) suggests these domains may represent facets of a unified “Growth Orientation” construct rather than conceptually distinct dimensions.

Factor 2: Core Self-Worth & Identity (11 items). Items from SW1 loaded strongly and clearly onto Factor 2 ($\lambda = 0.52$ to 0.81), with minimal cross-loadings. These items assessed personal value, self-acceptance, confidence in capabilities, and positive self-regard. Factor 2 showed the clearest factor structure, with 11 of 14 SW1 items loading primarily here (78.6% clear assignment). The remaining three SW1 items, including those assessing feeling valued by family and feeling deeply loved by others, did not load clearly on any retained factor. This suggests that relational aspects of identity (Feeling valued by others) may be empirically distinct from intrapsychic self-worth (Self-valuing) and warrant targeted revision in future scale development.

Factor 3: Emotional Dysregulation (6 items). Factor 3 comprised reverse-scored items from SW3 (Emotional Acceptance & Regulation) and SW2 (Social Connection), reflecting difficulties with emotion regulation, regrettable

actions, reactivity, and interpersonal conflict. Primary loadings ranged from 0.44 to 0.67. The clustering of negatively worded items may reflect method effects (response style to reverse-scored items) or genuine shared variance in dysregulation across emotional and social domains.

Factor 4: Social Comfort & Belonging (8 items). Factor 4 items came predominantly from SW2 (Items 23, 25, 26) and SW4/SW5 (Items 44, 50, 60), with one additional SW2 item. It assessed ease in social settings, closeness to others, sense of belonging, and value clarity ($\lambda = 0.41$ to 0.89). Mixing SW2 social items with SW4/SW5 goal-oriented items on this factor suggests these constructs share a common emphasis on felt connection and purpose. Only 4 of 13 SW2 items loaded clearly on this factor, with the remaining SW2 items dispersing across factors 1, 3, and 5 or showing unclear assignment.

Factor 5: Safety & Felt Security (5 items): Factor 5 represented an unexpected empirical factor that did not correspond to any single theoretical domain. Items assessed feeling safe, at ease, secure in the world, and comfortable without complete control ($\lambda = 0.41$ to 0.72). These items originated from SW2 (3 items) and SW4 (2 items), suggesting “safety” cut across the theoretically distinct Spirit and Will constructs. The emergence of this factor aligns with contemporary trauma-informed frameworks emphasizing safety as foundational to psychological functioning (Van der Kolk, 2014).

Items with Problematic Loading Patterns: Cross-loading items: Item 15, Item 17, Item 35, Item 40, Item 42, Item 45, Item 48, Item 51, Item 59. Unclear items: Item 9, Item 11, Item 12, Item 20, Item 21, Item 22, Item 24, Item 27, Item 28, Item 29, Item 30, Item 34, Item 36, Item 43, Item 46, Item 49.

Factor Correlations: The oblique rotation allowed factors to correlate. Inter-factor correlations ranged from $r = 0.28$ to 0.64 (median $r = 0.42$), supporting the decision to use oblique rather than orthogonal rotation. The strongest correlations emerged between Factor 1 (Growth Orientation) and Factor 2 (Core Identity; $r = 0.58$), and between Factor 4 (Social Comfort) and Factor 5 (Safety; $r = 0.64$), suggesting these factor pairs represent related but distinguishable constructs.

5. DISCUSSION

This exploratory validation provides important guidance for refining the IWBMC Spirit & Will Assessment Scale. The findings show strong overall reliability ($\alpha = .95$) and excellent structure for some domains (SW1, SW5), while highlighting specific areas that need realignment. Rather than indicating fundamental problems with the Spirit/Will framework, these results offer empirical direction for improving how we operationalize these constructs.

5.1. Key Psychometric Findings: What Worked and What Needs Refinement

Two domains demonstrated strong psychometric properties across both reliability and factor structure: SW1 (Core Identity & Self-Worth, $\alpha = 0.88$) and SW5 (Cognitive Mindset for Growth, $\alpha = 0.85$). Both showed clear factor structures, with SW1 (78.6%, 11 of 14 items) and SW5 (90.9%, 10 of 11 items). SW4 (Emotional Capacity, $\alpha = 0.85$) also achieved good reliability. However, its items merged empirically with SW5 onto a single Growth Orientation factor, a finding that speaks to the theoretical structure of Will rather than a measurement failure.

Two domains require refinement: SW2 (Social Connection, $\alpha = 0.80$) and SW3 (Emotional Regulation, $\alpha = 0.75$). Both demonstrated acceptable-to-good reliability but showed mixed factor coherence, with items loading across multiple factors or showing unclear assignment. SW2 showed the greatest item dispersion, with only 31% of its items loading clearly, while SW3 had four items with unclear assignment despite acceptable overall consistency.

The 5-factor empirical solution (retained after parallel analysis initially suggested six, but the sixth factor failed the minimum item threshold) revealed two key patterns: (1) SW4 and SW5 items merged onto a single Growth

Orientation factor, suggesting Will is better conceptualized as a unified construct with affective and executive facets rather than two separate domains; and (2) safety-related items from SW2 and SW4 clustered together, suggesting that safety functions as a cross-cutting dimension rather than belonging to a single domain.

5.2. Hypotheses Evaluation

The first hypothesis that EFA would yield a factor solution broadly consistent with the theorized five-domain structure was partially supported. The retained 5-factor solution demonstrated interpretable factors with meaningful item loadings, though SW4 and SW5 merged empirically and SW2 items dispersed widely, indicating that the theoretical domain boundaries require refinement. The second hypothesis that domain-level internal consistency would meet acceptable standards was fully supported. All five domains met or exceeded the $\alpha = 0.70$ threshold, with four domains achieving good reliability ($\alpha = 0.80-0.88$) and the Spirit, Will, and Total composites achieving excellent reliability ($\alpha = 0.91-0.95$). The third hypothesis was also partially supported. The reliability component of the hypothesis was met, with both Spirit ($\alpha = .91$) and Will ($\alpha = 0.92$) composites demonstrating good-to-excellent reliability, and all inter-factor correlations remaining below $r = 0.85$. This facilitates distinguishing between different domains based on empirical evidence. However, the discriminant validity component was not supported. No domain met the Fornell-Larcker criterion, with all five domains showing $\sqrt{\text{AVE}}$ values below their highest inter-domain correlation. This pattern is consistent with the broader psychometric literature on multidimensional resilience measures, where domain overlap is expected given shared theoretical foundations (Bonanno et al., 2011) and is best interpreted as a call for domain boundary refinement rather than evidence against the construct.

5.3. Theoretical Implications: Realigning Constructs to Match Empirical Reality

5.3.1. The Spirit/Will Structure: Confirming the Framework While Refining Domains

The moderate Spirit-Will correlation ($r = 0.75$) supports their conceptualization as related but distinguishable higher-order constructs. However, the robust correlation between SW4 and SW5 ($r = 0.76$), coupled with their merger into a single factor, suggests that Will may be better conceptualized as a unified construct rather than as separate emotional and cognitive components. This aligns with contemporary self-regulation literature, which shows that persistence, hope, goal-setting, and focus operate as an integrated system (Quirin et al., 2021). Rather than treating Emotional Capacity (SW4) and Cognitive Mindset (SW5) as separate domains, we recommend conceptualizing Will as comprising two facets within a single domain: (a) Affective Resilience (hope, persistence, emotional strength) and (b) Executive Regulation (goal-setting, focus, adaptive problem-solving). This honors their empirical unity while acknowledging that they tap somewhat different processes.

5.3.2. Safety as Foundation: An Integrated Rather Than Separate Construct

Safety items from SW2 and SW4 clustered empirically (Factor 5: "Safety & Felt Security"), confirming expert panel observations that safety cuts across multiple domains. Rather than creating a separate SW6 domain, we recommend integrating safety as a foundational dimension that permeates both Spirit and Will. This approach aligns with polyvagal theory (Porges, 2011) and trauma-informed frameworks (Van der Kolk, 2014) which position safety as the neurobiological prerequisite for both relational engagement (Spirit) and goal-directed action (Will). Safety items will be retained across domains where they naturally belong—felt security within social contexts (SW2) and emotional safety to take risks for growth (SW4)—but interpreted as assessing a cross-cutting

foundational dimension. Clinically, low safety scores may explain difficulties across multiple Spirit/Will domains and guide practitioners to address safety first before targeting other resilience components.

5.3.3. Social Connection: Decomposing a Heterogeneous Domain

SW2 achieved good reliability ($\alpha = .80$) but showed substantial construct heterogeneity, with only 31% of items showing clear factor loadings. The domain currently combines interpersonal closeness, social comfort, belonging, jealousy management, and responses to criticism, elements that the factor analysis suggests do not form a unitary construct. Future revisions should consider decomposing SW2 into two focused subdomains: Interpersonal Connection (closeness, meaningful relationships) and Social Emotional Functioning (social comfort, managing reactions). Items assessing belonging and felt acceptance, which clustered empirically with safety items, may fit better within safety-related measurement, where they reflect felt security as a precondition for belonging.

5.3.4. Identity Development: Honoring Intrapsychic and Relational Components

SW1 items split across two factors—intrapsychic self-worth (“I like who I am”) and relational trust/support (“I feel valued by others”), suggesting these are related but distinguishable aspects of identity. This aligns with attachment theory’s distinction between internal working models of self and models of others (Bowlby, 1979). Maintain SW1 as a single domain, but acknowledge two subfacets that can be scored separately when clinically useful. This preserves the theoretical unity of “Core Identity & Self-Worth” while recognizing that autonomous self-valuing and felt relational worth may require different intervention approaches. A client might demonstrate strong self-worth while struggling to trust others’ support, or vice versa.

5.3.5. Item 8 and Spirituality: A Measurement Lesson

The substantial negative correlation between Item 8 (“I believe in a power much greater than myself”) and Core Identity ($r = -0.624$) underscores an important measurement principle: items must directly assess intended constructs, not peripheral correlates. For some respondents, belief in a higher power emphasized limitation rather than empowerment; for others, theistic language did not resonate with their spiritual framework. If spirituality matters to the IWBM framework, measure it by its function (providing meaning, purpose, connection) rather than by specific beliefs. Revised items might assess: “My spiritual or philosophical beliefs give me a sense of purpose” or “I draw strength from something larger than myself.” This approach captures spirituality’s role in vitality, regardless of belief content.

5.4. Demographic Patterns: Appropriate Interpretation and Clinical Implications

5.4.1. Sex Differences in Will: Socialization, Not Bias

Males scored higher on Will domains ($d = 0.30-0.39$), likely reflecting socialization patterns—persistence, grit, and determination are often emphasized in masculine contexts—rather than scale bias. Importantly, Spirit domains showed no sex differences, suggesting resilience develops differently across dimensions based on socialization experiences. Practitioners should attend to how gender socialization may create barriers to Will development (e.g., girls receiving less encouragement for persistence), while recognizing that observed differences do not mean males should score higher—they suggest interventions may need to address socially constructed obstacles to support authentic development.

5.4.2. Race/Ethnicity and Age Patterns: Contextualizing Significant Effects

Two small but statistically significant race/ethnicity effects emerged for SW5, Cognitive Mindset for Growth, and the Will composite ($\eta^2 = .062-.067$). Closer examination shows these effects were driven primarily by lower scores among participants who preferred not to disclose their race/ethnicity ($n = 3-4$), a group too small to interpret meaningfully. The three largest racial/ethnic groups (White, Latino/Hispanic, and Black) showed closely comparable means across all domains, providing encouraging preliminary evidence of cross-group consistency. These findings should not be interpreted as indicating systematic racial/ethnic biases in the scale; rather, they underscore the need for future validation with larger, more balanced samples and formal measurement invariance testing (Putnick & Bornstein, 2016). No significant age differences were found across any domain or composite (all p s > 0.10 , all $\eta^2 < 0.037$). The descriptive trend of slightly higher scores in older adults, particularly on Will-related domains, aligns with lifespan developmental research showing continued growth in goal clarity and executive function across adulthood (Carstensen et al., 2011) while core identity and relational functioning remain relatively stable once established.

5.5. Current Applications: Using the Scale Responsibly and Effectively

In its current form, the Spirit & Will Scale has valuable research and clinical applications when used appropriately. For research purposes, the scale is well-suited for group-level studies examining Spirit/Will correlates with other psychological constructs, program evaluations measuring pre-post changes in resilience following interventions, and exploratory studies investigating relationships among domains. Researchers should prioritize Total and composite scores (Spirit, Will), which show the strongest reliability, while interpreting domain scores with caution, given the mixed evidence for validity. In clinical settings, the scale serves as an initial assessment tool to identify potential strengths and areas for growth for therapeutic discussion. It serves as a valuable conversation starter, helping clients articulate their experiences of identity, connection, regulation, and growth. The scale's strongest current clinical application is progress monitoring—tracking change over time provides meaningful feedback on treatment effectiveness and areas needing additional focus. Clinicians can also use the scale in team discussions during case conceptualization, recognizing scores as one data point among many sources of clinical information. For self-reflection and psychoeducation, individuals can complete the scale to gain insight and set goals. In educational settings, the scale can stimulate discussion of resilience dimensions and normalize conversations about psychological strengths and challenges. Do not use the scale for high-stakes decisions or as the sole basis for clinical formulation. Scores are preliminary indicators that should be integrated with clinical interview, behavioral observation, and other assessment data. Practitioners should interpret patterns (relative strengths and growth areas) rather than precise scores, treating results as hypothesis-generating rather than definitive.

5.6. The Path Forward: Systematic Scale Refinement Based on Empirical Evidence

This exploratory validation provides clear, data-driven guidance for scale revision. Five immediate priorities emerge from the findings. First, we will realign Will as a unified construct by merging SW4 and SW5 items into a single Will domain with two facets (Affective Resilience and Executive Regulation). This revision will select 10-12 items that balance emotional and cognitive content, ensuring they directly assess persistence/hope and goal-directed planning/focus rather than redundantly measuring the same underlying capacity. Second, we will decompose and refine Social Connection (SW2) by creating two subdomains: Interpersonal Connection (5-6 items) and Social-Emotional Functioning (5-6 items). This process will remove items with weak correlations (Items 19, 21,

22, 23, 24) and add positively worded items assessing relationship quality and social confidence. The revision will sharpen the distinction between intimate relationship capacity and broader social-emotional competence. Third, we will strengthen Emotional Regulation (SW3) by reducing reverse-scored items (currently 5 of 8) and adding positively worded items that assess emotion regulation skills, such as mindfulness, reappraisal, and emotional flexibility. This revision will also clarify whether the domain measures acceptance of emotions, regulation of expression, or both processes. Fourth, we will maintain and refine Core Identity (SW1) by preserving it as a single domain while acknowledging two distinguishable subfacets (intrapsychic worth and relational trust). The revision will report subfacet scores when clinical differentiation is functional and ensure balanced coverage of autonomous self-valuing and relational security items. Fifth, we will integrate safety conceptually by maintaining safety items within SW2 (social safety) and SW4 (emotional safety to take risks) and adding 2-3 items to each domain that explicitly assess felt safety in those contexts. Practitioners will be trained to interpret low safety scores as potentially explaining difficulties across multiple domains, consistent with trauma-informed frameworks. The goal is a refined scale with clearer construct boundaries, reduced cross-loading, and balanced coverage across domains.

6. STUDY LIMITATIONS AND INTERPRETIVE CAUTIONS

Several limitations qualify these conclusions. The sample was predominantly White (71.5%), female (60.4%), and young adult (63.8% ages 18-35), recruited through convenience sampling in the southeastern US. This limits generalizability to men, older adults, racial/ethnic minorities, and clinical populations. Small cell sizes for some groups prevented robust subgroup analyses. Methodologically, the cross-sectional design precludes examination of test-retest reliability or sensitivity to change. Single-method self-report assessment introduces method variance and social desirability bias. Most importantly, the study included no external criterion measures, preventing examination of convergent, discriminant, and predictive validity. All validity evidence derives from within-scale relationships. The exploratory factor analysis capitalizes on sample-specific variance and requires confirmation in independent samples. Listwise deletion for EFA (24% data loss) may have introduced bias. Although it met minimum guidelines, the sample size ($N=207$, 155 complete cases) is modest for stable factor solutions with 60 items. Despite these limitations, the study provides valuable exploratory evidence that establishes a foundation for systematic refinement and subsequent confirmatory validation.

7. CONCLUSION

This exploratory validation shows that the IWBMC Spirit & Will Assessment Scale successfully operationalizes core resilience constructs, with strong overall reliability ($\alpha = .95$) and a clear structure across well-defined domains. The divergence between theoretical and empirical structures—rather than indicating fundamental problems—provides actionable guidance for construct realignment. Key findings support: (1) reconceptualizing Will as a unified construct with affective and executive facets, (2) decomposing Social Connection into distinct subdomains, (3) integrating safety as a cross-cutting foundation rather than a separate domain, and (4) maintaining Core Identity while acknowledging intrapsychic and relational components. In its current form, the scale serves valuable purposes for research, clinical screening, and progress monitoring when used appropriately. With systematic refinement based on these empirical findings—particularly realigning domain boundaries to match factor structure—the Spirit & Will Scale can become a robust tool supporting both IWBMC research and practice. The path forward is clear. Targeted revision of problematic domains, confirmatory validation in larger samples, and external validation are needed to establish clinical utility. These results initiate that process by identifying exactly what needs to change and why.

REFERENCES

- American Educational Research Association, A. P. A., & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Barcaccia, B., Couyoumdjian, A., Di Consiglio, M., Papa, C., Cancellieri, U. G., & Cervin, M. (2023). Purpose in life as an asset for well-being and a protective factor against depression in adolescents. *Frontiers in Psychology*, 14, 1250279. <https://doi.org/10.3389/fpsyg.2023.1250279>
- Baumeister, R. F., & Vohs, K. D. (2004). *Self-regulation and self-control*. In C. Peterson & M. E. P. Seligman (Eds.), *Character strengths and virtues: A handbook and classification*. New York: Oxford University Press.
- Bonanno, G. A., Westphal, M., & Mancini, A. D. (2011). Resilience to loss and potential trauma. *Annual Review of Clinical Psychology*, 7(1), 511-535. <https://doi.org/10.1146/annurev-clinpsy-032210-104526>
- Boreham, I. D., & Schutte, N. S. (2023). The relationship between purpose in life and depression and anxiety: A meta-analysis. *Journal of Clinical Psychology*, 79(12), 2736-2767. <https://doi.org/10.1002/jclp.23576>
- Bowlby, J. (1979). The Bowlby-Ainsworth attachment theory. *Behavioral and Brain Sciences*, 2(4), 637-638. <https://doi.org/10.1017/S0140525X00064955>
- Carstensen, L. L., Turan, B., Scheibe, S., Ram, N., Ersner-Hershfield, H., Samanez-Larkin, G. R., . . . Nesselroade, J. R. (2011). Emotional experience improves with age: Evidence based on over 10 years of experience sampling. *Psychology and Aging*, 26(1), 21. <https://psycnet.apa.org/doi/10.1037/a0021285>
- Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, 31(12), 1412. <https://psycnet.apa.org/doi/10.1037/pas0000626>
- Colla, R., Williams, P., Oades, L. G., & Camacho-Morles, J. (2022). A New Hope” for positive psychology: A dynamic systems reconceptualization of hope theory. *Frontiers in Psychology*, 13, 809053. <https://doi.org/10.3389/fpsyg.2022.809053>
- Enders, C. K. (2022). *Applied missing data analysis* (2nd ed.). New York: The Guilford Press.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, 38, 69-119. [https://doi.org/10.1016/S0065-2601\(06\)38002-1](https://doi.org/10.1016/S0065-2601(06)38002-1)
- Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3), 238-247. <https://doi.org/10.1037/1040-3590.7.3.238>
- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1(1), 104-121. <https://doi.org/10.1177/109442819800100106>
- Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement*, 65(2), 202-226. <https://doi.org/10.1177/0013164404267287>
- Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Medicine*, 7(7), e1000316. <https://doi.org/10.1371/journal.pmed.1000316>
- Holt-Lunstad, J. (2024). Social connection as a critical factor for mental and physical health: Evidence, trends, challenges, and future implications. *World Psychiatry*, 23(3), 312-332. <https://doi.org/10.1002/wps.21224>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185. <https://doi.org/10.1007/BF02289447>
- Kim, E. S., Chen, Y., Nakamura, J. S., Ryff, C. D., & VanderWeele, T. J. (2022). Sense of purpose in life and subsequent physical, behavioral, and psychosocial health: An outcome-wide approach. *American Journal of Health Promotion*, 36(1), 137-147. <https://doi.org/10.1177/08901171211038545>

- Kroll, E., Lederman, M., Kohlmeier, J., Kumar, K., Ballard, J., Zant, I., & Fenkel, C. (2024). The positive impact of identity-affirming mental health treatment for neurodivergent individuals. *Frontiers in Psychology*, *15*, 1403129. <https://doi.org/10.3389/fpsyg.2024.1403129>
- Lebrón-Cruz, A., & Orvell, A. (2023). I am what I am: The role of essentialist beliefs and neurodivergent identification on individuals' self-efficacy. *Journal of Experimental Psychology: General*, *152*(11), 2995. <https://psycnet.apa.org/doi/10.1037/xge0001457>
- Logan, A. C., Berman, B. M., & Prescott, S. L. (2023). Vitality revisited: The evolving concept of flourishing and its relevance to personal and public health. *International Journal of Environmental Research and Public Health*, *20*(6), 5065. <https://doi.org/10.3390/ijerph20065065>
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84. <https://psycnet.apa.org/doi/10.1037/1082-989X.4.1.84>
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741. <https://psycnet.apa.org/doi/10.1037/0003-066X.50.9.741>
- Milton, D. E. (2012). On the ontological status of autism: The 'double empathy problem'. *Disability & Society*, *27*(6), 883-887. <https://doi.org/10.1080/09687599.2012.710008>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- O'Dell, A., Saunders, A., York, A., Haynes, H. A., & Gonzalez, J. (2025). The interpersonal whole brain model of care™: Utilizing the union of science & love to meet the needs of neurodivergent populations. *Innovation in Science and Technology*, *4*(3), 11-18. <https://doi.org/10.56397/IST.2025.03.02>
- Porges, S. W. (2011). *The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication, and self-regulation*. New York: W. W. Norton & Company.
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71-90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Quirin, M., Jais, M., Di Domenico, S. I., Kuhl, J., & Ryan, R. M. (2021). Effortless willpower? The integrative self and self-determined goal pursuit. *Frontiers in Psychology*, *12*, 653458. <https://doi.org/10.3389/fpsyg.2021.653458>
- Raimondi, G., Zacchilli, M., Frederick, C. M., Alivernini, F., Manganelli, S., Cavicchiolo, E., . . . Dawe, J. (2025). Measuring vitality and depletion during adolescence: Validation of the subjective vitality/subjective depletion scale in a sample of Italian students. *Pediatric Reports*, *17*(5), 98. <https://doi.org/10.3390/pediatric17050098>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68. <https://psycnet.apa.org/doi/10.1037/0003-066X.55.1.68>
- Stefana, A., Damiani, S., Granziol, U., Provenzani, U., Solmi, M., Youngstrom, E. A., & Fusar-Poli, P. (2025). Psychological, psychiatric, and behavioral sciences measurement scales: Best practice guidelines for their development and validation. *Frontiers in Psychology*, *15*, 1494261. <https://doi.org/10.3389/fpsyg.2024.1494261>
- Van der Kolk, B. A. (2014). *The body keeps the score: Brain, mind, and body in the healing of trauma*. New York: Viking.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*(6), 806-838. <https://doi.org/10.1177/0011000006288127>

Online Science Publishing is not responsible or answerable for any loss, damage or liability, etc. caused in relation to/arising out of the use of the content. Any queries should be directed to the corresponding author of the article.